



**HAL**  
open science

# Evolving genealogies for branching populations under selection and competition

Airam Blancas, Stephan Gufler, Sandra Kliem, Viet-Chi Tran, Anton Wakolbinger

► **To cite this version:**

Airam Blancas, Stephan Gufler, Sandra Kliem, Viet-Chi Tran, Anton Wakolbinger. Evolving genealogies for branching populations under selection and competition. 2021. hal-03275257v1

**HAL Id: hal-03275257**

**<https://hal.science/hal-03275257v1>**

Preprint submitted on 30 Jun 2021 (v1), last revised 9 Nov 2022 (v2)

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

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

# Evolving genealogies for branching populations under selection and competition

Airam Blancas\*, Stephan Gufler†, Sandra Kliem‡  
Viet Chi Tran§ and Anton Wakolbinger†

June 30, 2021

## Abstract

For a continuous state branching process with two types of individuals which are subject to selection and density dependent competition, we characterize the joint evolution of population size, type configurations and genealogies as the unique strong solution of a system of SDE's. Our construction is achieved in the lookdown framework and provides a synthesis as well as a generalization of cases considered separately in two seminal papers by Donnelly and Kurtz (1999), namely fluctuating population sizes under neutrality, and selection with constant population size. As a conceptual core in our approach, we introduce the *selective lookdown space* which is obtained from its neutral counterpart through a state-dependent thinning of “potential” selection/competition events whose rates interact with the evolution of the type densities. The updates of the genealogical distance matrix at the “active” selection/competition events are obtained through an appropriate sampling from the selective lookdown space. The solution of the above mentioned system of SDE's is then mapped into the joint evolution of population size and symmetrized type configurations and genealogies, i.e. marked distance matrix distributions. By means of Kurtz's Markov mapping theorem, we characterize the latter process as the unique solution of a martingale problem. For the sake of transparency we restrict the main part of our presentation to a prototypical example with two types, which contains the essential features. In the final section we outline an extension to processes with multiple types including mutation.

**Keywords.** Birth-death particle system, lookdown process, tree-valued processes, selection, density-dependent competition, selective lookdown space, fluctuating population size, genealogy.

**MSC2010.** 60J80, 60K35, 92D10.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Model and main results</b>	<b>4</b>
2.1	Population size and type frequencies . . . . .	4
2.2	Genealogies . . . . .	5
2.3	Lookdown representation of the joint process of population size, type frequencies, and genealogies . .	6

---

\*Department of Statistics, Instituto Tecnológico Autónomo de México, México; E-mail: [airam.blancas@itam.mx](mailto:airam.blancas@itam.mx)

†Goethe Universität, Institut für Mathematik, 60629 Frankfurt am Main, Germany; E-mail: [gufler@math.uni-frankfurt.de](mailto:gufler@math.uni-frankfurt.de), [wakolbinger@math.uni-frankfurt.de](mailto:wakolbinger@math.uni-frankfurt.de),

‡Universität Leipzig, Mathematisches Institut, Augustusplatz 10, 04109 Leipzig; E-mail: [kliem@math.uni-leipzig.de](mailto:kliem@math.uni-leipzig.de)

§LAMA, Univ Gustave Eiffel, Univ Paris Est Creteil, CNRS, F-77454 Marne-la-Vallée, France; E-mail: [chi.tran@univ-eiffel.fr](mailto:chi.tran@univ-eiffel.fr)

2.3.1	Type configuration and type frequencies . . . . .	6
2.3.2	Pathwise construction of the lockdown process . . . . .	7
2.3.3	Filling in the ancestry: from the neutral to the selective lockdown space . . . . .	10
2.3.4	Updating the distance matrix at neutral and selective reproduction events . . . . .	12
2.3.5	A well-posed martingale problem for the evolving lockdown genealogy . . . . .	13
2.4	A well-posed martingale problem for the evolving symmetrized genealogy . . . . .	15
<b>3</b>	<b>Building blocks from neutrality</b>	<b>16</b>
3.1	Neutral lockdown space and marked sampling measures . . . . .	16
3.2	Partitioning the lockdown space: roots and fragments . . . . .	17
<b>4</b>	<b>An SDE for type configuration and population size: proof of Theorem 2.2</b>	<b>20</b>
<b>5</b>	<b>From the neutral to the selective genealogy</b>	<b>26</b>
5.1	The selective lockdown genealogy . . . . .	27
5.2	Two well-posed martingale problems in the lockdown framework . . . . .	29
5.3	The symmetrized selective genealogy. Proof of Theorem 2.7 . . . . .	32
5.4	Proof of Propositions 2.9, 2.10 and 2.1 . . . . .	33
<b>6</b>	<b>Outlook: An extension to multiple types and mutations</b>	<b>36</b>

# 1 Introduction

The aim of our paper is to give a pathwise construction for the joint evolution of population size, type frequencies and genealogies in a continuous state branching process with interactions due to type dependent selective advantage in reproduction and type density dependent competition. Such processes model large populations whose individuals are distinguished by their types. The sizes of the populations and their type structures are fluctuating due to individual births and deaths, where certain types may have a selective advantage in the fecundity, and others may have a disadvantage against some other types, say in the competition for resources. We are interested here in these dynamics but also in that of the genealogies of the individuals composing these populations, which consist of the collection of their ancestral paths, i.e. the succession of their ancestors with their types. We demonstrate the strength of the approach in a prototypical example with two types, one of them having a selective advantage, the other one having a competitive disadvantage. This restriction is mainly for presentational reasons; in Section 6 we will outline an extension to more general processes with multiple types, including mutations.

We take the so-called lockdown approach that has been developed by Donnelly and Kurtz in order to construct and study the evolution of continuum populations with a general type space in terms of a countably infinite particle system. In two seminal papers, these authors treated two distinct cases: that of populations with constant sizes under selection (and recombination) [12] and that of neutral populations with fluctuating population sizes [13]. In the present work we consider selection and competition combined with fluctuating sizes. One of our key results, Theorem 2.2, extends ideas in the proof of [12, Theorem 4.1] to a situation where the total mass is a stochastic process whose dynamics depends on the type frequencies, and thus opens the way for a synthesis of the settings of [12] and [13]. In both of these papers the evolution of the (relative) type frequencies and the genealogies are encoded in an infinite particle system that describes the reproductive events. In [12] the population size (or total mass of the continuum population) is assumed to be constant, while in [13] it is accounted for in a separate process, which is autonomous due to the neutral setting considered in that paper. This is no longer the case in our setting where additional births and deaths occur in the infinite particle system due to selection and competition which depend on and also impact the evolution of the population size.

While many considerations pertaining to genealogies and ancestral lineages are already present in and between the lines of [12] and [13], the power of the lockdown approach for studying evolving genealogies has unfolded only more recently, several years after Evans [18] characterized Kingman’s coalescent as a random metric space. The lockdown representation of the evolving populations in terms of exchangeable particle systems comes with a graphical representation that provides a genealogy in a natural way. A central tool for proving Theorem 2.2 are the *sampling measures* on the *neutral lockdown space*, which is the completion of  $\mathbb{R} \times \mathbb{N}$  with respect to the (random) semi-metric given by the (neutral) genealogical distances. The concept of the (neutral) lockdown space has recently been introduced in [21] to obtain (in the neutral case and for constant population size) a pathwise construction of tree-valued Fleming-Viot processes. In Section 5 we will construct what we call the *selective lockdown space with fluctuating population size*. This space will carry the sampling measures which will serve to update the type configuration as well as the genealogical distances at selective events, see Sec. 2.3.4.

Theorem 2.5 then establishes a system of SDE’s (with unique strong solution) for the joint process of total mass, genealogical distance matrix, and type configuration. Exploiting the exchangeability that comes with the concept of sampling, we then turn to the *symmetrization* or “unlabelling” of the lockdown genealogies. As states that describe the type distributions and genealogies, we use here the isomorphy classes of marked ultrametric measure spaces introduced in [19, 9, 10] which can be thought of as marked distance matrix distributions. Our Theorem 2.7 then characterizes the joint evolution of genealogies and population size in terms of a well-posed martingale problem. This result is proved by a two-fold application of Kurtz’s Markov mapping theorem, see Sections 5.2 and 5.3.

While we provide a (and in a weak sense *the*) solution of this martingale problem from specified sources of randomness (the Brownian motion  $\mathcal{W}$  and Poisson point processes  $\mathcal{L}$  and  $\mathcal{K}$  defined in Section 2), a more common approach for showing the existence of a solution is to deal with tightness of finite approximations. This may sometimes be tedious, and cause serious technical problems. One of the few papers in which a martingale problem for continuum tree-valued processes including pairwise (competitive) individual interactions (and fluctuating total mass) has been treated in this vein is [27]. Note however that there the distances between individuals are measured in terms of numbers of mutations, whereas we measure distances in terms of times back to the most recent common ancestor of the two individuals. In [10] tree-valued Fleming-Viot processes with mutation and type-frequency dependent selection are constructed, but there constant population size is assumed. Our contribution here is to also bring a lockdown representation of the genealogies that completes the picture: the genealogies “locally” look like the neutral genealogies in absence of selection and competition with modifications related to the selective and competitive events.

While we are using the Poisson process  $\mathcal{L}$  of *lockdown events* (see Section 2) to encode the elements of neutral genealogies, there exist also alternative routes for doing this. One of them is along the *continuum random tree* and Brownian excursions ([1, 2, 31] or [20, Ch. 4]), with certain deformations of these objects to model competition ([3, 30, 34]) although the introduction of types is not straightforward in these models and in the cited references the competition depends on the individuals’ left-right order encoded in the excursion. Another one is Kurtz and Rodrigues’ *lockdown representation with a continuum of levels* [29], which has recently been extended by Etheridge and Kurtz [15] to a variety of models including selection and competition, but with less emphasis on evolving genealogies.

Recent work on evolving genealogies in the neutral case, with a focus on heavy-tailed offspring distributions, has been reviewed in [24]. Evolving ancestral path configurations under competition are studied in [33, 25] or [7], building on the framework of *historical processes* which in the non-interactive case was pioneered in [8, 14]. Inference methods in the presence of selection, varying population size and evolving population structure are described in [32], extending results of [5]; in the latter models, a time-scale separation allows to treat separately the type structure and population size on the one hand, and the genealogies on the other hand. In the present paper, however, we deal simultaneously with interactions,

demography and genealogies.

## 2 Model and main results

### 2.1 Population size and type frequencies

In order to make the conceptual novelties and essentials as transparent as possible, we will restrict ourselves in the main part of this work to a population with only two types,  $A$  and  $B$ . An extension of the results to more general type spaces and including mutations is outlined in Section 6. We will denote by  $\mathbb{I} = \{A, B\}$  the type space, and by  $\xi^A$  and  $\xi^B$  the processes in continuous time corresponding to the sizes of the type  $A$ - and type  $B$ -populations. Intuitively, these populations consist of a continuum of individuals with infinitesimal masses; the concept of *sampling measures*, which we will recover also in the *selective lockdown space*, makes this intuition rigorous.

The *population size* or *total population mass* at a time  $t > 0$  is  $\xi_t = \xi_t^A + \xi_t^B$ , and for  $\xi_t > 0$  we define the *type frequencies* or proportions of types  $A$  and  $B$  as

$$\mu_t^A = \xi_t^A / \xi_t; \quad \mu_t^B = \xi_t^B / \xi_t = 1 - \mu_t^A.$$

The system we are going to consider as a prototypical case is a two-type Feller branching diffusion with interactions

$$\begin{aligned} d\xi_t^A &= b\xi_t^A dt - c\xi_t^A \xi_t^B dt + \sqrt{\xi_t^A} dW_t^A \\ d\xi_t^B &= -c\xi_t^B \xi_t^A dt + \sqrt{\xi_t^B} dW_t^B, \end{aligned} \tag{1}$$

where  $W^A$  and  $W^B$  are independent standard Brownian motions. The processes  $W^A$  and  $W^B$  drive the fluctuations due to natural births and deaths in the diffusion limit of branching populations. The nonnegative constant  $b$  is the coefficient of the intensity of additional births of type  $A$ -individuals due to their enhanced fecundity, whereas  $c \geq 0$  is the intensity of additional deaths of individuals due to their competition against all individuals of the opposite type. Such a system of equations can be seen as arising from the limit of finite particle systems. In [17, Chap. 9, Section 2 p. 392] this is proved in the case  $c = 0$ ; then both  $\xi^A$  and  $\xi^B$  are independent Feller diffusions.

Existence and uniqueness of a strong solution of (1) is guaranteed by the following

**Proposition 2.1.** Given initial conditions  $(\xi_0^A, \xi_0^B)$  such that  $\mathbf{E}((\xi_0^A)^{2+\epsilon} + (\xi_0^B)^{2+\epsilon}) < +\infty$  for some  $\epsilon > 0$ , and given independent standard Brownian motions  $W^A$  and  $W^B$ , there exists a unique strong solution  $(\xi^A, \xi^B)$  to (1).

For the sake of self-containedness we will include a proof of this proposition at the end of Section 5. This will also reveal the long time behavior of the process  $(\xi^A, \xi^B)$  using comparisons with (possibly drifted) Feller diffusions: the process  $\xi^B$  gets extinct in finite time almost surely, while  $\xi^A$  is either trapped in 0 or diverges to  $+\infty$ .

As announced in the Introduction, our main goal is the characterization of *evolving marked genealogies* that underlie the system (1). It turns out that an accessible way to this goal leads via the *total mass process*  $\xi = \xi^A + \xi^B$ . A stochastic differential equation (SDE) for  $\xi$  and the type proportions  $\mu_t^A, \mu_t^B$  is obtained by adding the two equations in (1):

$$d\xi_t = (b\mu_t^A \xi_t - 2c\mu_t^A \mu_t^B \xi_t^2)dt + \sqrt{\xi_t} dW_t, \tag{2}$$

where  $W$  is a standard Brownian motion. Clearly, equation (2) is not autonomous. In addition to  $W$  which takes care of the fluctuations of the population size in the interplay with the current type proportions, the

other drivers of the evolving marked genealogy that will trigger the (neutral and selective) reproductive events will be Poisson point processes that come up in the lockdown framework described in Section 2.3. Closing the circle, Proposition 2.10 will then guarantee that the pair  $(\xi^A, \xi^B)$  can be restored from the total mass process  $\xi$  together with the evolving marked genealogy  $Y$ , thus rendering a weak solution of (1).

The following time change will be instrumental (see also [13]):

$$t \mapsto s = s(t) := \int_0^t \frac{1}{\xi_v} dv, \quad \zeta_s = \xi_t. \quad (3)$$

For reasons explained in Section 2.3, the timescale  $s$  will be called the *lockdown timescale*. In Section 2.3, we will provide a system of stochastic differential equations that describes, in this lockdown timescale, a population size process  $(\zeta_s)$  together with an evolving type configuration  $(G_s)$  to which we will be able to associate a process of type frequencies  $(\mu^{G_s})$  (see Theorem 2.2). As a corollary, transforming back to the timescale  $t$  via (3), the resulting process  $(\zeta_{s(t)}\mu^{G_{s(t)}}\{A\}, \zeta_{s(t)}\mu^{G_{s(t)}}\{B\})$  will provide a weak solution of (1), see Proposition 2.10.

In the neutral case ( $b = c = 0$ ),  $(\xi_t)$  is a standard Feller diffusion, and the process  $(\mu_t^A)$  after the time change (3) turns into a standard Wright-Fisher diffusion (e.g. [23, Chapter IV.8]). The correspondence  $(\xi_t^A, \xi_t^B) \leftrightarrow ((\xi_t), (\mu_t^A, \mu_t^B))$  is thus an interactive counterpart of Perkins' desintegration of super-Brownian motion into a Feller branching diffusion and a time-changed Fleming-Viot process (see [16] p. 83, [35]).

## 2.2 Genealogies

The *marked genealogy* of the continuum population at some fixed time is described by the joint distribution of pairwise genealogical distances and types of a sequence of individuals that is drawn i.i.d. according to a prescribed sampling measure. In order to formalize this, and to define the space of marked genealogies, we recall a couple of concepts. In our context, genealogical distances of contemporaneous individuals are described by a semi-ultrametric, i.e. a semi-metric  $d$  that satisfies the strong triangle inequality  $\max\{d(x, y), d(y, z)\} \geq d(x, z)$ . The prefix *semi* means that  $d(x, y) = 0$  does not imply  $x = y$ , corresponding to the fact that at the time of a reproduction event, the “mother” and her “daughter” have genealogical distance 0, while being considered as different individuals.

Marked metric measure spaces have been introduced by Depperschmidt, Greven, and Pfaffelhuber [9]. An  $\mathbb{I}$ -marked ultrametric measure space is a triple  $(\tau, d, m)$  where  $(\tau, d)$  is a complete, separable ultrametric space and  $m$  is a probability measure on the Borel sigma algebra on the product space  $\tau \times \mathbb{I}$ . In our context, such spaces  $(\tau, d)$  will arise as completions of semi-ultrametric spaces, after first identifying elements of distance zero, see Definition 2.3 a) in Sec. 2.3.3.

The marked distance matrix distribution of an  $\mathbb{I}$ -marked ultrametric measure space  $(\tau, d, m)$  is defined as the distribution of  $((d(V_i, V_j))_{i, j \in \mathbb{N}}, (H_i)_{i \in \mathbb{N}})$  where  $(V_i, H_i)_{i \in \mathbb{N}}$  is a sequence in  $\tau \times \mathbb{I}$ , i.i.d. with distribution  $m$ . (Here and below,  $\mathbb{N} = \{1, 2, \dots\}$  denotes the set of natural numbers.) Marked ultrametric measure spaces with the same marked distance matrix distribution are called isomorphic.

The space of isomorphy classes of  $\mathbb{I}$ -marked ultrametric measure spaces will be denoted by  $\mathbb{M}$ , and will be called the *space of marked genealogies*. This space  $\mathbb{M}$ , equipped with the marked Gromov-weak topology in which elements of  $\mathbb{M}$  converge if and only if the associated marked distance matrix distributions converge, is Polish by [9]. In Theorem 2.7 we will characterize an  $(0, \infty) \times \mathbb{M}$ -valued process  $(\xi, Y)$  by a stopped martingale problem (in the sense of [17, Ch. 4.6]). The first component of this process will describe the population size, and will give a weak solution of (2). The second component will describe the marked genealogy, with the type frequencies being a measurable function of the latter.

## 2.3 Lookdown representation of the joint process of population size, type frequencies, and genealogies

We are going to provide a representation of the just mentioned process  $(\xi, Y)$  in terms of a process  $(\zeta, R, G)$ , where  $(\zeta_s, G_s)_{s \geq 0}$  will be the unique strong solution of a system of SDE's in the time scale (3), see Theorem 2.2. The process  $R$  will take its values in the semi-ultrametrics on  $\mathbb{N}$  (which we will address as *distance matrices* for short). The underlying graphical representation includes, in addition to a Brownian motion  $\mathcal{W}$ , a pair  $(\mathcal{L}, \mathcal{K})$  of a family of Poisson point processes (defined in Sec. 2.3.2). The triple  $(\mathcal{W}, \mathcal{L}, \mathcal{K})$  does not only drive the process  $(\zeta, G)$  in terms of an SDE (see Theorem 2.2), but also the process  $R$ , see (24).

We will deduce that  $(\zeta, R, G)$  solves of a well-posed martingale problem. This will be an essential ingredient for the proof of Theorem 2.7, which provides the characterization of  $(\xi, Y)$  in terms of a well-posed martingale problem.

Individuals living in the lockdown system at time  $s$  are coded by  $(s, i)$ ,  $i = 1, 2, \dots$  (As we will see from the constructions explained in Subsection 2.3.3, this is only a subset of the uncountably many individuals living at time  $s$ , namely the subset consisting of those individuals who have an offspring that survives for some positive amount of time.) The second component is called *level*; it labels the individuals alive at time  $s$  and having an offspring at some time strictly larger than  $s$ . The graphical construction will allow to reconstitute the ancestral paths of the individuals  $(s, i)$ . The evolution of the genealogical distances and the types of these individuals will be described by the process

$$X = (R, G) = ((R_s, G_s)_{s \geq 0}).$$

The second component  $G_s = (G_s(i))_{i \in \mathbb{N}} \in \mathbb{I}^{\mathbb{N}}$  of this process is the type configuration at time  $s$ . The first component  $R_s = (R_s(i, j))_{i, j \in \mathbb{N}}$  is a random semi-ultrametric on  $\mathbb{N}$  that describes the genealogical distances between the individuals at time  $s$  in the time scale of the interactive branching system (1). That is, if the most recent common ancestor of  $(s, i)$  and  $(s, j)$  lived at time  $s' < s$ , then  $R_s(i, j) = 2(t(s) - t(s'))$ , with

$$t(s) := \int_0^s \zeta_u du \quad (4)$$

being the inverse of the time change (3).

We think of our initial value  $(R_0, G_0)$  as the genealogical distances and the types of a sequence of individuals that are drawn independently at random from an infinite population at time 0. More precisely, we always assume that  $(R_0, G_0)$  is distributed according to the marked distance matrix distribution of some  $\mathbb{I}$ -marked ultrametric measure space (as defined in Sec. 2.2).

Obviously this implies that the pair  $(R_0, G_0)$  is *exchangeable* in the sense that for all  $n \in \mathbb{N}$  and all permutations  $\pi$  of  $[n] = \{1, \dots, n\}$  one has

$$((R_0(i, j))_{1 \leq i, j \leq n}, (G_0(i))_{1 \leq i \leq n}) \stackrel{d}{=} ((R_0(\pi(i), \pi(j)))_{1 \leq i, j \leq n}, (G_0(\pi(i)))_{1 \leq i \leq n}). \quad (5)$$

Conversely, a version of the Gromov-Vershik representation theorem ([22, Corollary 3.12]) ensures that each  $(R_0, G_0)$  obeying (5) can be realized as the second step in a two-stage experiment, whose first step is the random choice of (an isomorphy class of) a marked ultrametric measure space (or equivalently of a marked distance matrix distribution), and whose second step is the marked distance matrix that arises by an i.i.d. drawing from that marked ultrametric measure space.

### 2.3.1 Type configuration and type frequencies

The process  $X$  provides “microscopic” information on the type configuration and genealogies of the individuals in the lockdown system. The fluctuations of the population mass obtained from (2) and the

time change (3) deal with “macroscopic” quantities and are not seen directly in the lockdown representation. However both scales are coupled: we will see that the type frequencies arise from the microscopic (i.e. individual-based) type configurations and appear in the coefficients of the SDE (2) whose solution in turn will impact the local dynamics of the lockdown levels.

For a type configuration  $g \in \mathbb{I}^{\mathbb{N}}$  we will say that  $g$  *admits type frequencies* if the limiting measure

$$\mu^g := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \delta_{g(i)} \quad (6)$$

exists in the weak topology on  $M^1(\mathbb{I})$ , the space of probability measures on  $\mathbb{I}$  (which in our case with two traits simply means that  $\mu^g\{A\} := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{g(i)=A\}}$  exists). We will then call  $\mu^g$  the *type distribution belonging to  $g$* .

We will construct the type process  $G$  in such a way that it a.s. admits type frequencies at every time  $s$ , hence allowing to read off the proportion  $\mu^{G_s}\{A\}$  of type  $A$  at time  $s$  from the configuration  $(G_s(i))_{i \in \mathbb{N}}$ .

These proportions will play a role in the dynamics of genealogies and type configurations (see (9), (10) and (11) below), and also in the SDE (2) for the total mass process, which in view of the time change (3), becomes:

$$d\zeta_s = (b\mu^{G_s}\{A\}\zeta_s^2 - 2c\mu^{G_s}\{A\}\mu^{G_s}\{B\}\zeta_s^3) ds + \zeta_s d\mathcal{W}_s, \quad (7)$$

where  $\mathcal{W}$  is a standard Brownian motion. Similarly, (1) becomes

$$\begin{aligned} d\zeta_s^A &= b\mu^{G_s}\{A\}\zeta_s^2 ds - c\mu^{G_s}\{A\}\mu^{G_s}\{B\}\zeta_s^3 ds + \sqrt{\zeta_s^A \zeta_s} d\mathcal{W}_s^A \\ d\zeta_s^B &= -c\mu^{G_s}\{B\}\mu^{G_s}\{A\}\zeta_s^3 ds + \sqrt{\zeta_s^B \zeta_s} d\mathcal{W}_s^B, \end{aligned} \quad (8)$$

where  $\mathcal{W}^A$  and  $\mathcal{W}^B$  are independent standard Brownian motions. Possible explosion or extinction events are treated at the beginning of Section 2.3.5.

### 2.3.2 Pathwise construction of the lockdown process

The construction of the process  $(\zeta_s, G_s)_{s \geq 0}$  is achieved via the so called ‘lookdown’ graphical construction. The ingredients are

- (I1) a standard Brownian motion  $\mathcal{W} = (\mathcal{W}_s)_{s \geq 0}$ ,
- (I2) a family  $(\mathcal{L}_{ij})_{i,j \in \mathbb{N}, i < j}$  of independent rate 1 Poisson point processes on  $\mathbb{R}_+$ ,
- (I3) a family  $(\mathcal{K}_i)_{i \in \mathbb{N}}$  of independent Poisson point processes on  $\mathbb{R}_+ \times \mathbb{R}_+ \times [0, 1] \times \{\beta, \delta\}$  whose intensity measure is the product of the Lebesgue measure on  $\mathbb{R}_+ \times \mathbb{R}_+ \times [0, 1]$  and of the counting measure on  $\{\beta, \delta\}$ ,

where the random elements in (I1), (I2) and (I3) are independent. The family  $(\mathcal{L}_{ij})$  can be superposed to a single Poisson point process  $\mathcal{L}$  on  $\mathbb{R}_+ \times \{(i, j) \in \mathbb{N}^2 : 1 \leq i < j < \infty\}$ , and the family  $(\mathcal{K}_i)$  can be superposed to a single Poisson point process  $\mathcal{K}$  on  $\mathbb{R}_+ \times \bigcup_{i \in \mathbb{N}} \{i\} \times \mathbb{R}_+ \times [0, 1] \times \{\beta, \delta\}$ . In this way  $\mathcal{L}_{ij}$  corresponds to the restriction of  $\mathcal{L}$  to  $\mathbb{R}_+ \times \{(i, j)\}$  and  $(\mathcal{K}_i)$  to the restriction of  $\mathcal{K}$  to  $\mathbb{R}_+ \times \{i\} \times \mathbb{R}_+ \times [0, 1] \times \{\beta, \delta\}$ . The Brownian motion  $\mathcal{W}$  drives the fluctuations of the population size,  $\mathcal{L}$  encodes the neutral birth events, and  $\mathcal{K}$  encodes the *potential* selective birth and death events affecting the levels in the graphical construction.

To each atom of  $\mathcal{L}_{ij}$  (for  $i < j$ ), say at time  $s$ , we associate an arrow starting from  $(s, i) \in \mathbb{R}_+ \times \mathbb{N}$  and directed to  $(s, j)$ . This arrow corresponds to a natural birth for the individual at level  $i$  at time  $s$ , placing an offspring of the same type at level  $j$ . Levels that were above  $j$  at time  $s-$  (*i.e.* levels  $k \in \mathbb{N}$  such that  $k > j$ ) are shifted up by 1. See Figure 1 and cf. also [36]. This lockdown process can be seen as the limit of a finite particle system as described in Donnelly and Kurtz [11] (see also [6]) where individuals with highest levels are removed at natural death events. That is why, heuristically, the natural death events are not seen any more on finite levels in the limit of infinitely many particles of small masses, as the highest level tends to infinity. It is now the varying population mass ( $\zeta_s$ ) that tracks the changing mass due to demographic events.

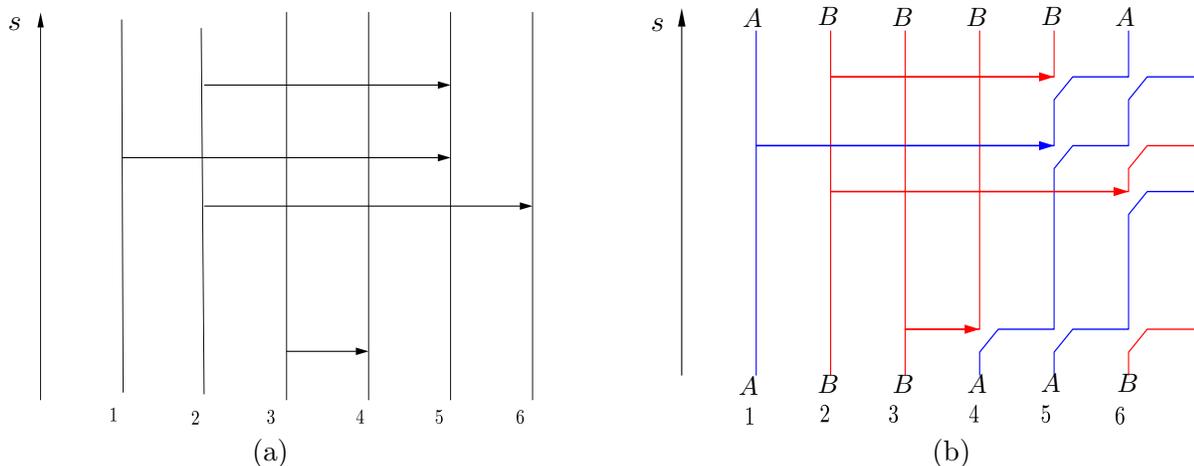


Figure 1: *Neutral genealogy in the lockdown representation: individuals correspond to levels (abscissa) and time is represented here along the ordinate axis. The neutral genealogy uses only atoms of the Poisson point process  $\mathcal{L}_{ij}$  (for  $i < j$ ). (a) The atoms of  $\mathcal{L}_{ij}$  can be associated with arrows from level  $i$  to level  $j$ . (b) Using the preceding arrows, we can define a ‘neutral’ genealogy: an arrow from  $i$  to  $j$  at time  $s$  corresponds to a natural birth for  $i$  at  $s$ , who places an offspring of same type at level  $j$ . Levels that were above  $j$  at time  $s-$  are shifted up by 1. Given the arrows and the type of the ancestral individuals (at the bottom of the picture), it is possible to reconstruct the types in the population at every time for the neutral genealogy by following the lineages back in time (travel along the arrows in opposite direction).*

To obtain the *selective births* and the *competitive deaths*, we will use a state dependent thinning of the Poisson point processes  $\mathcal{K}_i$ , to take into account the dependencies with respect to the rest of the population. The marks  $\beta$  and  $\delta$  specify whether an atom of  $\mathcal{K}_i$  corresponds to a potential selective birth or a potential competitive death. The variable  $s$  is the time at which the atom is encountered. The probabilities with which the potential selective births and the competitive selective deaths become effective involve interactions and thus depend on the state of the process. Consequently, our pathwise construction works with an acceptance-rejection rule that uses the marks  $(z, w) \in \mathbb{R}_+ \times [0, 1]$ . See Figure 2 for an illustration.

These rules are in accordance with (8): To an atom  $(s, z, w)$  of  $\mathcal{K}_i(\cdot \times \{\beta\})$  with  $z \leq b\mu^{G_{s-}}\{A\}\zeta_{s-}$  corresponds a selective birth: the individual sitting previously at level  $i$  is replaced by the offspring of a ‘uniformly’ chosen individual of type  $A$ .

To an atom of  $\mathcal{K}_i(\cdot \times \{\delta\})$  corresponds a competitive death if  $G_{s_-}(i) = B$  and  $z \leq c\mu^{G_{s_-}}\{A\}\zeta_{s_-}^2$ , or if  $G_{s_-}(i) = A$  and  $z \leq c\mu^{G_{s_-}}\{B\}\zeta_{s_-}^2$ . Then the individual at level  $i$  dies from the competition pressure exerted by the other type and is replaced by an individual chosen ‘uniformly’ among all living individuals. The way to sample an individual ‘uniformly’ among the infinite number of levels will be described in Sec. 2.3.3, and will be formally specified in Sec. 3. Indeed we will show that it is possible to define (random) sampling measures  $m_s(d\theta, dh)$  on  $\mathbb{T} \times \mathbb{I}$  where  $\mathbb{T}$  will be the completion of the set of levels  $\mathbb{N}$  with respect to the ultrametric  $R_s$  that will be described in Sec. 2.3.3 and constructed in Sec. 5.

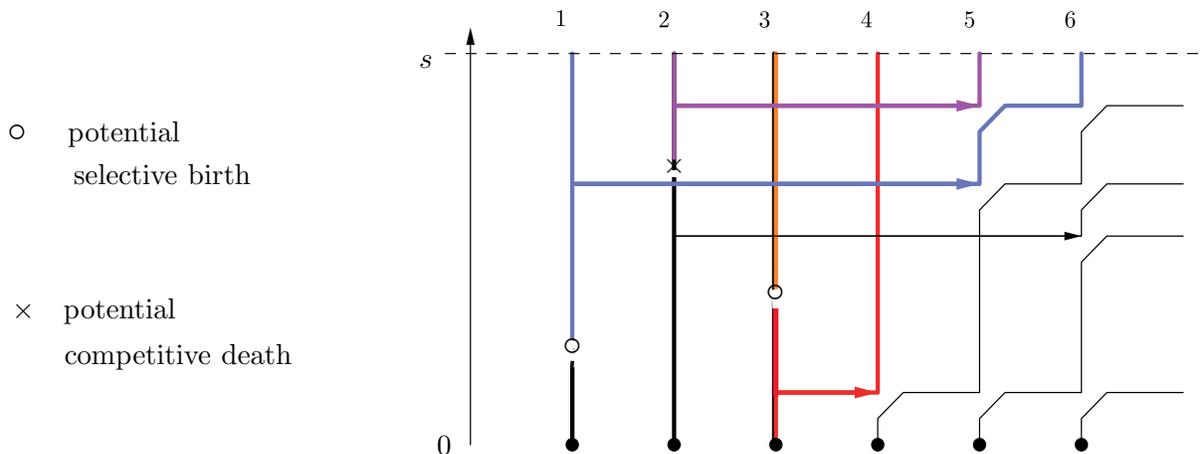


Figure 2: *The Poisson point measures  $\mathcal{K}_i$ ,  $i \in \mathbb{N}$  appear on top of the neutral lookdown representation. Their atoms correspond to potential selective births (marked as  $\circ$ ) or potential competitive deaths (marked as  $\times$ ). Each time such an event happens, the individual at the existing level may be replaced by another individual, sampled at random among the  $A$  individuals (selective births) or among the whole population (competitive death). The picture suggests (and indeed we will prove) that these events define a partition of the set  $\mathbb{N} \times \mathbb{R}_+$  (levels  $\times$  time) where the tree corresponding to the descent of each of these events is ruled by the neutral dynamics.*

To implement the acceptance-rejection rule mentioned in the previous paragraph, we will make use of a measurable mapping  $\kappa : M^1(\mathbb{I}) \times [0, 1] \rightarrow \mathbb{I}$  which is such that for a random variable  $\Upsilon$  that is uniformly distributed on the interval  $[0, 1]$ , and all  $\nu \in M^1(\mathbb{I})$ , the random variable  $\kappa(\nu, \Upsilon)$  has distribution  $\nu$ . This will allow us to construct random variables with a prescribed distribution  $\nu$ , using the third component of an atom of the Poisson point measures  $\mathcal{K}_i$  as input for  $\kappa(\nu, \cdot)$ . Recall that this third component is uniformly distributed on  $[0, 1]$ .

Our update rule for the process of type configurations  $G$  works by means of a mapping

$$q : \mathbb{I} \times \mathbb{I}^{\mathbb{N}} \times \mathbb{R}_+ \times \mathbb{R}_+ \times [0, 1] \times \{\beta, \delta\} \rightarrow \mathbb{I}$$

which prescribes how to change the type  $g(i)$  given  $G_{s_-}(i) = h$ ,  $G_{s_-} = g$  and  $\zeta_{s_-} = v$ , and given there is an atom of  $\mathcal{K}_i$  at  $(s, (z, w, \beta))$  or  $(s, (z, w, \delta))$ . Specifically, we put for a  $g$  that admits type frequencies

$$q(h, g, v, z, w, \beta) := \begin{cases} A & \text{if } z \leq b\mu^g\{A\}v, \\ h & \text{otherwise,} \end{cases} \quad (9)$$

$$q(h, g, v, z, w, \delta) := \begin{cases} \kappa(\mu^g, w) & \text{if } z \leq c\mu^g\{B\}v^2 \text{ and } h = A, \\ & \text{or } z \leq c\mu^g\{A\}v^2 \text{ and } h = B, \\ h & \text{otherwise.} \end{cases} \quad (10)$$

As in [12, Section 4] we identify the type space  $\mathbb{I} = \{A, B\}$  with the additive group  $\{0, 1\}$ . This corresponds to considering  $\mathbf{1}_{G_s(i)=B}$  instead of the type  $G_s(i)$  itself and allows to formulate more easily our SDE for the process of type configurations  $(G_s)$ . With an initial condition  $G_0$  admitting type frequencies according to (6), this is

$$\begin{aligned}
G_s(j) &= G_0(j) + \sum_{i=1}^{j-1} \int_{[0,s]} (G_{u-}(i) - G_{u-}(j)) d\mathcal{L}_{ij}(u) \\
&+ \sum_{1 \leq i < k < j} \int_{[0,s]} (G_{u-}(j-1) - G_{u-}(j)) d\mathcal{L}_{ik}(u) \\
&+ \int_{[0,s] \times \mathbb{R}_+ \times [0,1] \times \{\beta, \delta\}} (q(G_{u-}(j), G_{u-}, \zeta_{u-}, z, w, \omega) - G_{u-}(j)) \mathcal{K}_j(du, d(z, w, \omega))
\end{aligned} \tag{11}$$

for  $j \in \mathbb{N}$ .

The following theorem characterizes the mass and type configuration process  $(\zeta, G)$  with the triple  $(\mathcal{W}, \mathcal{L}, \mathcal{K})$  as the source of randomness, and also asserts the fact that a.s.  $G$  admits type frequencies at any time.

**Theorem 2.2.** Let  $G_0$  be exchangeable and admit type frequencies. Then the system (7), (11) for the total mass process  $(\zeta_s)$  and the type configurations  $(G_s)$  has a unique strong solution, up to the possibly infinite time  $\sigma := \inf\{s \geq 0 : \zeta_s = 0 \text{ or } \zeta_{s-} = \infty\}$  at which  $\zeta$  goes to extinction or explodes. For this unique solution, a.s. the type frequencies  $\mu^{G_s}$  (as defined in (6)) exist for all  $s \geq 0$ , i.e.

$$\mu^{G_s}\{A\} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{G_s(i)=A\}}. \tag{12}$$

The proof of this result will be given in Section 4, based on the preparations in Sections 3.1 and 3.2. In Section 5.1, we will build the genealogy with selection and competition on top of the neutral genealogy. The next two subsections will explain the main ideas and tools of this construction.

### 2.3.3 Filling in the ancestry: from the neutral to the selective lookdown space

With the total mass process  $\zeta$  and the type configuration process  $G$  being provided by Theorem 2.2, we can construct the ancestral lineages and “fill in” the process  $R$ . In this and the next subsection we will explain the graphical construction of the process  $R$  on top of  $(\zeta, G, \mathcal{L}, \mathcal{K})$ .

A crucial role will be played by a family of *sampling measures*. These arise as follows. As illustrated by Figure 1 (see Section 3 for a formal definition), the Poisson point process  $\mathcal{L}$  together with the initial random semi-ultrametric  $R_0$  define a random semi-metric  $\rho^{(0)}$  on  $\mathbb{R}_+ \times \mathbb{N}$ , where  $\rho^{(0)}((s, i), (s', i'))$  is the genealogical distance of  $(s, i)$  and  $(s', i')$  in the neutral case (i.e. without considering the atoms of the Poisson point measures  $\mathcal{K}$  associated with selective births and competitive deaths). The completion of  $(\mathbb{R}_+ \times \mathbb{N}, \rho^{(0)})$  is denoted by  $(Z, \rho^{(0)})$ , and called the *neutral lookdown space*. The random metric  $\rho^{(0)}$  on  $\mathbb{R}_+ \times \mathbb{N}$  is measurable with respect to  $(\mathcal{L}, R_0)$  and the completion is done realization-wise; in this sense one should think of  $(Z, \rho^{(0)})$  as a random metric space. In slight abuse of notation, we refer by  $(s, i) \in \mathbb{R}_+ \times \mathbb{N}$  also to the element of the metric space after the identification of elements with distance zero and the completion, that is we also assume  $\mathbb{R}_+ \times \mathbb{N} \subset Z$  in this sense. The space  $(Z, \rho^{(0)})$  describes the continuum of all individuals ever alive, together with their distances in the neutral genealogy.

It is known ([21], Thm 3.1) that there exists, on an event of probability 1 that does not depend on  $s$ , a family  $(\mathbf{m}_s)_{s>0}$  of probability measures on  $Z$  such that

$$\mathbf{m}_s = \text{w-}\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \delta_{(s,i)}, \quad \mathbf{m}_{s-} = \text{w-}\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \delta_{(s-,i)}, \quad (13)$$

where the w-lim in (13) refers to the weak topology on the set of probability measures on  $Z$  with respect to the metric  $\rho_0$ , and  $(s-, i) := \lim_{s' \uparrow s} (s', i)$  in  $(Z, \rho_0)$ . The measures  $\mathbf{m}_s$  allow to “sample uniformly” from the population at time  $s$ , and will be called the (family of) *neutral sampling measures*.

**Definition 2.3.** a) For a semi-ultrametric  $r \in \mathbb{R}^{\mathbb{N}^2}$ , we define  $\mathbb{T}^r$  as the completion of (the set of levels)  $\mathbb{N}$  with respect to  $r$ . For  $r \in \mathbb{R}^{\mathbb{N}^2}$  that is not a semi-ultrametric, we define  $\mathbb{T}^r$  in an arbitrary way, for definiteness as  $\mathbb{T}^r = \{1\}$ .

b) Given a marked distance matrix  $(r, g) \in \mathbb{R}^{\mathbb{N}^2} \times \mathbb{I}^{\mathbb{N}}$ , we say that  $(r, g)$  is *proper* if  $r$  is a semi-ultrametric on  $\mathbb{N}$  and

$$\mathbf{m}^{r,g} := \text{w-}\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \delta_{(i,g(i))} \quad (14)$$

exists on  $\mathbb{T}^r \times \mathbb{I}$ . We will then refer to  $\mathbf{m}^{r,g}$  as the *marked sampling measure* obtained from  $(r, g)$ . If  $(r, g)$  is not proper, we define  $\mathbf{m}^{r,g}$  in an arbitrary manner, for definiteness as  $\mathbf{m}^{r,g} := \delta_{(1,A)}$ .

c) Let  $\mathbf{m}^r = \text{w-}\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \delta_i$  denote the projection of  $\mathbf{m}^{r,g}$  to  $\mathbb{T}^r$ . If, for a proper marked distance matrix  $(r, g)$ , there exists a measurable function  $\bar{g} : \mathbb{T}^r \rightarrow \mathbb{I}$  such that

$$\mathbf{m}^{r,g}(d(\theta, h)) = \mathbf{m}^r(d\theta) \delta_{\bar{g}(\theta)}(dh), \quad \theta \in \mathbb{T}^r, h \in \mathbb{I}, \quad (15)$$

we will speak of  $\bar{g}(\theta)$  as *the type carried by the individual  $\theta$* . Note also that in this case, the type frequencies  $\mu^g\{A\}$ ,  $\mu^g\{B\}$  correspond to the projection of  $\mathbf{m}^{r,g}$  on the type component  $h \in \mathbb{I}$ .

Let

$$R_s^{(0)}(i, j) := \rho^{(0)}((s, i), (s, j)), \quad 1 \leq i \leq j \in \mathbb{N}, \quad (16)$$

be the *random distance matrix in the neutral genealogy*. In the sequel, we will prove:

**Proposition 2.4.** The random marked distance matrices  $(R_s^{(0)}, G_s)$ ,  $(R_{s-}^{(0)}, G_{s-})$   $s > 0$ , are proper on an event of probability 1 that does not depend on  $s$ .

The proof of Proposition 2.4 is given in Section 5 (see Corollary 5.2). According to Definition 2.3, this provides a.s. the *sampling measures*  $\mathbf{m}^{R_s^{(0)}, G_s}$  for all times  $s$ . At any time point  $s$  which is charged by  $\mathcal{K}$  and which is active according to the update rules (9) or (10), the individual which puts its offspring on  $(s, i)$  will be chosen by means of the sampling measure  $\mathbf{m}^{R_{s-}^{(0)}, G_{s-}}$ .

In Section 5, we will extend the concept of the neutral lookdown space to our present setting by constructing a *selective lookdown space*. Here is a short preview. On each level  $i$ , selective births and competitive deaths can occur only at the discrete time points given by  $\mathcal{K}_i$ . This discreteness allows to dissect the neutral lookdown space into countably many *fragments*, *rooted* in those points  $(s, j)$  that carry atoms of  $\mathcal{K}$  or belong to  $\{0\} \times \mathbb{N}$ . Each fragment consists of the completion of all the lineages descending from the ancestor  $(s, j)$  until a selective birth or a competitive death affects them. Hence, each fragment is monotypic, inheriting the type of its root, whose type, in turn, is determined by  $G$  from Theorem 2.2 (see Figure 2 for an illustration).

To describe all individuals ever alive by a *connected* metric space, we continue the ancestral lineages backwards in time until they hit a root, that is an atom of  $\mathcal{K}$ , say at time  $s$ . We say that a root is

*active* if the first conditions of (9) and (10) are fulfilled (otherwise, the selective fecundity event or the competition event proposed by  $\mathcal{K}$  do not happen). If a competition event occurs at time  $s$ , then choose the parent individual, i.e. the individual that continues the ancestral lineage backwards in time, according to  $\mathbf{m}_s$ . If a fecundity event occurs at time  $s$ , the individual which reproduces is drawn according to  $\mathbf{m}_s$  conditioned on the fragments being of type  $A$ . This yields a random metric space  $(\hat{Z}, \rho)$ , which is our selective lookdown space. For each  $(s, i), (s, j)$  whose ancestral lineages meet between times 0 and  $s$  we will consider the geodesic from  $(s, i)$  to  $(s, j)$  in  $(\hat{Z}, \rho)$  which we traverse with speed  $1/\zeta_{s-u}$  when passing through an ancestor that is time  $u$  back from  $s$ . The duration one needs to pass through this geodesic will be set equal to  $R_s(i, j)$ . If the two ancestral lineages lead back to two different levels  $a_i, a_j$  at time 0, then, in order to obtain  $R_s(i, j)$ , we add the distance  $R_0(a_i, a_j)$  to the sum of the durations to reach  $(0, a_i)$  from  $(s, i)$  and  $(0, a_j)$  from  $(s, j)$ .

### 2.3.4 Updating the distance matrix at neutral and selective reproduction events

In this subsection we describe the updating rule of  $R$  at the time of a reproductive event. First we consider the neutral events. When an atom of  $\mathcal{L}_{ij}$  ( $i < j$ ) is encountered at time  $s$ , the individual at level  $i$  puts a clonal offspring at level  $j$ , pushing the levels previously above and including  $j$  up by 1. The corresponding update  $\vartheta_{i,j}(r, g) = (\vartheta_{i,j}(r), \vartheta_{i,j}(g))$  of a marked distance matrix  $(r, g) \in \mathbb{R}^{\mathbb{N}^2} \times \mathbb{I}^{\mathbb{N}}$  is done by putting

$$(\vartheta_{i,j}(g))(\ell) = \begin{cases} g(\ell), & \ell < j, \\ g(\ell), & \ell = j, \\ g(\ell - 1), & \ell > j, \end{cases} \quad (17)$$

$$(\vartheta_{i,j}(r))(\ell, m) = \begin{cases} r(\ell, m), & 1 \leq \ell < m < j, \\ r(\ell, i), & 1 \leq \ell < j = m, \\ r(\ell, m - 1), & 1 \leq \ell < j < m, \\ r(i, m - 1), & \ell = j < m, \\ r(\ell - 1, m - 1), & j < \ell < m. \end{cases} \quad (18)$$

Recall that  $\vartheta_{i,j}(r)$  is symmetric with  $(\vartheta_{i,j}(r))(\ell, \ell) = 0$ .

In a selective birth or a competitive death, the individual at some level  $j \in \mathbb{N}$  is replaced by another individual from the closure of the present population. Specifically, for a marked distance matrix  $(r, g) \in \mathbb{R}^{\mathbb{N}^2} \times \mathbb{I}^{\mathbb{N}}$ , let  $\mathbb{T}$  be the completion of  $(\mathbb{N}, r)$  with respect to  $r$ , and let  $\theta \in \mathbb{T}$ ,  $h' \in \mathbb{I}$ . Then the corresponding update  $\tilde{\vartheta}_{j,\theta,h'}(r, g) = (\tilde{\vartheta}_{j,\theta}(r), \tilde{\vartheta}_{j,h'}(g))$  is done by putting

$$(\tilde{\vartheta}_{j,h'}(g))(\ell) = \begin{cases} g(\ell), & \ell \neq j, \\ h', & \ell = j, \end{cases} \quad (19)$$

$$(\tilde{\vartheta}_{j,\theta}(r))(\ell, m) = \begin{cases} r(\ell, m), & \ell, m \neq j \\ r(\ell, \theta), & \ell \neq j, m = j. \end{cases} \quad (20)$$

With the rule just described, we can read off the jumps of the marked distance matrix process  $(R, G)$  at those times  $s$  which are charged by the Poisson point process  $\mathcal{K}$ . Let us explain how the changes are parameterized by the variables attached to the atom of  $\mathcal{K}$  at time  $s$ .

If  $\mathcal{K}_j$  has an atom in  $(s, z, w, \delta)$ , then we need to pick an individual from  $m^{R_{s-}^{(0)}, G_{s-}}$ . Given  $(r, g) \in \mathbb{R}^{\mathbb{N}^2} \times \mathbb{I}^{\mathbb{N}}$ , the pick of an individual from  $m^{r, g}$  can be obtained from the measurable mapping  $w \mapsto (\theta'(w), h'(w))$  from  $[0, 1]$  to  $\mathbb{T}^r \times \mathbb{I}$  (see Def. 2.3) which transports the uniform distribution on  $[0, 1]$  into  $m^{r, g}$ . We can now specify the  $\kappa(\mu^g, w)$  appearing in (10) as

$$\kappa(\mu^g, w) = h'(w). \quad (21)$$

Then:

$$(R_s, G_s) = \begin{cases} \tilde{\vartheta}_{j, \theta'(w), h'(w)}(R_{s-}, G_{s-}) & \text{if } z \leq c \zeta_{s-}^2 \mu^{G_{s-}} \{A\} \text{ and } G_{s-}(j) = B \\ & \text{or } z \leq c \zeta_{s-}^2 \mu^{G_{s-}} \{B\} \text{ and } G_{s-}(j) = A, \\ (R_{s-}, G_{s-}) & \text{otherwise} \end{cases} \quad (22)$$

If  $\mathcal{K}_j$  has an atom in  $(s, z, w, \beta)$ , we need to sample an individual from  $m^{r, g}(d\theta, dh \mid h = A)$ . Let  $w \mapsto (\theta''(w), h''(w))$  be a measurable mapping from  $[0, 1]$  to  $\mathbb{T}^r \times \mathbb{I}$  which transports the uniform distribution on  $[0, 1]$  into  $m^{r, g}(d\theta, dh \mid h = A)$ . Notice that here, we necessarily have  $h''(w) = A$ . Then,

$$(R_s, G_s) = \begin{cases} \tilde{\vartheta}_{j, \theta''(w), A}(R_{s-}, G_{s-}) & \text{if } z \leq b \zeta_{s-} \mu^{G_{s-}} \{A\} \\ (R_{s-}, G_{s-}) & \text{otherwise.} \end{cases} \quad (23)$$

In a nutshell, we can embed all the preceding updating rules into a single SDE:

$$\begin{aligned} R_s(i, j) &= R_0(i, j) + 2 \int_{[0, s]} \zeta_u dR_u(i, j) + \sum_{1 \leq k < \ell \leq j} \int_{[0, s]} (\vartheta_{k, \ell}(R_{u-})(i, j) - R_{u-}(i, j)) d\mathcal{L}_{k\ell}(u) \\ &+ \sum_{k \in \{i, j\}} \int_{[0, s]} \mathbf{1}_{\{G_{u-}(k)=B, z \leq c \zeta_{u-}^2 \mu^{G_{u-}} \{A\}, \omega=\delta\}} (\tilde{\vartheta}_{k, \theta'(w), h'(w)}(R_{u-}, G_{u-})(i, j) - R_{u-}(i, j)) d\mathcal{K}_k(u, (z, w, \omega)) \\ &+ \sum_{k \in \{i, j\}} \int_{[0, s]} \mathbf{1}_{\{G_{u-}(k)=A, z \leq c \zeta_{u-}^2 \mu^{G_{u-}} \{B\}, \omega=\delta\}} (\tilde{\vartheta}_{k, \theta'(w), h'(w)}(R_{u-}, G_{u-})(i, j) - R_{u-}(i, j)) d\mathcal{K}_k(u, (z, w, \omega)) \\ &+ \sum_{k \in \{i, j\}} \int_{[0, s]} \mathbf{1}_{\{z \leq b \zeta_{u-} \mu^{G_{u-}} \{A\}, \omega=\beta\}} (\tilde{\vartheta}_{k, \theta''(w), h''(w)}(R_{u-}, G_{u-})(i, j) - R_{u-}(i, j)) d\mathcal{K}_k(u, (z, w, \omega)), \end{aligned} \quad (24)$$

where we write  $(r', g')(i, j) := r'(i, j)$ . In the light of the above constructions, the following result is now an immediate consequence of Theorem 2.2.

**Theorem 2.5.** Let  $G_0$  be exchangeable and admit type frequencies. Then the system (7), (11), (24) of SDE's has a unique strong solution  $(\zeta_s, R_s, G_s)_{s < \sigma}$  up to the stopping time  $\sigma$  defined in Theorem 2.2, and this process  $(\zeta, R, G)$  is Markovian.

### 2.3.5 A well-posed martingale problem for the evolving lockdown genealogy

Let  $(\mathcal{W}, \mathcal{L}, \mathcal{K})$  be as in Sec. 2.3.2, and let  $(\zeta, R, G)$  be the process provided by Theorem (2.5). The process  $\zeta$  can touch zero or explode in finite time; this happens on the event  $\sigma < \infty$  with  $\sigma$  being defined in Theorem 2.2. The time  $\sigma$  is announced by the following sequence of stopping times  $\sigma_M$ ,  $M \in \mathbb{N}$ :

$$\sigma_M := \inf \{s \geq 0, \zeta_s \notin [1/M, M]\}. \quad (25)$$

With  $b$  and  $c$  being the parameters that appear in (1), we set

$$C_M := (b \vee c)M^2. \quad (26)$$

Let us now introduce the state space for the process  $(\zeta, R, G)$  stopped at  $\sigma_M$ . We define:

$$E_M := \left( \left( \frac{1}{M}, M \right) \times \mathbb{R}^{\mathbb{N}^2} \times \mathbb{I}^{\mathbb{N}} \right) \cup \{ \Delta_M \} \quad (27)$$

where  $\left( \frac{1}{M}, M \right) \times \mathbb{R}^{\mathbb{N}^2} \times \mathbb{I}^{\mathbb{N}}$  is equipped with the product topology and where  $\Delta_M$  is a cemetery point such that a sequence  $(v_n, r_n, g_n)$  of  $E_M$  is said to converge to  $\Delta_M$  if either  $v_n \rightarrow \frac{1}{M}$  or  $v_n \rightarrow M$  as  $n \rightarrow \infty$ .

Next we display the generator of  $(\zeta, R, G)$  restricted to appropriate test functions  $F = F(v, r, g)$ , where  $v \in \mathbb{R}_+$ ,  $r \in \mathbb{R}^{\mathbb{N}^2}$  and  $g \in \mathbb{I}^{\mathbb{N}}$ .

For  $n \in \mathbb{N}$ , let  $\rho_n : \mathbb{R}^{\mathbb{N}^2} \times \mathbb{I}^{\mathbb{N}} \rightarrow \mathbb{R}^{n^2}$ ,  $(r, g) \mapsto (r(i, j))_{1 \leq i, j \leq n}$ , be the restriction map. We define  $D_{1, M}$  as the set of those functions  $f : \left( \frac{1}{M}, M \right) \times \mathbb{R}^{\mathbb{N}^2} \rightarrow \mathbb{R}$  for which there exists an  $n \in \mathbb{N}$ , a compact set  $C \subset \left( \frac{1}{M}, M \right)$  and an infinitely differentiable, bounded function  $\psi : \left( \frac{1}{M}, M \right) \times \mathbb{R}^{\mathbb{N}^2} \rightarrow \mathbb{R}$  such that  $\psi(v, r) = 0$  unless  $v \in C$ , and  $f = \psi \circ \rho_n$ .

Let  $D_2$  be the set of those functions  $\gamma : \mathbb{I}^{\mathbb{N}} \rightarrow \mathbb{R}$  for which there exists an  $n \in \mathbb{N}$  such that  $\gamma(g)$  depends only on the first  $n$  coordinates of  $g$ .

We now consider functions of the form

$$F(v, r, g) = f(v, r)\gamma(g) \quad (28)$$

where  $f \in D_{1, M}$  and  $\gamma \in D_2$ . The smallest possible  $n \in \mathbb{N}$  which fits to the required representations of  $f$  and  $\gamma$  will be called the *degree* of  $F$ . We write  $F_{r(i, j)}$  for the partial derivative of  $F$  with respect to the variable  $r(i, j)$ , and  $F_v$  for partial derivatives of  $F$  with respect to  $v$ .

Let  $\vartheta_{i, j}$  and  $\tilde{\vartheta}_{j, \theta, h'}$  be as in (17), (18), (19) and (20). Let  $F$  be as in (28) with degree  $n$ . For a proper pair  $(r, g)$ , let  $m^{r, g}$  be the sampling measure as in (14), and  $\mu^g$  be the second marginal of  $m^{r, g}$ , which is equal to the type distribution belonging to  $g$ . For functions  $F$  that are of the form (28) we define  $\mathbf{A}F$  as follows:

$$\begin{aligned} \mathbf{A}F(v, r, g) &= \frac{v^2}{2} F_{vv}(v, r, g) + (bv^2 \mu^g\{A\} - 2cv^3 \mu^g\{A\} \mu^g\{B\}) F_v(v, r, g) \\ &\quad + 2v \sum_{1 \leq i \neq j \leq n} F_{r(i, j)}(v, r, g) \\ &\quad + \sum_{1 \leq i < j \leq n} (F(v, \vartheta_{i, j}(r, g)) - F(v, r, g)) \\ &\quad + cv^2 \mu^g\{A\} \sum_{j=1}^n \int m^{r, g}(d\theta, dh') \mathbf{1}_{\{g(j)=B\}} (F(v, \tilde{\vartheta}_{j, \theta, h'}(r, g)) - F(v, r, g)) \\ &\quad + cv^2 \mu^g\{B\} \sum_{j=1}^n \int m^{r, g}(d\theta, dh') \mathbf{1}_{\{g(j)=A\}} (F(v, \tilde{\vartheta}_{j, \theta, h'}(r, g)) - F(v, r, g)) \\ &\quad + bv \sum_{j=1}^n \int m^{r, g}(d\theta, dh') \mathbf{1}_{\{h'=A\}} (F(v, \tilde{\vartheta}_{j, \theta, h'}(r, g)) - F(v, r, g)). \end{aligned} \quad (29)$$

Let  $D_M$  be the linear span of the constant real-valued functions on  $E_M$  (defined in (27)) and all functions of the form (28). The linear extension of (29) to  $D_M$  will again be denoted by  $\mathbf{A}$ .

**Proposition 2.6.** For all  $M > 0$ , the process  $(\zeta_{s \wedge \sigma_M}, R_{s \wedge \sigma_M}, G_{s \wedge \sigma_M})_{s \geq 0}$  solves the martingale problem  $(\mathbf{A}, D_M)$ , and this martingale problem is well-posed.

The proof will be given in Section 5.2. There, our first step (Proposition 5.4) will be to establish a well-posed martingale problem for a refinement  $(\zeta, R, G, \Lambda)$ , where  $\Lambda$  counts the points of  $\mathcal{L}$  and  $\mathcal{K}$  and thus

keeps track of all the essential graphical ingredients that are needed to specify the jump distribution of  $R$  at these points. The second step will complete the proof of Proposition 2.6 by applying Kurtz's Markov Mapping Theorem, thus projecting to a well-posed martingale problem for the first three components  $(\zeta, R, G)$ . Let us also mention that a similar strategy has been applied in Lemma 4.2 in [12] in a situation without the components  $\zeta$  and  $R$ , i.e. for a dynamics with constant population size and without consideration of the genealogies.

## 2.4 A well-posed martingale problem for the evolving symmetrized genealogy

Let  $(\zeta_s, R_s, G_s)$ ,  $s < \sigma$ , be the process provided by Theorem 2.5. In Corollary 5.2 we will prove that a.s.  $(R_s, G_s)$  is proper in the sense of Definition 2.3. Hence we can define a process of marked genealogies (see Sec. 2.2) whose state at each time  $s$  is the isomorphy class of the marked ultrametric measure space  $(\mathbb{T}^{R_s}, R_s, m^{R_s, G_s})$  (again see Definition 2.3). Recalling the notation  $X_s = (R_s, G_s)$ , we denote this isomorphy class by  $\psi(X_s)$ .

For  $\chi \in \mathbb{M}$ , the space of marked genealogies (see Sec. 2.2), we will write  $\nu^\chi$  for the *marked distance matrix distribution* obtained from  $\chi$ , i.e. the distribution of  $(R^\chi, G^\chi)$  where  $R^\chi$  is the distance matrix and  $G^\chi$  is the type configuration of a sequence drawn i.i.d. from the sampling measure belonging to (an arbitrary representative of)  $\chi$ . For a prescribed initial condition  $(v_0, \chi_0) \in (0, \infty) \times \mathbb{M}$ , we define  $X_0 := (R^{\chi_0}, G^{\chi_0})$  and take  $(v_0, X_0)$  as initial condition for the process  $(\zeta, X)$ . Let the time change  $t \rightarrow s(t)$  be as in (3). For  $t$  such that  $s(t) < \sigma$  we define

$$(\xi_t, Y_t) := (\zeta_{s(t)}, \psi(X_{s(t)})). \quad (30)$$

We now set out to describe the process  $(\xi, Y)$  by a stopped martingale problem. That  $\xi$  can reach zero or converge to infinity may be problematic for the change of time (3). That is why it is natural to introduce, for a fixed positive integer  $M > 0$ , the stopping time

$$\tau_M = \inf\{t \in \mathbb{R}_+, \xi_t \notin [1/M, M]\} \quad (31)$$

and the stopped processes  $\xi^{\tau_M} = \xi_{\cdot \wedge \tau_M}$  and  $Y^{\tau_M} = Y_{\cdot \wedge \tau_M}$ . Let us also define  $\tau_0 = \inf\{t \geq 0, \xi_t = 0\}$ . Without restriction, we can choose  $M$  so that the initial condition  $v_0$  of the mass process satisfies  $v_0 \in (\frac{1}{M}, M)$ . We then have that  $0 < \tau_M \leq \tau_0$  a.s., and  $\lim_{M \rightarrow +\infty} \tau_M \leq \tau_0$  by the continuity of  $\xi$ .

For  $F \in D_M$  (defined just after (29)),  $v > 0$  and  $\chi \in \mathbb{M}$  we put

$$\Phi_F(v, \chi) := \int F(v, r, g) \nu^\chi(dr, dg), \quad \mathbb{A}\Phi_F(v, \chi) := \int \frac{1}{v} \mathbf{A}F(v, r, g) \nu^\chi(dr, dg). \quad (32)$$

In analogy to (27) we now consider the state space

$$S_M := ((\frac{1}{M}, M) \times \mathbb{M}) \cup \{\Delta_M\}, \quad (33)$$

where  $(\frac{1}{M}, M) \times \mathbb{M}$  is equipped with the product topology and a sequence  $(v_n, \chi_n)$  is said to converge to  $\Delta_M$  if either  $v_n \rightarrow \frac{1}{M}$  or  $v_n \rightarrow M$  as  $n \rightarrow \infty$ . In other words, this corresponds to a ‘‘lumping’’ of all states  $(\frac{1}{M}, \chi)$  and  $(M, \chi)$  with  $\chi \in \mathbb{M}$  into one state  $\Delta_M$ .

**Theorem 2.7.** For  $(v_0, \chi_0) \in (0, \infty) \times \mathbb{M}$  and  $M > 1/v_0$ , the process  $(\xi^{\tau_M}, Y^{\tau_M})$  is Markovian and gives the unique solution of the martingale problem

$$(\xi_0, Y_0) = (v_0, \chi_0), \quad \Phi_F(\xi_t^{\tau_M}, Y_t^{\tau_M}) - \int_0^{t \wedge \tau_M} \mathbb{A}\Phi_F(\xi_u, Y_u) du = \text{martingale}, \quad F \in D_M. \quad (34)$$

Theorem 2.7 will be proven in Section 5.3.

**Remark 2.8.** Since  $\tau_M \uparrow \tau_0$  a.s. as  $M \rightarrow \infty$ , the process  $(\xi_t, Y_t)_{t < \tau_0}$  is characterized in distribution by the requirement that, when stopped at  $\tau_M$ , it solves the martingale problem (34) for all  $M \in \mathbb{N}$ .

In contrast to  $(X_s)_{s < \sigma}$ , which has jumps, the process  $(Y_t)_{t < \tau_0}$  is continuous. This is contained in the next result, which will be proved in Section 5.3.

**Proposition 2.9.** For  $(v_0, \chi_0) \in (0, \infty) \times \mathbb{M}$  and  $M \in \mathbb{N}$ , the process  $(\xi_{t \wedge \tau_M}, Y_{t \wedge \tau_M})_{t \geq 0}$  has a.s. continuous paths in  $\mathbb{R}_+ \times \mathbb{M}$ .

If  $(\xi, Y)$  is the solution of (34), then  $\xi$  is a weak solution of (2), as can be seen immediately by projecting (34) to its first component. We can also recover the equations for  $\xi^A$  and  $\xi^B$  given in (1). For this, we first recall that the type frequencies  $\mu_t\{A\}$  and  $\mu_t\{B\}$  can be recovered from the projection of  $m^{Y_t}(d(\theta, h))$  on its second component. Consequently,  $(\xi_t^A, \xi_t^B)$  is defined in terms of  $(\xi, Y)$  as  $(\xi_t \mu_t^A, \xi_t \mu_t^B)$ .

**Proposition 2.10.** Let  $(\xi, Y)$  be the solution of (34). Then  $(\xi_t^A, \xi_t^B)_{t \geq 0}$  is a weak solution of solution of the SDE (1).

The proof will be given in Section 5.4.

### 3 Building blocks from neutrality

#### 3.1 Neutral lookdown space and marked sampling measures

The *neutral setting* will provide the building blocks for the analysis of the genealogy also in the presence of selection and competition, and we study it specifically in this section. Its only ingredients are the initial condition  $(R_0, G_0)$  (being distributed according to the marked distance matrix distribution of a marked ultrametric measure space), and the neutral birth events given by the Poisson point measures  $\{\mathcal{L}_{i,j}, 1 \leq i < j\}$  (see (I2) in Section 2). As illustrated by Figure 1, each of the points of  $\mathcal{L}_{i,j}$  can be seen as a *merger* of two ancestral lineages: if  $\mathcal{L}_{i,j}$  has an atom at time  $s$ , then the ancestral lineage of  $(s, j)$  starts, back into the past, from  $(s-, i)$ , from there on being identical with the ancestral lineage of  $(s, i)$ . In this case, the (neutral) genealogical distance of  $(s, i)$  and  $(s, j)$  equals zero; more generally, the neutral genealogical distance of  $(s_1, i)$  and  $(s_2, j)$  is determined as follows: trace the neutral ancestral lineages back from  $(s_1, i)$  and  $(s_2, j)$ . If they merge at time  $u \geq 0$ , then the distance is  $(s_1 - u) + (s_2 - u)$ . Otherwise, if  $a_1$  and  $a_2$  are the labels of the two neutral ancestors at time 0, the distance is  $s_1 + s_2 + R_0(a_1, a_2)$ . For given  $R_0$ , this gives rise to an  $\mathcal{L}$ -measurable random semi-ultrametric  $\rho^{(0)}$  on  $\mathbb{R}_+ \times \mathbb{N}$ . The *neutral lookdown space* is the metric completion of  $(\mathbb{R}_+ \times \mathbb{N}, \rho^{(0)})$ , denoted by  $(Z, \rho^{(0)})$ . It carries the family of *sampling measures*  $\mathfrak{m}_s$ ,  $s > 0$ , defined by (13).

By the Glivenko-Cantelli lemma, the assumption that  $(R_0, G_0)$  has the marked distance matrix distribution of a marked ultrametric measure space ensures that a.s.

$$\text{w-} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \delta_{(i, G_0(i))} \quad \text{exists on } \mathbb{T}_0 \times \mathbb{I}, \quad (35)$$

where  $\mathbb{T}_0$  is the metric completion of  $(\mathbb{N}, R_0)$ . This clearly implies that

$$\mathfrak{m}_0 := \text{w-} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \delta_{(0, i)} \quad (36)$$

exists on  $Z$ , including the case  $s = 0$  into (13). Likewise, (35) implies that

$$\mathfrak{m}_0 := \text{w-} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \delta_{((0, i), G_0(i))} \quad (37)$$

exists on  $Z \times \mathbb{I}$  a.s. With the notation introduced in Definition 2.3, and since  $R_0$  is a random semi-ultrametric, (37) says that  $(R_0, G_0)$  is a.s. proper.

In the next lemma we show, based on the existence of the neutral sampling measure (13), that the corresponding statement also holds true for  $s > 0$ .

**Lemma 3.1.** For the neutral genealogy, and for a time  $s \geq 0$ , denote by  $G_s^{(0)}(i)$  the type of  $(s, i)$  given by

$$G_s^{(0)}(i) := G_0(a), \quad (38)$$

where  $a$  is the level of the neutral ancestor of  $(s, i)$  at time 0, and recall the notation  $R_s^{(0)}(i, j)$  given in (16) for the random distance matrix. Almost surely,  $(R_s^{(0)}, G_s^{(0)})$  and the left limit  $(R_{s-}^{(0)}, G_{s-}^{(0)})$  are proper for all  $s > 0$ . In particular, the weak limit

$$m_s^{(0)} := \text{w-} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \delta_{((s,i), G_s^{(0)}(i))} \quad (39)$$

exists on  $Z \times \mathbb{I}$  on an event of probability 1 that does not depend on  $s$ .

*Proof.* Fix  $s > 0$ . The map  $\hat{f} : \{s\} \times \mathbb{N} \rightarrow \mathbb{I}$ ,  $(s, i) \mapsto G_s^{(0)}(i)$ , is uniformly continuous with respect to  $\rho^{(0)}$ . To see this, let  $\delta < 2s$  arbitrary and suppose  $\rho^{(0)}((s, i_1), (s, i_2)) < \delta$ . Then  $(s, i_1)$  and  $(s, i_2)$  have a common ancestor at time  $s - \delta/2$ ; consequently their types coincide.

Thus the map  $\hat{f}$  can be extended to a (uniformly) continuous function  $\hat{f} : \overline{\{s\} \times \mathbb{N}} \rightarrow \mathbb{I}$ , where  $\overline{\{s\} \times \mathbb{N}}$  denotes the closure of  $\{s\} \times \mathbb{N}$  with respect to  $\rho^{(0)}$ . Then the map  $f : \overline{\{s\} \times \mathbb{N}} \rightarrow \overline{\{s\} \times \mathbb{N}} \times \mathbb{I}$ ,  $\theta \mapsto (\theta, \hat{f}(\theta))$  is also continuous. It satisfies

$$\frac{1}{n} \sum_{i=1}^n \delta_{((s,i), G_s^{(0)}(i))} = \left( \frac{1}{n} \sum_{i=1}^n \delta_{(s,i)} \right) \circ f^{-1},$$

where the right hand side denotes the image measure under  $f$ . The continuous mapping theorem and (13) imply that  $\frac{1}{n} \sum_{i=1}^n \delta_{((s,i), G_s^{(0)}(i))}$  converges weakly to  $m_s^{(0)} = \mathbf{m}_s \circ f^{-1}$  on an event of probability 1 that does not depend on  $s$ . This shows  $(R_s^{(0)}, G_s^{(0)})$  is proper. The assertion on  $(R_{s-}^{(0)}, G_{s-}^{(0)})$  follows using the left limit in (13).  $\square$

**Remark 3.2.** The function  $\hat{f}$  is called mark function in the sense of e.g. [26]. The  $Z$ -component of  $m_s^{(0)}$  is  $\mathbf{m}_s$  given by (13), and the  $\mathbb{I}$ -component of  $m_s^{(0)}$  is the type distribution  $\mu_s^{(0)}$  under the neutral transport, which in view of (39) and (6) obeys a.s.

$$\mu_s^{(0)} = m_s^{(0)}(Z \times \cdot) = \text{w-} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \delta_{G_s^{(0)}(i)}. \quad (40)$$

### 3.2 Partitioning the lookdown space: roots and fragments

Throughout the article, we assume that our initial state  $(R_0, G_0)$  has the marked distance matrix distribution of a marked ultrametric measure space, so that in particular we have the neutral sampling measure (39) at hand. Recall the families of independent Poisson point measures  $\{\mathcal{K}_i : i \in \mathbb{N}\}$  from ingredient (I3). We now partition (up to a set that is not charged by any of the sampling measures  $\mathbf{m}_s$ ) the entire space  $Z$  into (what we call) *fragments*  $\Gamma_\gamma \subset Z$  with *roots*  $\gamma \in \mathbb{R}_+ \times \mathbb{N}$  as follows. On top of the neutral

lookdown construction, we think of a competition respectively a fecundity “cross” added at  $(s, i)$  when  $\mathcal{K}_i$  places an atom at  $(s, i)$ . To be more precise, fix  $C < \infty$  and put

$$\tilde{\mathcal{R}} := \bigcup_{i \in \mathbb{N}} \left( \{s \in \mathbb{R}_+ : \mathcal{K}_i(\{s\} \times [0, C] \times [0, 1] \times \{\beta, \delta\}) \geq 1\} \times \{i\} \right) \subset \mathbb{R}_+ \times \mathbb{N} \quad (41)$$

Note that the restriction to  $[0, C]$  guarantees that the overall rate of potential events on a fixed level is bounded on any finite time-interval. As a consequence the points in  $\tilde{\mathcal{R}} \cap (\mathbb{R}_+ \times \{i\})$  do not accumulate for fixed  $i \in \mathbb{N}$  almost surely.

Now let

$$\mathcal{R} := \tilde{\mathcal{R}} \cup (\{0\} \times \mathbb{N}) \subset \mathbb{R}_+ \times \mathbb{N}$$

be the set of *roots*. The types and lineages in the subtree above a root evolve according to the dynamics of the neutral model until they hit another root.

**Remark 3.3.** Note that the (neutral) ancestral lineage of each element  $\theta \in Z \setminus (\mathbb{R}_+ \times \mathbb{N})$  is well-defined. Indeed, take a sequence  $(s_n, \ell_n) \in \mathbb{R}_+ \times \mathbb{N}$ ,  $n \in \mathbb{N}$ , such that  $(s_n, \ell_n) \rightarrow \theta$  for  $n \rightarrow \infty$ . Then we have in particular  $s_n \rightarrow s \in \mathbb{R}_+$ . Without loss of generality, assume  $(s_n)_{n \in \mathbb{N}}$  is monotonically increasing. How to determine the ancestor of  $\theta$  at time  $s - \epsilon$  for  $0 < \epsilon < s$  arbitrary? There exists  $n_0 \in \mathbb{N}$  such that  $\rho^{(0)}((s_n, \ell_n), \theta) < \epsilon$  for all  $n \geq n_0$ , that is,  $(s_n, \ell_n)$  and  $\theta$  have a common ancestor at time  $s - \epsilon$  for all  $n \geq n_0$ . Thus, take the ancestor of  $\theta$  at time  $s - \epsilon$  to be the one of  $(s_{n_0}, \ell_{n_0})$ .

For each  $\gamma \in \mathcal{R}$  let

$$\Gamma_\gamma := \{\theta \in Z : \theta \text{ descends from } \gamma \text{ and there are no points in } \mathcal{R} \text{ on the} \\ \text{lineage connecting } \theta \text{ with } \gamma\} \subset Z.$$

**Remark 3.4.** We make the following observations.

- 1) Interpret  $\Gamma_\gamma$  as descendants of  $\gamma$  in a *neutral* infinite alleles model with mutation. Here, the frequencies exist and  $Z$  can be broken into a countable number of fragments, rooted in  $\mathcal{R}$ . This construction yields  $\mathbf{m}_s(\Gamma_\gamma)$  for *all* times  $s \geq 0$ . It can be shown that a. s.,

$$\mathbf{m}_s\left(\bigcup_{\gamma \in \mathcal{R}} \Gamma_\gamma\right) = 1$$

for all  $s > 0$ . Further details are given in the proof of Lemma 3.5 below.

- 2) By restricting to  $\{s\} \times \mathbb{N}$ , a partition of  $\mathbb{N}$  is inherited. This partition depends on  $s$ , where  $i \sim j$  if  $(s, i)$  and  $(s, j)$  have a common ancestor living between times 0 and  $s$  and there is no root on their geodesics.
- 3) In contrast to a tree-valued process whose states describe genealogical trees at fixed times, the lookdown space describes all individuals which live at any time. From this object, we can read off the state of the tree-valued processes at time  $s$  using a restriction of the lookdown space. The lookdown space itself however is universal for all  $s$ .

The sets  $\Gamma_\gamma$ ,  $\gamma \in \mathcal{R}$ , form a partition of the set that is obtained from  $Z$  by removing the accumulation points of  $\mathcal{R}$  in  $Z$ . Almost surely, the set of these accumulation points has zero mass under all  $\mathbf{m}_s$ ,  $s > 0$ . This is the contents of the following lemma.

**Lemma 3.5.** Almost surely,  $\mathbf{m}_s(Z \setminus \bigcup_{\gamma \in \mathcal{R}} \Gamma_\gamma) = 0$  for all  $s > 0$ .

*Proof.* The points of  $\mathcal{R}$  can be thought of as mutation events in an infinite alleles model that come with rate  $C$  along the lineages in a lookdown model, say with type space  $[0, 1]$  and parent independent mutation where the type in each mutation event is drawn uniformly and independently. To obtain a contradiction, assume that there exists an  $s$  such that the set of accumulation points of  $\mathcal{R}$  has nonzero mass under  $\mathbf{m}_s$ . As all such accumulation points have different types in the infinite alleles model, this results in a type distribution of mass smaller than 1. The lookdown construction for the infinite alleles model [13, Theorem 3.2] shows, however, that there are a.s. no exceptional time points with defective type distribution.  $\square$

**Corollary 3.6.** Almost surely,

$$\mathbf{m}_s(\Gamma_\gamma) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(s,i) \in \Gamma_\gamma}$$

for all  $\gamma \in \mathcal{R}$  and  $s > 0$ .

*Proof.* Almost surely, for each  $s > 0$  and  $\gamma \in \mathcal{R}$ , the boundary of  $\Gamma_\gamma$  as a subset of  $\overline{\{s\} \times \mathbb{N}} \subset Z$  is not charged by  $\mathbf{m}_s$ . Hence, the Portmanteau theorem and (13) yield the result.  $\square$

**Lemma 3.7.** Almost surely,

- (i) the fragment masses  $\mathbf{m}_s(\Gamma_\gamma)$  are continuous in  $s$  for each  $\gamma$ ,
- (ii) for each fragment  $\Gamma_\gamma$ , the restriction  $\mathbf{m}_s(\cdot \cap \Gamma_\gamma)$  is continuous in  $s$  with respect to the weak topology on  $(\Gamma_\gamma, \rho^{(0)})$ .

*Proof.* The statement (i) follows by relating the assertion to an infinite alleles model, similar as in the proof of Lemma 3.5.

To prove (ii), note that a discontinuity at a time  $s$  implies the existence of a closed subset  $A$  of  $(\Gamma_\gamma, \rho^{(0)})$  with  $\limsup_{s' \rightarrow s} \mathbf{m}_{s'}(A) > \mathbf{m}_s(A)$ . Since  $\mathbf{m}_s(\overline{\Gamma_\gamma} \setminus \Gamma_\gamma) = 0$  by Lemma 3.5, the inequality holds also for the closure of  $A$  in  $(Z, \rho^{(0)})$ . Thus, a discontinuity at time  $s$  results in a discontinuity of  $\mathbf{m}_s$  on the neutral lookdown space in contradiction to [21, Theorem 3.1].  $\square$

**Lemma 3.8.** For all  $\epsilon, T > 0$  there exists almost surely a random  $\ell \in \mathbb{N}$  such that

$$\sum_{\gamma=(u,i) \in \mathcal{R}, \text{ with } u \in \mathbb{R}_+, i \leq \ell} \mathbf{m}_s(\Gamma_\gamma) \geq 1 - \epsilon \text{ for all } s \in [0, T].$$

*Proof.* For fixed  $s \in [0, T]$ , this follows from Lemma 3.5. Let

$$\vartheta_k = \inf \left\{ s \in [0, 2T] : \sum_{\gamma=(u,i) \text{ with } u \in \mathbb{R}_+, i \leq k} \mathbf{m}_s(\Gamma_\gamma) < 1 - \epsilon \right\},$$

where we set  $\inf \emptyset = 2T$ . Then  $\vartheta_k$  is monotonically increasing in  $k$ . Set

$$\vartheta = \sup_{k \in \mathbb{N}} \vartheta_k.$$

It suffices to show that  $\vartheta = 2T$  almost surely.

On the event that  $\vartheta < 2T$ , there exists by Lemma 3.5 a.s.  $k \in \mathbb{N}$  with

$$\sum_{\gamma=(u,i) \text{ with } u \in \mathbb{R}_+, i \leq k} \mathbf{m}_\vartheta(\Gamma_\gamma) \geq 1 - \epsilon/2.$$

By Lemma 3.7(i), almost surely, the fragment masses  $\mathbf{m}_s(\Gamma_\gamma)$  are continuous in  $s$  for each  $\gamma$ . Hence, there exists  $\delta > 0$  with

$$\sum_{\gamma=(u,i) \text{ with } u \in \mathbb{R}_+, i \leq k} \mathbf{m}_s(\Gamma_\gamma) \geq 1 - \epsilon$$

for all  $s \in (\vartheta - \delta, \vartheta + \delta)$ . This implies that for all  $\ell \geq k$  with  $\vartheta_\ell > \vartheta - \delta$ , we have  $\vartheta_\ell \geq \vartheta + \delta$ , in contradiction to the definition of  $\vartheta$ . Thus  $\{\vartheta < 2T\}$  must be a null event and the claim follows.  $\square$

## 4 An SDE for type configuration and population size: proof of Theorem 2.2

In this section we provide an iteration scheme which leads to the proof of Theorem 2.2. We will be guided by the proof of Theorem 4.1 in [12]. The additional (and substantial) challenge that is overcome in our proof is that the total mass, which in [12] was assumed constant, now is a stochastic process which depends on the type configurations.

Recalling the ingredients from Section 2.3.2, we will work with the filtration  $\mathcal{F} = (\mathcal{F}_s)$ , where  $\mathcal{F}_s$  is generated by  $\mathcal{W}_u$ ,  $u \leq s$  and those points in  $\mathcal{L}$  and  $\mathcal{K}$  whose time component is less than  $s$ . Following the steps described in Section 2.3, we will prove the existence and uniqueness of the type process  $G$  in (11). A substantial difficulty is that the SDEs for  $G$  depend on the mass process  $\zeta$  (in the lookdown time-scale) that itself depends on the process  $\mu^{G_\cdot}\{A\}$  of proportions of type  $A$ . Also, Theorem 2.2 asserts that  $G_s$  admits type frequencies for all times  $s > 0$ , so that  $\mu^{G_s}\{A\}$  is well-defined.

Let us introduce the following function, describing the drift of the process  $\xi$ :

$$f(v, p) = bpv - 2cp(1-p)v^2, \quad v \geq 0, p \in [0, 1]. \quad (42)$$

Fix a constant  $C \in (0, \infty)$  in (41) that bounds the rate at which selective and competitive events occur. For a modification of the system of SDEs (11) and (7), where we use  $C$  and  $M$  to control the dynamics (see third term in the r.h.s. of (44) below), we prove existence and strong uniqueness by a Picard iteration-like argument. For this we put

$$f_M(v, p) = f((v \vee \frac{1}{M}) \wedge M, p). \quad (43)$$

The following key proposition treats SDEs similar to the ones in Theorem 2.2, but with  $f_M$  instead of  $f$ , which simplifies the problem of controlling the population size. Replacement of  $f_M$  by  $f$  will be treated at the end of the section, in the *completion of the proof of Theorem 2.2*.

**Proposition 4.1.** Let  $G_0$  be an exchangeable initial type configuration in  $\mathbb{I}^{\mathbb{N}}$  that admits type frequencies (see (6)). The following system (44), (45) of SDEs has a unique strong solution:

$$\begin{aligned} G_s(j) &= G_0(j) + \sum_{i=1}^{j-1} \int_{[0,s]} (G_{u-}(i) - G_{u-}(j)) d\mathcal{L}_{ij}(u) \\ &+ \sum_{1 \leq i < k < j} \int_{[0,s]} (G_{u-}(j-1) - G_{u-}(j)) d\mathcal{L}_{ik}(u) \\ &+ \int_{[0,s] \times [0,C] \times [0,1] \times \{\beta, \delta\}} (q(G_{u-}(j), G_{u-}, (\zeta_{u-} \vee \frac{1}{M}) \wedge M, z, w, \omega) - G_{u-}(j)) \mathcal{K}_j(du, d(z, w, \omega)), \end{aligned} \quad (44)$$

$$\zeta_0 = v_0, \quad d\zeta_s = \zeta_s f_M(\zeta_s, \mu^{G_s}\{A\}) ds + \zeta_s d\mathcal{W}_s, \quad j \in \mathbb{N}, \quad s \geq 0. \quad (45)$$

For this unique solution, a.s. the type frequencies for  $G_s$  exist for all  $s \geq 0$ .

In order to prepare the proof of this proposition, we first show a statement on the continuous dependence of (45) on its input  $\mu^{G_s}\{A\}$ .

**Lemma 4.2.** Let  $\mathcal{W}$  be an  $(\mathcal{F}_s)$ -adapted Brownian motion and let  $\varphi, \tilde{\varphi}$  be  $(\mathcal{F}_s)$ -adapted,  $[0, 1]$ -valued and continuous. Let  $\alpha, \tilde{\alpha}$  obey

$$\begin{aligned} d\alpha_s &= \alpha_s f_M(\alpha_s, \varphi_s) ds + \alpha_s d\mathcal{W}_s, \\ d\tilde{\alpha}_s &= \tilde{\alpha}_s f_M(\tilde{\alpha}_s, \tilde{\varphi}_s) ds + \tilde{\alpha}_s d\mathcal{W}_s, \end{aligned}$$

with the same initial condition  $v_0$  in  $[\frac{1}{M}, M]$  at time 0. Then there exists a constant  $\tilde{C}$  (depending on  $M$  but not depending on  $s$ ) such that for all  $s \geq 0$  we have

$$|(\alpha_s \vee \frac{1}{M}) \wedge M - (\tilde{\alpha}_s \vee \frac{1}{M}) \wedge M| \leq \tilde{C} s e^s \int_0^s |\varphi_u - \tilde{\varphi}_u| du \quad \text{a.s.} \quad (46)$$

*Proof of Lemma 4.2.* By Itô's formula,

$$\begin{aligned} d \ln \alpha &= \frac{1}{\alpha} d\alpha - \frac{1}{2\alpha^2} d[\alpha] = f_M(\alpha, \varphi) ds + d\mathcal{W} - \frac{1}{2} ds, \\ d \ln \tilde{\alpha} &= \frac{1}{\tilde{\alpha}} d\tilde{\alpha} - \frac{1}{2\tilde{\alpha}^2} d[\tilde{\alpha}] = f_M(\tilde{\alpha}, \tilde{\varphi}) ds + d\mathcal{W} - \frac{1}{2} ds. \end{aligned}$$

Subtracting one equation from the other and using the triangle inequality we get

$$|\ln \alpha_s - \ln \tilde{\alpha}_s| \leq \int_0^s |f_M(\alpha_u, \varphi_u) - f_M(\tilde{\alpha}_u, \tilde{\varphi}_u)| du. \quad (47)$$

There exists a constant  $c_1$  (depending on  $M$  but not depending on  $s$ ) such that for all  $s \geq 0$

$$|(\alpha_s \vee \frac{1}{M}) \wedge M - (\tilde{\alpha}_s \vee \frac{1}{M}) \wedge M| \leq c_1 |\ln((\alpha_s \vee \frac{1}{M}) \wedge M) - \ln((\tilde{\alpha}_s \vee \frac{1}{M}) \wedge M)|. \quad (48)$$

From (47), (48) and the Lipschitz property of  $f_M$  we obtain

$$|(\alpha_s \vee \frac{1}{M}) \wedge M - (\tilde{\alpha}_s \vee \frac{1}{M}) \wedge M| \leq c_2 \int_0^s (|(\alpha_u \vee \frac{1}{M}) \wedge M - (\tilde{\alpha}_u \vee \frac{1}{M}) \wedge M| + |\varphi_u - \tilde{\varphi}_u|) du.$$

Using Gronwall's inequality we arrive at

$$\begin{aligned} |(\alpha_s \vee \frac{1}{M}) \wedge M - (\tilde{\alpha}_s \vee \frac{1}{M}) \wedge M| &\leq c_2 \int_0^s |\varphi_u - \tilde{\varphi}_u| du + c_2^2 \int_0^s \int_0^u |\varphi_w - \tilde{\varphi}_w| dw e^u du \\ &\leq \tilde{C} s e^s \int_0^s |\varphi_u - \tilde{\varphi}_u| du. \end{aligned}$$

□

Most of the remainder of this section is devoted to the proof of Proposition 4.1 which uses an iteration scheme. To get this scheme started, we take  $G_s^{(0)}$  as the neutral type transport defined by (38) and  $\mu_s^{(0)}$  as the neutral type distributions given by (40).

**Step 1, Recursion hypothesis:** Assume that for  $\ell = 0, \dots, k-1$  we have defined  $\mathcal{F}$ -adapted  $\mathbb{I}^{\mathbb{N}}$ -valued processes  $G^{(\ell)}$  and continuous  $\mathbb{R}_+$ -valued processes  $\zeta^{(\ell)}$  such that:

- almost surely,  $G_s^{(\ell)}$  admits type frequencies  $\mu_s^{(\ell)}$  for all  $s > 0$ , with

$$\mu_s^{(\ell)} = \text{w-}\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \delta_{G_s^{(\ell)}(i)},$$

and  $s \mapsto \mu_s^{(\ell)}$  is continuous,

- the process  $\zeta^{(\ell)}$  is the unique strong solution of the SDE

$$\zeta_0^{(\ell)} = v_0, \quad d\zeta_s^{(\ell)} = \zeta_s^{(\ell)} f_M(\zeta_s^{(\ell)}, \mu_s^{(\ell)} \{A\}) ds + \zeta_s^{(\ell)} d\mathcal{W}_s. \quad (49)$$

The fact that (49) indeed has a unique strong solution follows e.g. from [37, Theorem 5.3].

**Step 2, Setting up the iteration step:** In order to define  $G^{(k)}$  in terms of  $G^{(k-1)}$ ,  $\zeta^{(k-1)}$ ,  $\mathcal{L}$  and  $\mathcal{K}$ , we consider the following system of SDE's where the function  $q$  (cf. (9)–(10)) uses the type frequencies  $\mu^{(k-1)}$  which are well-defined by our recursion hypothesis.

$$\begin{aligned} G_s^{(k)}(j) &= G_0(j) + \sum_{i=1}^{j-1} \int_{[0,s]} (G_{u-}^{(k)}(i) - G_{u-}^{(k)}(j)) d\mathcal{L}_{ij}(u) \\ &+ \sum_{1 \leq i < l < j} \int_{[0,s]} (G_{u-}^{(k)}(j-1) - G_{u-}^{(k)}(j)) d\mathcal{L}_{il}(u) \\ &+ \int_{[0,s] \times [0,C] \times [0,1] \times \{\beta, \delta\}} (q(G_{u-}^{(k)}(j), G_{u-}^{(k-1)}, (\zeta_{u-}^{(k-1)} \vee \frac{1}{M}) \wedge M, z, w, \omega) - G_{u-}^{(k)}(j)) \mathcal{K}_j(du, d(z, w, \omega)). \end{aligned} \quad (50)$$

This has the following interpretation. While the type transport through the neutral lookdown events (given by the points of  $\mathcal{L}$ ) happens as usual, the activation levels (appearing in the update rules (9) and (10)) for the potential selective events (given by those points  $(u, (z, w, \omega))$  of  $\mathcal{K}$  with  $z \leq C$ ) are controlled by the mass process and the type frequencies *from the previous iteration*. Notice that  $G_{u-}^{(k)}(j)$ , that is the type *in the current iteration*, enters as the first argument in the update rule  $q$ , which is relevant at a competitive death event. This amounts to having a frozen environment for the competition. Also, notice that we use the mass process  $(\zeta_{u-}^{(k-1)} \vee \frac{1}{M}) \wedge M$  truncated at  $M$  and  $1/M$ . For the sequel, let us define the first time at which the truncation is effective:  $\sigma_M^{(\ell)} = \inf \left\{ s \geq 0, \zeta_s^{(\ell)} \in \left\{ \frac{1}{M}, M \right\} \right\}$ .

We now use (50) to successively update the types of  $\gamma \in \mathcal{R}_{>0} := \{\gamma = (s, i) \in \mathcal{R} \mid s > 0\}$  in the  $k$ -th iteration. This we do by first recording all the roots  $(s_0, i_0), \dots, (s_n, i_n) \in \mathcal{R}$  that lie on the neutral ancestral lineage of  $\gamma$ , with  $0 = s_0 < \dots < s_n = s$ . The type of  $(s_0, i_0)$  remains to be  $G_0(i_0)$ ; the new type  $G_{s_1}^{(k)}(i_1)$  of  $(s_1, i_1)$  is determined by taking  $G_0(i_0)$  as the first argument in the update rule  $q$ , the new type  $G_{s_2}^{(k)}(i_2)$  of  $(s_2, i_2)$  is determined by taking  $G_{s_1}^{(k)}(i_1)$  as the first argument of the update rule  $q$ , etc.

Having thus re-colored all  $\gamma \in \mathcal{R}_{>0}$  in the  $k$ -th iteration, we complete the recoloring by letting  $\Gamma_\gamma$  inherit the type of its root, i.e. by setting, for each  $(s, j) \in \mathbb{R}_+ \times \mathbb{N}$ , its type  $G_s^{(k)}(j)$  equal to the type of that  $\gamma$  for which  $(s, j) \in \Gamma_\gamma$ .

For  $s > 0$  and  $h \in \mathbb{I}$  we now put

$$\mu_s^{(k)} \{h\} := \sum_{\gamma=(u,i) \in \mathcal{R}: G_u^{(k)}(i)=h} \mathbf{m}_s(\Gamma_\gamma). \quad (51)$$

Recall (13):  $\mathbf{m}_s(\Gamma_\gamma)$  is the weight which the neutral sampling measure at time  $s$  assigns to that part of the neutral offspring of  $\gamma$  whose ancestral lineages are not separated from  $\gamma$  by some other root.

The next assertion, which will also be used in the uniqueness part (Step 4 of the proof of Proposition 4.1), and will therefore be singled out as a lemma, shows that  $G^{(k)}$  a.s. admits type frequencies at all times.

**Lemma 4.3.** In each iteration step  $k = 1, 2, \dots$  we have a.s.

$$\mu_s^{(k)} = \text{w-}\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \delta_{G_s^{(k)}(i)} \quad \text{for all } s > 0. \quad (52)$$

*Proof.* Let  $s, \epsilon > 0$  be arbitrarily fixed, and take  $h \in \mathbb{I}$ . On an a.s. event that does not depend on  $s$ , there exists by Lemma 3.5 a finite set  $\{\gamma_1, \dots, \gamma_\ell\} \subset \mathcal{R}$  of roots such that

$$\sum_{j=1}^{\ell} \mathbf{m}_s(\Gamma_{\gamma_j}) > 1 - \epsilon.$$

By (13) and Corollary 3.6, it follows that on an a.s. event that does not depend on  $s$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(s,i) \in \bigcup_{j=1}^{\ell} \Gamma_{\gamma_j}} > 1 - \epsilon.$$

For all iterations  $k$ , we have

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{G_s^{(k)}(i)=h} &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \left( \sum_{j=1}^{\ell} \mathbf{1}_{(s,i) \in \Gamma_{\gamma_j}, G_s^{(k)}(\gamma_j)=h} + \mathbf{1}_{(s,i) \notin \bigcup_{j=1}^{\ell} \Gamma_{\gamma_j}} \right) \\ &\leq \sum_{j=1}^{\ell} \mathbf{1}_{G_s^{(k)}(\gamma_j)=h} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(s,i) \in \Gamma_{\gamma_j}} + \epsilon \\ &\leq \mu_s^{(k)}\{h\} + \epsilon, \end{aligned}$$

where we used the definition (51) of  $\mu_s^{(k)}$ . Similarly,

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{G_s^{(k)}(i)=h} &\geq \sum_{j=1}^{\ell} \mathbf{1}_{G_s^{(k)}(\gamma_j)=h} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(s,i) \in \Gamma_{\gamma_j}} \\ &\geq \sum_{\gamma \in \mathcal{R}} \mathbf{1}_{G_s^{(k)}(\gamma)=h} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(s,i) \in \Gamma_{\gamma}} - \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(s,i) \notin \bigcup_{j=1}^{\ell} \Gamma_{\gamma_j}} \\ &\geq \mu_s^{(k)}\{h\} - \epsilon. \end{aligned}$$

□

Summarizing the results so far, we are able to define the values  $G_s^{(k)}(j)$  and  $\mu_s^{(k)}$  for  $j \in \mathbb{N}$  and  $s > 0$ . Using Equation (49) we can define  $\zeta_s^{(k)}$ ,  $s > 0$ .

**Step 3, Convergence of the iteration scheme:** For two type vectors  $g$  and  $g'$  admitting type frequencies we have (with  $\|\mu^g - \mu^{g'}\|$  denoting the total variation distance of  $\mu^g$  and  $\mu^{g'}$ )

$$\|\mu^g - \mu^{g'}\| \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{g(i) \neq g'(i)\}}. \quad (53)$$

Denoting by  $\lambda$  the ‘‘Lebesgue times uniform’’ measure on  $B_C := [0, C] \times [0, 1] \times \{\beta, \delta\}$ , we see from the definition of  $q$  in (9) and (10) that for all  $g, g'$  admitting type frequencies, all  $v, v' \geq 0$  and all  $h \in \mathbb{I}$  we have

$$\begin{aligned} & \int_{B_C} \mathbf{1}_{\{q(h,g,v,a) \neq q(h,g',v',a)\}} \lambda(da) \\ & \leq C(b|\mu^g\{A\}v - \mu^{g'}\{A\}v'| + c|\mu^g\{A\}v^2 - \mu^{g'}\{A\}(v')^2| + c|\mu^g\{B\}v^2 - \mu^{g'}\{B\}(v')^2|). \end{aligned} \quad (54)$$

Using (53), we infer that for all  $g, g'$  admitting type frequencies, all  $v, v' \geq 0$  and all  $h, h' \in \mathbb{I}$ ,

$$\begin{aligned} & \int_{B_C} \mathbf{1}_{\{q(h,g,(v \vee \frac{1}{M}) \wedge M, a) \neq q(h',g',(v' \vee \frac{1}{M}) \wedge M, a)\}} \lambda(da) \\ & \leq D \left( \mathbf{1}_{\{h \neq h'\}} + \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{g(i) \neq g'(i)\}} + |v - v'| \right), \end{aligned} \quad (55)$$

where the constant  $D$  may depend on  $M$  and  $C$ . (Note that this is an analogue of [12, (4.14)].)

Fix  $T \in \mathbb{R}_+$  and  $j \in \mathbb{N}$ . For  $0 \leq s \leq T$  and  $j \in \mathbb{N}$ , let us define by  $\mathcal{A}_s^T(j)$  the level of the ancestor of  $(T, j)$  at time  $s$  in the neutral genealogy. Note that  $\mathcal{A}^T(j)$  is  $\mathcal{L}$ -measurable, obeying the SDE

$$\begin{aligned} \mathcal{A}_s^T(j) &= j - \sum_{1 \leq i < \ell < j} \int_{(s,T]} \mathbf{1}_{\{\mathcal{A}_u^T(j) > \ell\}} d\mathcal{L}_{i\ell}(u) \\ &\quad - \sum_{1 \leq i < \ell \leq j} \int_{(s,T]} (\ell - i) \mathbf{1}_{\{\mathcal{A}_u^T(k) = \ell\}} d\mathcal{L}_{i\ell}(u). \end{aligned} \quad (56)$$

For  $s > 0$ , we abbreviate  $\tilde{G}_s^{T,k}(j) := G_s^{(k)}(\mathcal{A}_s^T(j))$ .

Let  $\tilde{\mathcal{K}}_j$  have an atom in  $(u, a)$  if and only if  $\mathcal{K}_{\mathcal{A}_u^T(j)}$  has an atom in  $(u, a)$ ,  $0 \leq u \leq T$ ,  $a \in B_C$ . For notational reasons we now consider the induction step from  $k$  to  $k+1$  instead of  $k-1$  to  $k$ . We get as an analogue to the estimate starting at p. 1112 line -3 in [12], that for  $s > 0$ :

$$\begin{aligned} & \mathbf{1}_{\{\tilde{G}_{s-}^{T,k+1}(j) \neq \tilde{G}_{s-}^{T,k}(j)\}} \leq \int_{[0,s] \times B_C} (1 - \mathbf{1}_{\{\tilde{G}_{u-}^{T,k+1}(j) \neq \tilde{G}_{u-}^{T,k}(j)\}}) \\ & \quad \times \mathbf{1}_{\{q(\tilde{G}_{u-}^{T,k}(j), G_{u-}^{(k)}, (\zeta_{u-}^{(k)} \vee \frac{1}{M}) \wedge M, a) \neq q(\tilde{G}_{u-}^{T,k}(j), G_{u-}^{(k-1)}, (\zeta_{u-}^{(k-1)} \vee \frac{1}{M}) \wedge M, a)\}} \tilde{\mathcal{K}}_j(du, da) \\ & = \int_{[0,s] \times B_C} (1 - \mathbf{1}_{\{\tilde{G}_{u-}^{T,k+1}(j) \neq \tilde{G}_{u-}^{T,k}(j)\}}) \\ & \quad \times \mathbf{1}_{\{q(\tilde{G}_{u-}^{T,k}(j), G_{u-}^{(k)}, (\zeta_{u-}^{(k)} \vee \frac{1}{M}) \wedge M, a) \neq q(\tilde{G}_{u-}^{T,k}(j), G_{u-}^{(k-1)}, (\zeta_{u-}^{(k-1)} \vee \frac{1}{M}) \wedge M, a)\}} (\tilde{\mathcal{K}}_j(du, da) - du \lambda(da)) \\ & \quad + \int_{[0,s] \times B_C} (1 - \mathbf{1}_{\{\tilde{G}_{u-}^{T,k+1}(j) \neq \tilde{G}_{u-}^{T,k}(j)\}}) \\ & \quad \times \mathbf{1}_{\{q(\tilde{G}_{u-}^{T,k}(j), G_{u-}^{(k)}, (\zeta_{u-}^{(k)} \vee \frac{1}{M}) \wedge M, a) \neq q(\tilde{G}_{u-}^{T,k}(j), G_{u-}^{(k-1)}, (\zeta_{u-}^{(k-1)} \vee \frac{1}{M}) \wedge M, a)\}} du \lambda(da) \\ & \leq \text{martingale} + D \int_0^s \eta_u^{(k)} du + D \int_0^s |(\zeta_u^{(k)} \vee \frac{1}{M}) \wedge M - (\zeta_u^{(k-1)} \vee \frac{1}{M}) \wedge M| du \end{aligned} \quad (57)$$

where

$$\eta_u^{(k)} := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{G_{u-}^{(k)}(i) \neq G_{u-}^{(k-1)}(i)\}} \quad (58)$$

and where we used (55) in the last estimate. To see that the limit exists we can argue as in the proof of Lemma 4.3. Indeed, the limit equals the sum of the masses at time  $u$  of the fragments whose roots are colored differently at iterations  $k$  and  $k - 1$ . Taking expectations of both sides in the above estimate, putting  $s := T$  and noting that  $\tilde{G}_T^{T,k}(j) = G_T^{(k)}(j)$ , we obtain the estimate

$$\mathbf{P}(G_T^{(k+1)}(i) \neq G_T^{(k)}(i)) \leq D \left( \mathbf{E} \left( \int_0^T \eta_u^{(k)} du \right) + \mathbf{E} \left( \int_0^T |(\zeta_u^{(k)} \vee \frac{1}{M}) \wedge M - (\zeta_u^{(k-1)} \vee \frac{1}{M}) \wedge M| du \right) \right) \quad (59)$$

which in turn implies

$$\mathbf{E}(\eta_T^{(k+1)}) \leq D \left( \int_0^T \mathbf{E}(\eta_u^{(k)}) du + \int_0^T \mathbf{E}(|(\zeta_s^{(k)} \vee \frac{1}{M}) \wedge M - (\zeta_s^{(k-1)} \vee \frac{1}{M}) \wedge M|) ds \right) \quad (60)$$

by dominated convergence. Equation (60) and Lemma 4.2 give that

$$\mathbf{E}(\eta_T^{(k+1)}) \leq D \left( \int_0^T \mathbf{E}(\eta_u^{(k)}) du + \tilde{C}T e^T \int_0^T \int_0^s \mathbf{E}(|\mu_u^{(k)}\{A\} - \mu_u^{(k-1)}\{A\}|) du ds \right).$$

Using (53) we arrive at

$$\mathbf{E}(\eta_T^{(k+1)}) \leq D \int_0^T (1 + \tilde{C}T^2 e^T) \mathbf{E}(\eta_u^{(k)}) du \quad (61)$$

for all  $T \geq 0$ . By a direct reiteration, this gives:

$$\begin{aligned} \mathbf{E}(\eta_T^{(k+1)}) &\leq D(1 + \tilde{C}T^2 e^T) \int_0^T \left[ D(1 + \tilde{C}s_1^2 e^{s_1}) \int_0^{s_1} \mathbf{E}(\eta_{s_2}^{(k-1)}) ds_2 \right] ds_1 \\ &\leq (D(1 + \tilde{C}T^2 e^T))^2 \int_0^T \int_0^{s_1} \mathbf{E}(\eta_{s_2}^{(k-1)}) ds_2 ds_1 \\ &\leq (D(1 + \tilde{C}T^2 e^T))^k \int_0^T \int_0^{s_1} \cdots \int_0^{s_{k-1}} \mathbf{E}(\eta_{s_k}^{(1)}) ds_k \cdots ds_1. \end{aligned}$$

Because  $\mathbf{E}(\eta_s^{(1)})$  is uniformly bounded by 1, we obtain that:

$$\mathbf{E}(\eta_T^{(k+1)}) \leq (D(1 + \tilde{C}T^2 e^T))^k \frac{T^k}{k!}. \quad (62)$$

Combining (53), (58) and (62) we infer that for all  $s \in [0, T]$

$$\mathbf{E}[|\mu_s^{(k+1)} - \mu_s^{(k)}|] \text{ is summable over } k. \quad (63)$$

Moreover, from Lemma 4.2, (49) and (63) we conclude that

$$\sup_{s \leq T} \mathbf{E}[|(\zeta_s^{(k)} \vee \frac{1}{M}) \wedge M - (\zeta_s^{(k-1)} \vee \frac{1}{M}) \wedge M|] \text{ is summable over } k. \quad (64)$$

From (59), Fubini, (62) and (64), we conclude that for all  $s \in [0, T]$  and  $i \in \mathbb{N}$

$$\mathbf{P}(G_s^{(k+1)}(i) \neq G_s^{(k)}(i)) \text{ is summable over } k. \quad (65)$$

Hence for  $\epsilon > 0$  arbitrary and each finite subset  $\mathcal{M} \subset [0, T] \times \mathbb{N}$ , we have by Borel-Cantelli that there exists  $k_0 \in \mathbb{N}$  with

$$\mathbf{P}(G_s^{(k)}(i) = G_s^{(k_0)}(i) \text{ for all } k \geq k_0 \text{ and } (s, i) \in \mathcal{M}) \geq 1 - \epsilon. \quad (66)$$

Now take  $\ell$  as in Lemma 3.8 and choose the (nonrandom) finite set  $\mathcal{M} \subset [0, T] \times \mathbb{N}$  so large and “dense” that with probability  $1 - \epsilon$  every fragment  $\Gamma_\gamma$  whose root  $\gamma$  is an element of  $[0, T] \times \{1, \dots, \ell\}$ , contains an element of  $\mathcal{M}$ . Lemma 3.8 together with (66) imply that

$$\mathbf{P}(\sup_{s \leq T} \|\mu_s^{(k)} - \mu_s^{(k_0)}\| \geq \epsilon) \leq 2\epsilon \text{ for all } k \geq k_0. \quad (67)$$

From (64), (67), (65) and (46) and the choice of  $\mathcal{M}$  we infer that  $((G_s^{(k)}(i))_{i \in \mathbb{N}}, \mu_s^{(k)}, (\zeta_s^{(k)} \vee \frac{1}{M}) \wedge M)$  converges uniformly in  $s \in [0, T]$  as  $k \rightarrow \infty$ . In order to see that the limit satisfies (44) we recall that the Poisson point measures  $\mathcal{L}$  and  $\mathcal{K}$  do not change over the iterations, and note that the distribution of the mark  $z$  which figures in (44) and (50) is continuous, which provides the adequate continuity in the coefficient  $q$  that is given by the activation conditions (9) and (10). Theorem (6.4) of [37] shows that the limit also satisfies (45).

**Step 4, Uniqueness:** The argument from Lemma 4.3 shows that for any solution of (44), (45),  $G_s$  admits type frequencies for all  $s \geq 0$  a.s. Uniqueness of the strong solution of (44), (45) follows by the same argument as in Step 3 where we now compare in (57) two solutions instead of two approximations (Note that this strategy was also successful in the simpler setting of [13]). This concludes the proof of Proposition 4.1.  $\square$

For the *completion of the proof of Theorem 2.2* let us now relax the control of the total mass with the constant  $M$ . Fix again the constant  $C \in (0, \infty)$  and consider the following system of SDEs

$$\begin{aligned} G_s(j) &= G_0(j) + \sum_{i=1}^{j-1} \int_{[0,s]} (G_{u-}(i) - G_{u-}(j)) d\mathcal{L}_{ij}(u) \\ &+ \sum_{1 \leq i < k < j} \int_{[0,s]} (G_{u-}(j-1) - G_{u-}(j)) d\mathcal{L}_{ik}(u) \\ &+ \int_{[0,s] \times [0,C] \times [0,1] \times \{\beta, \delta\}} (q(G_{u-}(j), G_{u-}, \zeta_{u-}, z, w, \omega) - G_{u-}(j)) \mathcal{K}_j(du, d(z, w, \omega)), \\ \zeta_0 &= v_0, \quad d\zeta_s = \zeta_s f(\zeta_s, \mu^{G_s}\{A\}) ds + \zeta_s d\mathcal{W}_s, \quad s \geq 0. \end{aligned} \quad (68)$$

Proposition 4.1 tells us that this system has for each  $M$  a unique pathwise solution up to the stopping time  $\sigma_M$  defined by (25). By projectivity, this shows that (68) has a unique pathwise solution up to the time at which its mass process  $\zeta$  goes to extinction or explodes. In view of (9) and (10), the solution of (11), (7) coincides with that of (68) up to that time at which  $\zeta$  exceeds  $\sqrt{\frac{C}{c}}$  or  $\frac{C}{b}$ . Again by projectivity, this implies the assertion of Theorem 2.2.  $\square$

## 5 From the neutral to the selective genealogy

For an (exchangeable) initial type configuration  $G_0 \in \mathbb{I}^{\mathbb{N}}$  and the independent stochastic input  $(\mathcal{W}, \mathcal{L}, \mathcal{K})$  specified in (I1), (I2), (I3) in Section 2, Theorem 2.2 provides an a.s. unique solution  $(\zeta, G) = (\zeta_s, G_s)_{s \geq 0}$  of (7) and (11). From this lookdown representation we will construct in Sec. 5.1 the process  $(\zeta, X) = (\zeta_s, R_s, G_s)_{s \geq 0}$  of type configurations and genealogical distance matrices, which will be turned in Sec. 5.3 into the process  $(\xi_t, Y_t)_{t \geq 0}$  of isomorphy classes of marked metric measure spaces that describe type distributions and sample genealogies. As will be proved in Sec. 5.2, the latter will provide the unique solution to the martingale problem formulated in Prop. 2.7. We recall that we always assume that  $(R_0, G_0)$  has the marked distance matrix distribution of a marked ultrametric measure space.

## 5.1 The selective lookdown genealogy

In this subsection we define the *selective lookdown space* for fixed  $M \geq 1$ . As in (26) we set  $C := C_M = (b \vee c)M^2$ . With regard to (9), (10) and (11) we say that a point  $(s, i) \in \mathbb{R}_+ \times \mathbb{N}$  is *active* if

$$\begin{aligned} \mathcal{K}_i \text{ has an atom in } (s, (z, w, \beta)) \text{ for some } z \in \mathbb{R}_+, w \in [0, 1] \text{ such that} \\ z \leq (b\mu^{G_s}\{A\}\zeta_s) \wedge C \end{aligned}$$

or if

$$\begin{aligned} \mathcal{K}_i \text{ has an atom in } (s, (z, w, \delta)) \text{ for some } z \in \mathbb{R}_+, w \in [0, 1] \text{ such that} \\ z \leq (c\mu^{G_s}\{B\}(\zeta_s)^2) \wedge C \quad \text{and} \quad G_{s-}(i) = A, \text{ or} \\ z \leq (c\mu^{G_s}\{A\}(\zeta_s)^2) \wedge C \quad \text{and} \quad G_{s-}(i) = B. \end{aligned}$$

In the first case we say that a fecundity event takes place at  $(s, i)$ , in the second case we say that a competition event happens at  $(s, i)$ . Note that because  $s \mapsto \zeta_s$  and  $s \mapsto \mu^{G_s}$  are a.s. continuous (see Prop. 2.9), we can as well replace  $s$  by  $s-$  in the three inequalities.

Next we define the *selective ancestral lineage* of an element  $\theta \in \bigcup_\gamma \Gamma_\gamma \subset Z$ . For this we trace the lineage of  $\theta$  back into the past according to Remark 3.3, until it hits an active point  $(s, i) \in \mathbb{R}_+ \times \mathbb{N}$ . If a competition event occurs at that point, then we continue the lineage at an element of  $Z$  picked independently according to the neutral sampling measure  $\mathbf{m}_s$  defined by (13). If a fecundity event happens at  $(s, i)$ , then we continue the lineage at an element of  $Z$  picked independently according to  $\mathbf{m}_s$  conditioned on the fragments of type  $A$ . (For the definition of the fragments  $\Gamma_\gamma$ , we refer back to Sec 3.2). The individuals on the selective ancestral lineage of  $\theta$  will be called the *selective ancestors* of  $\theta$ .

We now modify the metric  $\rho^{(0)}$  from Section 3.1 as follows: Trace the selective ancestral lineages back from  $(s_1, i_1)$  and  $(s_2, i_2)$ . If they merge at time  $u \geq 0$ , then put  $\rho((s_1, i_1), (s_2, i_2)) := (s_1 - u) + (s_2 - u)$ . Otherwise, if  $a_1$  and  $a_2$  are the labels of the two (selective) ancestors at time 0, then define  $\rho((s_1, i_1), (s_2, i_2)) := s_1 + s_2 + R_0(a_1, a_2)$ , where  $R_0$  is the semi-ultrametric on  $\mathbb{N}$  given by the initial condition. We define  $\hat{Z}$  as the completion of  $\bigcup_\gamma \Gamma_\gamma$  with respect to  $\rho$  and call  $(\hat{Z}, \rho)$  the *selective lookdown space*. (Note that  $(\hat{Z}, \rho)$  depends on  $M$  but the projective limit of the restrictions to the completions of  $[0, \sigma_M] \times \mathbb{N}$  with  $\sigma_M$  defined in (25) exists.)

The selective lookdown space  $(\hat{Z}, \rho)$  inherits the family of *sampling measures*  $\mathbf{m}_s$ ,  $s > 0$  from (13), which remain probability measures by Lemma 3.5. As each fragment  $\Gamma_\gamma$  is monotypic, we can also endow  $\hat{Z} \times \mathbb{I}$  with the sampling measures  $m_s$ , defined by

$$m_s(E \times E') = \sum_{\gamma=(u,i) \text{ with } G_u(i) \in E'} \mathbf{m}_s(\Gamma_\gamma \cap E), \quad E' \subset \mathbb{I}, E \subset \hat{Z} \text{ measurable.} \quad (69)$$

**Lemma 5.1.** The measures  $m_s$  defined in (69) obey

$$m_s = \text{w-} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \delta_{((s,i), G_s(i))} \quad (70)$$

for all  $s > 0$  on an event of probability 1 that does not depend on  $s$ . Here, the weak limit in (70) can be understood either with respect to the metric  $\rho_0$  or with respect to the metric  $\rho$ .

*Proof.* This follows directly from Lemma 3.8 and Corollary 3.6. □

With regard to (4), which defines the mapping  $s \mapsto t(s)$ , we define the time-changed distance  $\tilde{\rho}$  as follows: For  $(s_1, i_1)$  and  $(s_2, i_2)$  that lie on the same selective ancestral lineage in the lookdown graph, we put

$$\tilde{\rho}((s_1, i_1), (s_2, i_2)) := |t(s_1) - t(s_2)| = \left| \int_{s_1}^{s_2} \zeta_u du \right|;$$

this is the time it takes from one point to the other when traveling along the selective ancestral lineage with speed  $1/\zeta_u$  (cf. (3)) at an intermediate point  $(u, j)$ . More generally we put

$$\tilde{\rho}((s_1, i_1), (s_2, i_2)) := \tilde{\rho}((s_1, i_1), (s', j)) + \tilde{\rho}((s_2, i_2), (s', j))$$

if the two selective ancestral lineages merge at some point  $(s', j)$  with  $s' \in [0, s]$ ; otherwise, if  $(s_1, i_1), (s_2, i_2)$  have two distinct selective ancestors  $(0, a_1)$  and  $(0, a_2)$  at time 0, we put

$$\tilde{\rho}((s_1, i_1), (s_2, i_2)) := \tilde{\rho}((s_1, i_1), (0, a_1)) + \tilde{\rho}((s_2, i_2), (0, a_2)) + R_0(a_1, a_2).$$

For  $s \geq 0$  and  $i_1, i_2 \in \mathbb{N}$  we set

$$R_s(i_1, i_2) := \tilde{\rho}((s, i_1), (s, i_2)).$$

In other words,

$$R_s(i_1, i_2) := \begin{cases} 2(t(s) - t(s')) & \text{if the two selective ancestral lineages merge at time } s' \in [0, s], \\ 2t(s) + R_0(a_1, a_2) & \text{if the selective ancestors } (0, a_1), (0, a_2) \text{ at time 0 are different.} \end{cases}$$

**Corollary 5.2.** (i) The weak limit in (70) can be understood also with respect to the metric  $\tilde{\rho}$ .

(ii)  $(R_s^{(0)}, G_s)$ ,  $(R_{s-}, G_{s-})$  as well as  $(R_s, G_s)$ ,  $s > 0$  are proper on an event of probability 1 that does not depend on  $s$ .

*Proof.* This is a consequence of (69), (70) and Lemma 3.1. □

Let us also prove that:

**Proposition 5.3.** For each  $s \geq 0$ , the pair  $(R_s, G_s)$  is exchangeable conditionally given  $(\zeta_u)_{u \leq s}$ .

*Proof.* It suffices to work along the sequence of jump times of the restriction of the process  $(R_s, G_s)_{s \geq 0}$  of marked distance matrices to the first  $n$  levels, where  $n \in \mathbb{N}$  is arbitrarily fixed. Between these jump times, the types of the individuals on the first  $n$  levels remain unchanged, and the genealogical distance between each pair of such individuals grows deterministically with slope 2. The jump times of this process are given by the Poisson processes  $(\mathcal{L}_{ij})_{1 \leq i < j \leq n}$  and  $(\mathcal{K}_i)_{1 \leq i \leq n}$ . Let us also recall that  $(R_0, G_0)$  is exchangeable by the assumption that it has the marked distance matrix distribution of a marked ultrametric measure space.

(i) We first consider the jumps given by the neutral events, i.e. by the processes  $(\mathcal{L}_{ij})_{1 \leq i < j \leq n}$ . We assume that the restriction of the process  $(R_u, G_u)_{u \geq 0}$  to the first  $n$  levels jumps at some time  $s$  due to a neutral reproduction event.

Proceeding inductively we assume that the restriction of  $(R_{s-}, G_{s-})$  to the first  $n$  levels is exchangeable. Consequently,  $((R(\ell, m))_{1 \leq \ell, m \leq n}, (G(\ell))_{1 \leq \ell \leq n}) := ((R_{s-}(\ell, m))_{1 \leq \ell, m \leq n}, (G_{s-}(\ell))_{1 \leq \ell \leq n})$  has the same distribution as

$$(\tilde{R}, \tilde{G}) := ((R(\Pi(\ell), \Pi(m)))_{1 \leq \ell, m \leq n}, (G(\Pi(\ell)))_{1 \leq \ell \leq n}), \tag{71}$$

where  $\Pi$  is a uniformly distributed random pick from the permutations of  $[n]$ . Let  $(I, J)$  be uniformly distributed on  $\{(i, j) : 1 \leq i < j \leq n\}$ , and let, given  $\{(I, J) = (i, j)\}$ , the random array  $(\tilde{R}', \tilde{G}')$  be constructed from  $(\tilde{R}, \tilde{G})$  according to (17) and (18).

Putting

$$f_j(m) := \begin{cases} \ell & \text{if } \ell < m \\ \ell - 1 & \text{if } j < m \leq n \end{cases}$$

we can write  $\tilde{R}'$  as

$$\tilde{R}'(\ell, m) = \begin{cases} 0 & \text{if } \ell, m \in \{I, J\}, \\ \tilde{R}(I, f_J(\ell)) & \text{if } \ell \in \{I, J\} \text{ and } \ell \in [n] \setminus \{I, J\}, \\ \tilde{R}(f_J(\ell), f_J(m)) & \text{if } \ell, m \in [n] \setminus \{I, J\}. \end{cases} \quad (72)$$

Using (71) one checks readily that, for each permutation  $\sigma$  of the numbers  $1, \dots, n$ , the random array  $((\tilde{R}'(\sigma(\ell), \sigma(m)))_{1 \leq \ell, m \leq n}, (\tilde{G}'(\sigma(\ell)))_{1 \leq \ell \leq n})$  has the same distribution as  $(\tilde{R}'(\ell, m))_{1 \leq \ell, m \leq n}, \tilde{G}'(\ell)_{1 \leq \ell \leq n}$ , which gives the desired exchangeability of  $(R_s, G_s)$ .

(ii) We now turn to the non-neutral events. Let  $u$  be a time point at which one of the counting measures  $\mathcal{K}_i(\cdot \times [0, C] \times [0, 1] \times \{\beta, \delta\})$ ,  $1 \leq i \leq n$ , has an atom for which the corresponding activation condition on the r.h.s. of (9) resp. (10) is satisfied for  $v = \zeta_u$ ,  $g = G_{u-}$  and  $h = G_{u-(i)}$ . Making use of part (i) and proceeding by induction, we assume that the random array  $(R_{u-}^{(0)}(\ell, m), G_{u-}(\ell))_{1 \leq \ell, m \leq n}$  is exchangeable given  $(\zeta_w^{(k-1)})_{w \leq u}$ . Let  $i$  be that element of  $\{1, \dots, n\}$  for which  $\mathcal{K}_i(\{u\} \times [0, C] \times [0, 1] \times \{\beta, \delta\}) = 1$ ; because all the Poisson point measures  $\mathcal{K}_i$  have the same intensity given  $\zeta$ , the level  $i$  is uniformly chosen from  $\{1, \dots, n\}$ . According to the update rule (21), (23), conditionally given  $(\zeta_w)_{w \leq u}$ , the exchangeability of the restriction of  $(R_{u-}, G_{u-})$  to the first  $n$  levels propagates to the exchangeability of the restriction of  $(R_u, G_u)$  to the first  $n$  levels.  $\square$

By letting  $M \rightarrow \infty$  we have thus constructed the process  $(\zeta, X) := (\zeta, R, G)$  up to the time  $\sigma$  defined in Theorem 2.2.

We remark that by using the sampling measures  $\mathbf{m}_s$  for the independent picks needed to continue the ancestral lineage at competitive and selective events, we avoid the formalism of genetic markers which is used in Section 6 of [12] to trace ancestral lineages.

The above construction and Theorem 2.2 show that the process  $(\zeta, X)$  is the pathwise unique solution the system of SDEs given by (7), (11) and (24), which is driven by  $(\mathcal{L}_{ij}), (\mathcal{K}_i), \mathcal{W}$  and extends (7), (11) to include also the genealogical distances.

## 5.2 Two well-posed martingale problems in the lockdown framework

The aim of this subsection is prove Proposition 2.6, and thus to establish a well-posed martingale problem for the (suitably stopped) process  $(\zeta, R, G)$ . For this, we follow the strategy outlined at the end of Sec. 2.3 (right after the statement of Proposition 2.6). Let  $(\mathcal{W}, \mathcal{L}, \mathcal{K})$  be as in Sec. 2.3.2, choose some  $M > 0$  and let  $(\zeta, R, G)$  be the unique strong solution of the system of SDE's (7), (11), (24). With  $C_M$  defined in (26), we set

$$\mathbb{B}_M := [0, C_M] \times [0, 1] \times \{\beta, \delta\} \quad (73)$$

and  $\mathcal{B}_M :=$  the  $\sigma$ -algebra of Borel-sets on  $\mathbb{B}_M$ . We are now going to define the family of counting processes  $(\Lambda_s^M)$  that will figure as a fourth component (in addition to  $(\zeta, R, G)$ ) in the first one of our two martingale problems. The process  $(\Lambda_s^M)$  takes its values in the set  $\mathcal{N}_M$  of finite counting measures on  $\{(i, j) : 1 \leq i < j < \infty\} \cup (\bigcup_{i \in \mathbb{N}} \{i\} \times \mathbb{B}_M)$  and is defined by

$$\Lambda_s^M(\{(i, j)\}) := \mathcal{L}_{ij}([0, s]), \quad 1 \leq i < j < \infty, \quad (74)$$

$$\Lambda_s^M(\{i\} \times H) := \mathcal{K}_i([0, s] \times H), \quad i \in \mathbb{N}, H \in \mathcal{B}_M. \quad (75)$$

Thus the process  $\Lambda^M$  jumps at time  $s$  if and only if either the process  $\mathcal{K}$  (restricted to  $\mathbb{R}_+ \times \bigcup_{i \in \mathbb{N}} (\{i\} \times \mathbb{B}_M)$ ) or the process  $\mathcal{L}$  has an atom whose time component is  $s$ . To prepare for a martingale problem for  $(\zeta, R, G, \Lambda^M)$  stopped at  $\sigma_M$ , we define the state space

$$\widehat{E}_M := \left( \left( \frac{1}{M}, M \right) \times \mathbb{R}^{\mathbb{N}^2} \times \mathbb{I}^{\mathbb{N}} \times \mathcal{N}_M \right) \cup \{\Delta_M\} \quad (76)$$

where  $\left( \frac{1}{M}, M \right) \times \mathbb{R}^{\mathbb{N}^2} \times \mathbb{I}^{\mathbb{N}} \times \mathcal{N}_M$  is equipped with the product topology and a sequence  $(v_n, r_n, g_n, \lambda_n)$  is said to converge to  $\Delta_M$  if either  $v_n \rightarrow \frac{1}{M}$  or  $v_n \rightarrow M$  as  $n \rightarrow \infty$ . Note that the space (76) is an extension of (27).

Next we display the generator of  $(\zeta, R, G, \Lambda^M)$  restricted to appropriate test functions  $F = F(v, r, g, \lambda)$ , where  $v \in \mathbb{R}_+$ ,  $r \in \mathbb{R}^{\mathbb{N}^2}$ ,  $g \in \mathbb{I}^{\mathbb{N}}$  and  $\lambda \in \mathcal{N}_M$ . For  $\lambda \in \mathcal{N}_M$  we define

$$\begin{aligned} \varphi_{ij}(\lambda) &:= (-1)^{\lambda(\{(i,j)\})}, & 1 \leq i < j < \infty, \\ \varphi_i^H(\lambda) &:= (-1)^{\lambda(\{i\} \times H)}, & i \in \mathbb{N}, H \in \mathcal{B}_M. \end{aligned} \quad (77)$$

We note that  $\varphi_{ij}(\Lambda_s^M)$  jumps from  $+1$  to  $-1$  (or vice versa) whenever  $\mathcal{L}_{ij}$  has an atom whose time component is  $s$ .

With  $D_{1,M}$  and  $D_{2,M}$  as in Section 2.3.5, let  $D_{3,M}$  be the set of those functions  $\varphi : \mathcal{N}_M \rightarrow \mathbb{R}$  which are of the form

$$\varphi = \prod_{\substack{k \in L' \\ V \in \mathcal{V}}} \varphi_k^V \prod_{(i,j) \in L} \varphi_{ij} \quad (78)$$

for some  $n \in \mathbb{N}$ ,  $L \subset \{(i, j) : 1 \leq i < j \leq n\}$ ,  $L' \subset \{1, \dots, n\}$ ,  $\mathcal{V} \in \mathcal{B}_M$  finite.

We now consider functions of the form

$$F(v, r, g, \lambda) = f(v, r) \gamma(g) \varphi(\lambda) \quad (79)$$

where  $f \in D_{1,M}$ ,  $\gamma \in D_2$  and  $\varphi \in D_{3,M}$ . The smallest possible  $n \in \mathbb{N}$  which fits to the required representations of  $f$ ,  $\gamma$  and  $\varphi$  will be called the *degree* of  $F$ . We write  $F_{r(i,j)}$  for the partial derivative of  $F$  with respect to the variable  $r(i, j)$ , and  $F_v$  for partial derivative of  $F$  with respect to  $v$ .

Let  $\vartheta_{i,j}$  and  $\tilde{\vartheta}_{j,\theta,h'}$  be as in (17), (18), (19) and (20). Let  $F$  be as in (79) with degree  $n$ . For a proper pair  $(r, g)$ , let  $\kappa_{\mathfrak{m}^{r,g}}$  be a measurable mapping defined on  $[0, 1]$  that transports the uniform distribution on  $[0, 1]$  into the sampling measure  $\mathfrak{m}^{r,g}$ . Writing  $\mu^g$  for the second marginal of  $\mathfrak{m}^{r,g}$  and putting  $(\theta(w), h'(w)) := \kappa_{\mathfrak{m}^{r,g}}(w)$ ,  $w \in [0, 1]$ , we put for all  $v > 0$ , all  $\lambda \in \mathcal{N}_M$ , and all proper pairs  $(r, g)$

$$\begin{aligned}
\widehat{\mathbf{A}}F(v, r, g, \lambda) &= \frac{v^2}{2}F_{vv}(v, r, g, \lambda) + (bv^2\mu^g\{A\} - 2cv^3\mu^g\{A\}\mu^g\{B\})F_v(v, r, g, \lambda) \\
&+ 2v \sum_{1 \leq i \neq j \leq n} F_{r(i,j)}(v, r, g, \lambda) \\
&+ \sum_{1 \leq i < j \leq n} ((-1)^{\mathbf{1}_{L((i,j))}} F(v, \vartheta_{i,j}(r, g), \lambda) - F(v, r, g, \lambda)) \\
&+ \sum_{j=1}^n \mathbf{1}_{\{g(j)=B\}} \int_0^{cv^2\mu^g\{A\}} dz \int_0^1 dw \left( (-1)^{\mathbf{1}_{\{k\} \times v((j,z,w))}} F(v, \tilde{\vartheta}_{j,\theta(w),h'(w)}(r, g), \lambda) - F(v, r, g, \lambda) \right) \\
&+ \sum_{j=1}^n \mathbf{1}_{\{g(j)=B\}} \int_{cv^2\mu^g\{A\}}^{C_M} dz \int_0^1 dw \left( (-1)^{\mathbf{1}_{\{k\} \times v((j,z,w))}} F(v, r, g, \lambda) - F(v, r, g, \lambda) \right) \\
&+ \sum_{j=1}^n \mathbf{1}_{\{g(j)=A\}} \int_0^{cv^2\mu^g\{B\}} dz \int_0^1 dw \left( (-1)^{\mathbf{1}_{\{k\} \times v((j,z,w))}} F(v, \tilde{\vartheta}_{j,\theta(w),h'(w)}(r, g), \lambda) - F(v, r, g, \lambda) \right) \\
&+ \sum_{j=1}^n \mathbf{1}_{\{g(j)=A\}} \int_{cv^2\mu^g\{B\}}^{C_M} dz \int_0^1 dw \left( (-1)^{\mathbf{1}_{\{k\} \times v((j,z,w))}} F(v, r, g, \lambda) - F(v, r, g, \lambda) \right) \\
&+ \sum_{j=1}^n \int_0^{bv\mu^g\{A\}} dz \int_0^1 dw \left( (-1)^{\mathbf{1}_{\{k\} \times v((j,z,w))}} F(v, \tilde{\vartheta}_{j,\theta(w),h''(w)}(r, g), \lambda) - F(v, r, g, \lambda) \right) \\
&+ \sum_{j=1}^n \int_{bv\mu^g\{A\}}^{C_M} dz \int_0^1 dw \left( (-1)^{\mathbf{1}_{\{k\} \times v((j,z,w))}} F(v, r, g, \lambda) - F(v, r, g, \lambda) \right)
\end{aligned} \tag{80}$$

Let  $\widehat{D}_M$  be the linear span of the constant real-valued functions on  $\widehat{E}_M$  and all functions of the form (79), and denote the extension of (80) to  $\widehat{D}_M$  again by  $\widehat{\mathbf{A}}$ .

**Proposition 5.4.** The process  $(\zeta_{s \wedge \sigma_M}, R_{s \wedge \sigma_M}, G_{s \wedge \sigma_M}, \Lambda_{s \wedge \sigma_M}^M)_{s \geq 0}$  solves the martingale problem  $(\widehat{\mathbf{A}}, \widehat{D}_M)$ , and this martingale problem is well-posed.

*Proof.* a) For all  $F \in \widehat{D}_M$ ,

$$F(\zeta_{s \wedge \sigma_M}, R_{s \wedge \sigma_M}, G_{s \wedge \sigma_M}, \Lambda_{s \wedge \sigma_M}^M) - \int_0^{s \wedge \sigma_M} \widehat{\mathbf{A}}F(\zeta_u, R_u, G_u, \Lambda_u^M) du, \quad s \geq 0, \tag{81}$$

is a martingale by Itô's formula, since  $(\zeta_{s \wedge \sigma_M}, R_{s \wedge \sigma_M}, G_{s \wedge \sigma_M}, \Lambda_{s \wedge \sigma_M}^M)_{s \geq 0}$  obeys (7), (11), (24) and the SDE for  $\Lambda^M$  driven by  $(\mathcal{K}_i)$ .

b) Conversely, from any solution  $(\widehat{\zeta}_{s \wedge \widehat{\sigma}_M}, \widehat{R}_{s \wedge \widehat{\sigma}_M}, \widehat{G}_{s \wedge \widehat{\sigma}_M}, \widehat{\Lambda}_{s \wedge \widehat{\sigma}_M}^M)_{s \geq 0}$  to the martingale problem  $(\widehat{\mathbf{A}}, \widehat{D}_M)$  we can extract, up to the stopping time  $\widehat{\sigma}_M$  (which is defined as in (25) but now for  $\widehat{\zeta}$  instead of  $\zeta$ ) a Poisson point process  $\widehat{\mathcal{L}}$  on  $\mathbb{R}_+ \times \{(i, j) : 1 \leq i < j < \infty\}$  and a Poisson point process  $\widehat{\mathcal{K}}$  on  $\mathbb{R}_+ \times \mathbb{N} \times [0, M] \times [0, 1] \times \{\beta, \delta\}$ , such that  $(\widehat{\mathcal{L}}, \widehat{\mathcal{K}})$  is equal in distribution to the corresponding restriction of  $(\mathcal{L}, \mathcal{K})$ . We can then extract from  $(\widehat{\zeta}, \widehat{G}, \widehat{\mathcal{L}}, \widehat{\mathcal{K}})$  also a Brownian motion  $\widehat{\mathcal{W}}$  up to the stopping time  $\widehat{\sigma}_M$ . Taking  $(\widehat{\mathcal{W}}, \widehat{\mathcal{L}}, \widehat{\mathcal{K}})$  as the source of randomness in the system given by (7), (11), (24) and in the definition of  $\Lambda^M$ , we infer from the pathwise uniqueness of that system that

$$(\widehat{\zeta}_{s \wedge \widehat{\sigma}_M}, \widehat{R}_{s \wedge \widehat{\sigma}_M}, \widehat{G}_{s \wedge \widehat{\sigma}_M}, \widehat{\Lambda}_{s \wedge \widehat{\sigma}_M}^M)_{s \geq 0} \stackrel{d}{=} (\zeta_{s \wedge \sigma_M}, R_{s \wedge \sigma_M}, G_{s \wedge \sigma_M}, \Lambda_{s \wedge \sigma_M}^M)_{s \geq 0}, \tag{82}$$

as asserted.  $\square$

Now we turn to the completion of the proof of Proposition 2.6, by establishing a well-posed martingale problem for  $(\zeta, R, G)$ . Recall from Sec. 2.3.4 that the mapping  $w \mapsto (\theta(w), h'(w))$  which appears in (80) is chosen such that, given  $(R_{s-}, G_{s-}) = (r, g)$ , it transports the uniform distribution on  $[0, 1]$  into the sampling measure  $m^{r,g}$ . Thus, for functions  $F$  that are of the form (79) with  $\varphi \equiv 1$  (and hence do not depend on  $\lambda$ ) the operator  $\widehat{\mathbf{A}}$  defined in (80) turns into the operator  $\mathbf{A}$  defined in (29). The proof of Proposition 2.6 now follows from Proposition 5.4 together with Kurtz's Markov mapping theorem, see Theorem A.2 in [15] (which is an extension of Corollary 3.5 in [28]). The role of the processes  $X$  and  $Y$  there is played by our processes  $(\zeta, R, G, \Lambda^M)$  and  $(\zeta, R, G)$ , respectively. In the initial distribution, we take the components of  $\Lambda_0^M$  to be i.i.d.  $\text{Unif}\{-1, 1\}$  distributed, and also for the kernel  $\alpha$  appearing in Theorem A.2 in [15], we take  $\alpha(v, r, g, \cdot)$  to be the iid  $\text{Unif}\{-1, 1\}$  distribution.

### 5.3 The symmetrized selective genealogy. Proof of Theorem 2.7

*Proof of Theorem 2.7.* The proof is divided into several steps.

**Step 1.** Recall the definition of  $D_M$  just after (29), and that of  $\Phi_F$  in (32). Recalling the definition (30), and with regard to the time change (3) between the processes  $(\zeta_s, R_s, G_s)$  and  $(\xi_t, Y_t)$ , we first claim that  $(\zeta_{s \wedge \sigma_M}, \psi(R_{s \wedge \sigma_M}, G_{s \wedge \sigma_M}))_{s \geq 0}$  solves the martingale problem  $(\widetilde{\mathbf{A}}, \mathbb{D}_M)$ , where the generator  $\widetilde{\mathbf{A}}$  is defined by

$$\widetilde{\mathbf{A}}\Phi_F(v, \chi) := \int \mathbf{A}F(v, r, g) \nu^\chi(dr, dg), \quad F \in D_M,$$

with the marked distance matrix distribution  $\nu^\chi$  defined at the beginning of Section 2.4, and where  $\mathbb{D}_M$  is the domain of  $\widetilde{\mathbf{A}}$  containing the linear span of all the functions  $\Phi_F$  for  $F \in D_M$ .

The process  $(\zeta_{s \wedge \sigma_M}, \psi(R_{s \wedge \sigma_M}, G_{s \wedge \sigma_M}))_{s \geq 0}$  arises from the process  $(\zeta_{s \wedge \sigma_M}, R_{s \wedge \sigma_M}, G_{s \wedge \sigma_M})_{s \geq 0}$  through a mapping (from one state space to the other). This mapping is given by  $(v, r, g) \mapsto (v, \psi(r, g))$ , where  $\psi(r, g)$  is the isomorphism class of the marked ultrametric measure space  $(\mathbb{T}^r, r, m^{r,g})$ .

Due to Theorem 2.5, the process  $(\zeta_{s \wedge \sigma_M}, R_{s \wedge \sigma_M}, G_{s \wedge \sigma_M})_{s \geq 0}$  is Markovian. Hence, in order to prove our first claim it suffices to show that

$$\mathbf{E} \left( \Phi_F(\zeta_{s \wedge \sigma_M}, \psi(R_{s \wedge \sigma_M}, G_{s \wedge \sigma_M})) - \int_0^{s \wedge \sigma_M} \widetilde{\mathbf{A}}\Phi_F(\zeta_u, \psi(R_u, G_u)) du \right) = 0$$

for all  $F \in D_M$ . By definition of  $\widetilde{\mathbf{A}}$  and Fubini, this follows provided

$$\mathbf{E} \left( \int \mathbf{A}F(\zeta_{u \wedge \sigma_M}, r, g) \nu^{\psi(R_{u \wedge \sigma_M}, G_{u \wedge \sigma_M})}(dr, dg) \right) = \mathbf{E}(\mathbf{A}F(\zeta_{u \wedge \sigma_M}, R_{u \wedge \sigma_M}, G_{u \wedge \sigma_M})).$$

This in turn follows as in e. g. Proposition 10.3 of [22] as  $(R_s, G_s)$  is exchangeable conditionally given  $\zeta_s$  by Proposition 5.3, and  $(R_s, G_s)$  is a.s. proper by Corollary 5.2.

**Step 2.** Next we show the well-posedness of the martingale problem  $(\widetilde{\mathbf{A}}, \mathbb{D}_M)$ .

We know from Proposition 2.6 that the process  $(\zeta_{s \wedge \sigma_M}, R_{s \wedge \sigma_M}, G_{s \wedge \sigma_M})_{s \geq 0}$  solves the martingale problem  $(\mathbf{A}, D_M)$ , and that this martingale problem is well-posed. We are thus in a convenient setting of Kurtz' Markov mapping theorem, see Theorem A.2 in [15] (which is an extension of Corollary 3.5 in [28]). Let us check the validity of the assumptions of Kurtz' theorem using the notation from there. The state space of the "coarse" process  $(\zeta_{s \wedge \sigma_M}, \psi(R_{s \wedge \sigma_M}, G_{s \wedge \sigma_M}))_{s \geq 0}$  is  $S_0 := \mathbb{R}_+ \times \mathbb{M}$ , which we can endow with

$$d_0((v, \chi), (v', \chi')) = |v - v'| + d_{\text{MGP}}(\chi, \chi'),$$

where  $d_{\text{MGP}}$  denotes the marked Gromov-Prohorov metric [9]. The state space of the “fine” process  $(\zeta^{\sigma_M}, R^{\sigma_M}, G^{\sigma_M})$  is  $S = \mathbb{R}_+ \times \mathbb{R}^{\mathbb{N}^2} \times \mathbb{I}^{\mathbb{N}}$ , which we can endow with the metric

$$d((v, r, g), (v', r', g')) = |v - v'| + d'((r, g), (r', g')),$$

where  $d'$  denotes the following refinement of the marked Gromov-Prohorov metric: For  $(r, g), (r', g') \in \mathbb{R}^{\mathbb{N}^2} \times \mathbb{I}^{\mathbb{N}}$ , we define  $d'((r, g), (r', g'))$  as the infimum of those  $\epsilon > 0$  for which there exists a complete and separable metric space  $(U, d_U)$  and isometries  $\iota : (\mathbb{T}^r, r) \rightarrow (U, d_U)$ ,  $\iota' : (\mathbb{T}^{r'}, r') \rightarrow (U, d_U)$  such that for  $\hat{\iota} : \mathbb{T}^r \times \mathbb{I} \rightarrow U \times \mathbb{I}$ ,  $(x, h) \mapsto (\iota(x), h)$  and  $\hat{\iota}' : \mathbb{T}^{r'} \times \mathbb{I} \rightarrow U \times \mathbb{I}$ ,  $(x, h) \mapsto (\iota'(x), h)$ , we have

$$d_{\text{Prohorov}}(\hat{\iota}(m^{r,g}), \hat{\iota}'(m^{r',g'})) < \epsilon \text{ and } d_U(\iota(k), \iota'(k)) + \mathbf{1}_{g(k) \neq g'(k)} < \epsilon \text{ for all } k = 1, \dots, \lfloor \epsilon^{-1} \rfloor.$$

Here  $\hat{\iota}(m^{r,g})$  denotes the image measure of  $m^{r,g}$  under the mapping  $\hat{\iota}$ , and the metric on  $U \times \mathbb{I}$  that underlies  $d_{\text{Prohorov}}$  is defined to be  $d_{U \times \mathbb{I}}((u, h), (u', h')) := d_U(u, u') + \mathbf{1}_{h \neq h'}$ .

With respect to  $d$ ,  $\mathbf{A}F$  is continuous for all  $F \in D_M$ . As the mapping  $\gamma$  from  $S$  to  $S_0$  and the probability kernel  $\alpha$  from  $S_0$  to  $S$  that figure in [15] Theorem A.2 we take

$$\gamma(v, r, g) := (v, \psi(r, g)) \in S_0, \quad \alpha((v, \chi), dv' dr dg) = \delta_v(dv') \otimes \nu^\chi(dr dg)$$

Then we have  $\gamma(\alpha(\chi, \cdot)) = \delta_\chi$  by a reconstruction argument as in Proposition 10.5 of [22]. We can rewrite the above defined operator  $\tilde{\mathbb{A}}$  as

$$\tilde{\mathbb{A}}\Phi_F(v, \chi) = \int \mathbf{A}F(v', r, g) \alpha((v, \chi), dv' dr dg), \quad F \in D_M,$$

In view of Step 1 and the well-posedness of the martingale problem  $(\mathbf{A}, D_M)$ , we can now infer the well-posedness of the martingale problem  $(\tilde{\mathbb{A}}, \mathbb{D}_M)$  as well as the Markov property of its solution from [15] Theorem A.2.

3. Steps 1 and 2 together show that  $(\zeta_{s \wedge \sigma_M}, \psi(R_{s \wedge \sigma_M}, G_{s \wedge \sigma_M}))_{s \geq 0}$  is the unique solution of the martingale problem  $(\tilde{\mathbb{A}}, \mathbb{D}_M)$  and is Markovian. The assertion of Theorem 2.7 now follows from the time-change relation  $\tilde{\mathbb{A}} = v\mathbb{A}$ .  $\square$

## 5.4 Proof of Propositions 2.9, 2.10 and 2.1

*Proof of Proposition 2.9.* The process  $Y$  takes its values in the space  $\mathbb{M}$  of marked genealogies that is equipped with the marked Gromov-weak topology, see Sec 2.2. According to [9], this topology is metrized by the so-called Gromov-Prohorov metric, and [9] Definition 3.1 ensures that the Gromov-Prohorov distance of two elements  $\chi, \chi' \in \mathbb{M}$  is bounded from above by the Prohorov distance of  $m$  and  $m'$ , where the marked ultrametric measure spaces  $(\tau, d, m)$  and  $(\tau, d, m')$  are representatives of the isomorphism classes  $\chi$  and  $\chi'$  in a common embedding. In our situation the common embedding of the representatives of  $Y_{t_1}$  and  $Y_{t_2}$  happens in the selective lookdown space  $(\hat{Z}, \rho)$ , and the two measures in the embedding are the sampling measures  $m_{s(t_1)}$  and  $m_{s(t_2)}$ , with  $m_s$  defined in (69). Since for each  $M \in \mathbb{N}$  the time change  $t \mapsto s(t)$  given by (3) is bi-continuous up to the stopping time  $\tau_M$ , it suffices to show that a.s. the map  $s \mapsto m_s$  is continuous in the weak topology on  $(\hat{Z}, \rho)$ . This latter continuity, however, is a consequence of Lemmas 3.7 and 3.8 and the fact that the fragments  $\Gamma_\gamma$  are monotypic.  $\square$

*Proof of Proposition 2.10.* We proceed in two steps. In the whole proof, let  $M > 0$  be fixed and let us consider all the processes stopped at  $\tau_M$ . For the sake of notation, we omit here the stopping times  $\tau_M$ . First, we explain how to obtain the martingale problems for  $\xi^A$  and  $\xi^B$  and second, we compute the brackets of the corresponding martingales.

**Step 1:** For any  $f : \mathbb{R}_+ \times \mathbb{I} \rightarrow \mathbb{R}$  of class  $\mathcal{C}^\infty$  with respect to its first component and bounded, we can associate a function  $F = f \circ \gamma_1$  of degree 1 on  $\mathbb{R}_+ \times \mathbb{R}^{\mathbb{N}^2} \times \mathbb{I}^{\mathbb{N}}$ . Such a function  $F$  belongs to  $D_M$  with  $F(v, r, g) = f(v, g(1))$  for all  $(v, r, g) \in \mathbb{R}_+ \times \mathbb{R}^{\mathbb{N}^2} \times \mathbb{I}^{\mathbb{N}}$ , and our purpose is to rewrite the martingale problem (34) for such test function  $F$ .

For the first term in the left hand side of (34), we have that:

$$\Phi_F(\xi_t, Y_t) = \int f(\xi_t, g(1)) \nu^{Y_t}(dr, dg) = f(\xi_t, A) \mu_t^A + f(\xi_t, B) (1 - \mu_t^A), \quad (83)$$

since under the marked distance matrix distribution  $\nu^{Y_t}$ , the type configuration corresponds to an i.i.d. sequence drawn from  $\mu_t^A \delta_A(dh) + (1 - \mu_t^A) \delta_B(dh)$ , which results from Theorem 2.2 and the explanations in Section 2.3.3.

Now, let us compute the second term of the left hand side of (34). For our choice of function  $F$ , we have from (29) that:

$$\begin{aligned} \mathbb{A}\Phi_F(v, Y_t) &= \int \nu^{Y_t}(dr, dg) \left\{ \frac{v}{2} f_{vv}(v, g(1)) + (bv\mu^g\{A\} - 2cv^2\mu^g\{A\}\mu^g\{B\}) f_v(v, g(1)) \right. \\ &\quad \left. + cv \int m^{r,g}(d\theta, dh) \left[ \mathbf{1}_{\{g(1)=B\}} \mu^g\{A\} (f(v, h) - f(v, B)) + \mathbf{1}_{\{g(1)=A\}} \mu^g\{B\} (f(v, h) - f(v, A)) \right] \right. \\ &\quad \left. + b \int m^{r,g}(d\theta, dh) \mathbf{1}_{\{h=A\}} (f(v, A) - f(v, g(1))) \right\} \\ &= \int \nu^{Y_t}(dr, dg) \left\{ \frac{v}{2} f_{vv}(v, g(1)) + (bv\mu^g\{A\} - 2cv^2\mu^g\{A\}\mu^g\{B\}) f_v(v, g(1)) \right. \\ &\quad \left. + cv \left( \mathbf{1}_{\{g(1)=B\}} (\mu^g\{A\})^2 - \mathbf{1}_{\{g(1)=A\}} (\mu^g\{B\})^2 \right) (f(v, A) - f(v, B)) \right. \\ &\quad \left. + b\mu^g\{A\} (f(v, A) - f(v, g(1))) \right\}, \end{aligned}$$

by recalling that the projection of  $m^{r,g}(d\theta, dh)$  on its second component gives the type frequencies  $\mu^g\{A\}$  and  $\mu^g\{B\}$ . Under  $\nu^{Y_t}(dr, dg)$ ,  $\mu^g\{A\}$  (resp.  $\mu^g\{B\} = 1 - \mu_t^A$ ) is constant and equal to  $\mu_t^A$  (resp.  $\mu_t^B$ ) and  $g(1)$  is a random variable that takes the values  $A$  and  $B$  with probabilities  $\mu_t^A$  and  $1 - \mu_t^A$ . We then deduce that:

$$\begin{aligned} \mathbb{A}\Phi_F(v, Y_t) &= \mu_t^A \frac{v}{2} f_{vv}(v, A) + (1 - \mu_t^A) \frac{v}{2} f_{vv}(v, B) \\ &\quad + (bv\mu_t^A - 2cv^2\mu_t^A(1 - \mu_t^A)) (\mu_t^A f_v(v, A) + (1 - \mu_t^A) f_v(v, B)) \\ &\quad + cv\mu_t^A(1 - \mu_t^A) (2\mu_t^A - 1) (f(v, A) - f(v, B)) \\ &\quad + b\mu_t^A(1 - \mu_t^A) (f(v, A) - f(v, B)). \end{aligned} \quad (84)$$

Choosing  $f$  in (83) and (84) such that  $f(v, h) = v\mathbf{1}_{\{h=A\}}$  and replacing  $v$  with  $\xi_t$ , we find that:

$$M_t^A := \xi_t^A - \xi_0^A - \int_0^t (b\xi_u^A - c\xi_u^A \xi_u^B) du \quad (85)$$

is a local martingale (when stopped at  $\tau_M$ ,  $M_{\cdot \wedge \tau_M}^A$  is a square integrable martingale) started at 0. This martingale is also continuous by Proposition 2.9. We can proceed similarly to find that  $M_t^B :=$

$\xi_t^B - \xi_0^B + \int_0^t c \xi_u^A \xi_u^B du$  is also a continuous local martingale started at 0.

**Step 2:** Let us now compute the brackets  $\langle M^A \rangle$ ,  $\langle M^B \rangle$  and  $\langle M^A, M^B \rangle$ . Proceed similarly as in Step 1 for functions  $f : \mathbb{R}_+ \times \mathbb{I}^2 \rightarrow \mathbb{R}$  of class  $\mathcal{C}^\infty$  with respect to its first component, and to which we associate  $F = f \circ \gamma_2$  of degree 2 on  $\mathbb{R}_+ \times \mathbb{R}^{\mathbb{N}^2} \times \mathbb{I}^{\mathbb{N}}$ . Such a function  $F$  belongs to  $D_\infty$  with  $F(v, r, g) = f(v, g(1), g(2))$  for all  $(v, r, g) \in \mathbb{R}_+ \times \mathbb{R}^{\mathbb{N}^2} \times \mathbb{I}^{\mathbb{N}}$ .

For the choice of  $f(v, h_1, h_2) = v^2 \mathbf{1}_{\{h_1=A\}} \mathbf{1}_{\{h_2=A\}}$ , we obtain that:

$$(\xi_t^A)^2 - \int_0^t \left( \xi_u^A + 2b(\xi_u^A)^2 - 2c(\xi_u^A)^2 \xi_u^B \right) du \quad (86)$$

is a continuous local martingale. Using Itô's formula on (85), we also have that:

$$(\xi_t^A)^2 - \int_0^t \left( 2b(\xi_u^A)^2 - 2c(\xi_u^A)^2 \xi_u^B \right) du - \langle M^A \rangle_t \quad (87)$$

is a continuous local martingale. From the comparison of these two expressions, we deduce that:

$$\langle M^A \rangle_t = \int_0^t \xi_u^A du. \quad (88)$$

In a similar way, the choices of  $f(v, h_1, h_2) = v^2 \mathbf{1}_{\{h_1=B\}} \mathbf{1}_{\{h_2=B\}}$  and  $f(v, h_1, h_2) = v^2 \mathbf{1}_{\{h_1=A\}} \mathbf{1}_{\{h_2=B\}}$  allow us to compute  $\langle M^B \rangle$  and  $\langle M^A, M^B \rangle$ . Using Levy's representation theorem [38, Th. IV.3.6, p.141], we deduce that there exists on an enlarged probability space two independent Brownian motions  $W^A$  and  $W^B$  such that  $dM_t^A = \sqrt{\xi_t^A} dW_t^A$  and  $dM_t^B = \sqrt{\xi_t^B} dW_t^B$ .  $\square$

*Proof of Proposition 2.1.* Let us consider the following stopping times, for any  $\varepsilon > 0$ :

$$\tau_\varepsilon^A = \inf\{t \geq 0, \xi_t^A \leq \varepsilon\}, \quad \text{and} \quad \tau_\varepsilon^B = \inf\{t \geq 0, \xi_t^B \leq \varepsilon\}, \quad (89)$$

with the usual convention that  $\inf \emptyset = +\infty$ . Before  $\tau_\varepsilon^A \wedge \tau_\varepsilon^B$ , the diffusion coefficients are Lipschitz continuous (with a Lipschitz constant of order  $1/\sqrt{\varepsilon}$ ). Classical results (e.g. [23, Chap. IV]) ensure strong uniqueness of the stopped processes  $(\xi_{\cdot \wedge \tau_\varepsilon^A \wedge \tau_\varepsilon^B}^A, \xi_{\cdot \wedge \tau_\varepsilon^A \wedge \tau_\varepsilon^B}^B)$  for all  $\varepsilon > 0$ . Let  $\tau_0 = \inf\{t \geq 0, \xi_t^A = 0 \text{ or } \xi_t^B = 0\}$ . By the continuity of the processes,  $\lim_{\varepsilon \downarrow 0} \tau_\varepsilon^A \wedge \tau_\varepsilon^B = \tau_0$ . Once one of the processes  $\xi^A$  or  $\xi^B$  has touched zero, it remains trapped there and the other process coincides with a standard (possibly drifted) Feller diffusion  $\bar{\xi}^A$  or  $\bar{\xi}^B$ :

$$\begin{aligned} d\bar{\xi}_t^A &= b\bar{\xi}_t^A dt + \sqrt{\bar{\xi}_t^A} dW_t^A \quad \text{and} \quad \bar{\xi}_t^B = 0 \\ \text{or } \bar{\xi}_t^A &= 0 \quad \text{and} \quad d\bar{\xi}_t^B = \sqrt{\bar{\xi}_t^B} dW_t^B. \end{aligned} \quad (90)$$

The latter diffusions are well studied (see e.g. [23, Chap. IV.8]) and we have strong existence and uniqueness for (90).  $\square$

Note that the two (independent) Feller diffusions appearing in that proof (see eq. (90)) also provide dominating processes for  $(\xi^A, \xi^B)$  (see e.g. [23, Th. VI.1.1]). Let us define the process  $\bar{\xi}^A$  as the Feller diffusion with nonnegative growth rate  $b \geq 0$ ; it remains nonnegative for all  $t \geq 0$ , and 0 is a trap. Denoting  $\bar{\tau}_0^A = \inf\{t \geq 0, \bar{\xi}_t^A = 0\}$ , it is known that  $\mathbf{P}(\bar{\tau}_0^A < +\infty) \in (0, 1)$  and that on the set  $\{\bar{\tau}_0^A = +\infty\}$ ,  $\lim_{t \rightarrow +\infty} \bar{\xi}_t^A = +\infty$  a.s. (see [4, Corollary 2, p.190]). Also, let us define the process  $\bar{\xi}^B$  as a critical Feller diffusion without finite variational part and that gets extinct almost surely in finite

time:  $\mathbf{P}(\bar{\tau}_0^B < +\infty) = 1$  with  $\bar{\tau}_0^B = \inf\{t \geq 0, \bar{\xi}_t^B = 0\}$ . The process  $(\bar{\xi}^A, \bar{\xi}^B)$  dominates stochastically  $(\xi^A, \xi^B)$ . As a consequence,  $\xi^B$  gets extinct in finite time almost surely. Overall, either  $\xi^A$  touches zero before  $\xi^B$  and the whole process then goes to extinction, or  $\xi_{\bar{\tau}_0^B}^A > 0$  and there is a positive probability that  $\tau_0^A = +\infty$  and when this happens,  $\lim_{t \rightarrow +\infty} \xi_t^A = +\infty$  a.s.

## 6 Outlook: An extension to multiple types and mutations

The previous sections were restricted to a prototype example with two types and without mutation. Indeed, we believe that this example, which allowed for a trade-off between conciseness and elaboration, is best suited for displaying our novel *pathwise* approach to the joint evolution of population size, type configuration and genealogy.

In this concluding section we give a brief outlook to a more general situation, without going into further details. Let now the type space  $\mathbb{I}$  be a compact group. Again we write  $\mu_t$  for the relative type frequencies and  $\xi_t$  for the total mass of the population at time  $t$ , and we put  $\Xi_t := \xi_t \mu_t$ . The state space of  $(\mu_t)$  is  $M^1(\mathbb{I})$ , the set of probability measures on  $\mathbb{I}$ , and that of  $(\Xi_t)$  is  $M(\mathbb{I})$ , the set of finite measures on  $\mathbb{I}$ , equipped with the weak topology.

Let  $\underline{b} = \underline{b}(h)$  and  $\underline{c} = \underline{c}(h, h')$  be bounded, measurable mappings from  $\mathbb{I}$  to  $\mathbb{R}_+$  and  $\mathbb{I} \times \mathbb{I}$  to  $\mathbb{R}_+$ , respectively. For  $\rho \in M^1(\mathbb{I})$  and  $h \in \mathbb{I}$  we put  $c(h, \rho) := \int_{\mathbb{I}} \underline{c}(h, h') \rho(dh')$ . Finally, let  $h \mapsto \ell(h, \cdot)$  be a measurable map from  $\mathbb{I}$  to  $M^1(\mathbb{I})$ .

We say that  $\Xi$  is an interactive Dawson-Watanabe process with fecundity function  $b$ , competition kernel  $c$  and mutation kernel  $\ell$  if for all  $f \in \mathcal{C}(\mathbb{I})$ , the continuous functions on  $\mathbb{I}$ ,

$$\int_{\mathbb{I}} f(h) \Xi_t(dh) - \int_0^t \int_{\mathbb{I}} \left( f(h)(b(h) - c(h, \mu_u) \xi_u) + \int_{\mathbb{I}} (f(h') - f(h)) \ell(h, dh') \right) \Xi_u(dh) du$$

is a continuous martingale with quadratic variation

$$\int_0^t \int_{\mathbb{I}} f^2(h) \Xi_u(dh) du,$$

cf. [13], Example 4.6 for the non-interactive case. Putting  $f \equiv 1$  we see that the total mass process  $\xi_t := \Xi_t(\mathbb{I})$ ,  $t \geq 0$ , is required to be a weak solution of the SDE

$$d\xi_t = \left( \xi_t \int_{\mathbb{I}} b(h) \mu_t(dh) - \xi_t^2 \int_{\mathbb{I}} c(h, \mu_t) \mu_t(dh) \right) dt + \sqrt{\xi_t} dW_t. \quad (91)$$

Our prototype example fits into this framework with  $\mathbb{I} = \{A, B\}$ ,  $\ell \equiv 0$  and  $\underline{b}(A) = b$ ,  $\underline{b}(B) = 0$ ,  $\underline{c}(A, A) = \underline{c}(B, B) = 0$ ,  $\underline{c}(A, B) = \underline{c}(B, A) = c$ .

In order to arrive at an analogue of Theorem 2.2 in this more general framework one has to modify the update rules (9) and (10). In addition to the symbols  $\beta$  and  $\delta$  that indicate “birth” or “death” as the 4<sup>th</sup> component of the Poisson point measures  $\mathcal{K}_i$ , we now have a third symbol  $\lambda$  that figures for “mutation”: the  $\mathcal{K}_i$  are now a family of independent Poisson processes on  $\mathbb{R}_+ \times \mathbb{R}_+ \times [0, 1] \times \{\beta, \delta, \lambda\}$ , with  $\mathcal{K}_i(\cdot \times \{\beta\})$ ,  $\mathcal{K}_i(\cdot \times \{\delta\})$  and  $\mathcal{K}_i(\cdot \times \{\lambda\})$  having Lebesgue intensity measure.

With the abbreviation  $\mu^{g,b}(dh') := \frac{b(h')\mu^g(dh')}{\int b d\mu^g}$ , the update rules (9) and (10) are modified to

$$\begin{aligned}
q(h, g, v, z, w, \beta) &:= \begin{cases} \kappa(\mu^{g,b}, w) & \text{if } z \leq \int b d\mu^g v, \\ h & \text{otherwise,} \end{cases} \\
q(h, g, v, z, w, \delta) &:= \begin{cases} \kappa(\mu^g, w) & \text{if } z \leq c(h, \mu^g) v^2, \\ h & \text{otherwise.} \end{cases} \\
q(h, g, v, z, w, \lambda) &:= \begin{cases} \kappa(\ell(h, \cdot), w) & \text{if } z \leq v, \\ h & \text{otherwise.} \end{cases}
\end{aligned}$$

Likewise, the total mass process  $(\zeta_s)$  in the lookdown timescale, which is another ingredient of Theorem 2.2, will be a time change of  $(\xi_t)$  under (3), turning (91) into

$$d\zeta_s = \left( \zeta_s \int_{\mathbb{I}} b(h) \mu^{G_s}(dh) - \zeta_s^2 \int_{\mathbb{I}} c(h, \mu^{G_s}) \mu^{G_s}(dh) \right) \zeta_s ds + \zeta_s d\mathcal{W}_s.$$

**Acknowledgements:** We thank Götz Kersting for helping us with the proof of Lemma 4.2. A.B. was supported as a postdoc by CONACyT in an earlier phase of this project. She gratefully acknowledges the kind hospitality of Goethe-University Frankfurt. S.G. has been supported in part at the Technion by a Zeff Fellowship, a postdoctoral fellowship of the Minerva Foundation, and through Israel Science Foundation (ISF) grant No. 1382/17, Binational Science Foundation (BSF) award 2018330, and Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) research grant contract number 2337/1-1. S.K. has been funded by the DFG - Project number 393092071. V.C.T. has been funded by Labex CEMPI (ANR-11-LABX-0007) and Labex Bézout (ANR-10-LABX-58). V.C.T. also acknowledges support from the Chaire "Modélisation Mathématique et Biodiversité" of Veolia Environnement-Ecole Polytechnique-Museum National d'Histoire Naturelle-Fondation X. A.W. received partial support through DFG project WA 967/4-2 in the SPP 1590. Also, A.B., S.K., V.C.T. and A.W. would like to thank the Institute for Mathematical Sciences for supporting their visit to the IMS, National University of Singapore in 2017, where progress on this project was made.

## References

- [1] D. Aldous. The continuum random tree. II. An overview. In *Stochastic analysis (Durham, 1990)*, volume 167 of *London Math. Soc. Lecture Note Ser.*, pages 23–70. Cambridge Univ. Press, Cambridge, 1991.
- [2] J. Berestycki and N. Berestycki. Kingman's coalescent and Brownian motion. *ALEA Lat. Am. J. Probab. Math. Stat.*, 6:239–259, 2009.
- [3] J. Berestycki, M. C. Fittipaldi, and J. Fontbona. Ray–Knight representation of flows of branching processes with competition by pruning of Lévy trees. *Probab. Theory Related Fields*, 172(3-4):725–788, 2018.
- [4] J. Bertoin. *Lévy processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996.
- [5] S. Billiard, R. Ferrière, S. Méléard, and V.C. Tran. Stochastic dynamics of adaptive trait and neutral marker driven by eco-evolutionary feedbacks. *Journal of Mathematical Biology*, 71(5):1211–1242, 2015.
- [6] A. Blancas, S. Gufler, S. Kliem, V.C. Tran, and A. Wakolbinger. Evolving genealogies for finite branching populations under selection and competition, 2021. Work in progress.
- [7] V. Calvez, B. Henry, S. Méléard, and V.C. Tran. Dynamics of lineages in adaptation to a gradual environmental change. arXiv:2104.10427, 2021.
- [8] D. Dawson and E. Perkins. *Historical Processes*, volume 93. American Mathematical Society, Memoirs of the American Mathematical Society edition, 1991.

- [9] A. Depperschmidt, A. Greven, and P. Pfaffelhuber. Marked metric measure spaces. *Electron. Commun. Probab.*, 16:174–188, 2011.
- [10] A. Depperschmidt, A. Greven, and P. Pfaffelhuber. Tree-valued Fleming-Viot dynamics with mutation and selection. *Ann. Appl. Probab.*, 22(6):2560–2615, 2012.
- [11] P. Donnelly and T. Kurtz. Particle representations for measure-valued population models. *Annals of Probability*, 27(1):166–205, 1999.
- [12] P. Donnelly and T. G. Kurtz. Genealogical processes for Fleming-Viot models with selection and recombination. *Ann. Appl. Probab.*, 9(4):1091–1148, 1999.
- [13] P. Donnelly and T. G. Kurtz. Particle representations for measure-valued population models. *Ann. Probab.*, 27(1):166–205, 1999.
- [14] E. Dynkin. Branching particle systems and superprocesses. *Annals of Probability*, 19:1157–1194, 1991.
- [15] A. Etheridge and T. G. Kurtz. Genealogical constructions of population models. *Ann. Probab.*, 47(4):1827–1910, 2019.
- [16] A. M. Etheridge. *An introduction to superprocesses*, volume 20 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2000.
- [17] S. Ethier and T. Kurtz. *Markov Processes, Characterization and Convergence*. John Wiley & Sons, New York, 1986.
- [18] S. N. Evans. Kingman’s coalescent as a random metric space. In *Stochastic models (Ottawa, ON, 1998)*, volume 26 of *CMS Conf. Proc.*, pages 105–114. Amer. Math. Soc., Providence, RI, 2000.
- [19] A. Greven, P. Pfaffelhuber, and A. Winter. Convergence in distribution of random metric measure spaces (lambda-coalescent measure trees). *Probability Theory and Related Fields*, 145(1):285–322, 2009.
- [20] S. Gufler. Tree-valued Fleming-Viot processes: a generalization, pathwise constructions, and invariance principles, 2017. PhD Thesis, Goethe-Universität Frankfurt.
- [21] S. Gufler. Pathwise construction of tree-valued Fleming-Viot processes. *Electron. J. Probab.*, 23:Paper No. 42, 58, 2018.
- [22] S. Gufler. A representation for exchangeable coalescent trees and generalized tree-valued Fleming-Viot processes. *Electron. J. Probab.*, 23:Paper No. 41, 42, 2018.
- [23] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*, volume 24. North-Holland Publishing Company, 1989. Second Edition.
- [24] G. Kersting and A. Wakolbinger. Probabilistic aspects of  $\Lambda$ -coalescents in equilibrium and in evolution. To appear in *Probabilistic Structures in Evolution*, E. Baake, A. Wakolbinger, eds., EMS Series of Congress Reports Vol. 17, EMS Press 2021, arXiv:2002.05250 [math.PR].
- [25] S. Kliem. A compact containment result for nonlinear historical superprocess approximations for population models with trait-dependence. *Electronic Journal of Probability*, 19(97):1–13, 2014.
- [26] S. Kliem and W. Löhr. Existence of mark functions in marked metric measure spaces. *Electron. J. Probab.*, 20(73):1–24, 2015.
- [27] S. Kliem and A. Winter. Evolving phylogenies of trait-dependent branching with mutation and competition. part i: Existence. *Stochastic Processes and their Applications*, 129(12):4837–4877, 2019.
- [28] T. G. Kurtz. Martingale problems for conditional distributions of markov processes. *Electron. J. Probab.*, 3(4):Paper no. 9, 1998.
- [29] T. G. Kurtz and E. R. Rodrigues. Poisson representations of branching Markov and measure-valued branching processes. *Ann. Probab.*, 39(3):939–984, 2011.
- [30] V. Le, E. Pardoux, and A. Wakolbinger. “Trees under attack”: a Ray-Knight representation of Feller’s branching diffusion with logistic growth. *Probab. Theory Related Fields*, 155(3-4):583–619, 2013.
- [31] J.-F. Le Gall. Random trees and applications. *Probab. Surv.*, 2:245–311, 2005.
- [32] C. Lepers, S. Billiard, M. Porte, S. Méléard, and V.C. Tran. Inference with selection, varying population size and evolving population structure: Application of abc to a forward-backward coalescent process with interactions. arXiv:1910.10201 [q-bio.PE], submitted, 2019.
- [33] S. Méléard and V.C. Tran. Nonlinear historical superprocess approximations for population models with past dependence. *Electronic Journal of Probability*, 17(47), 2012.
- [34] E. Pardoux and A. Wakolbinger. From exploration paths to mass excursions—variations on a theme of Ray and Knight. In *Surveys in stochastic processes*, EMS Ser. Congr. Rep., pages 87–106. Eur. Math. Soc., Zürich, 2011.

- [35] E. A. Perkins. Conditional Dawson-Watanabe processes and Fleming-Viot processes. In *Seminar on Stochastic Processes, 1991 (Los Angeles, CA, 1991)*, volume 29 of *Progr. Probab.*, pages 143–156. Birkhäuser Boston, Boston, MA, 1992.
- [36] P. Pfaffelhuber and A. Wakolbinger. The process of most recent common ancestors in an evolving coalescent. *Stochastic Process. Appl.*, 116(12), 2006.
- [37] P. E. Protter. On the existence, uniqueness, convergence and explosions of solutions of systems of stochastic integral equations. *Ann. Probability*, 5(2):243–261, 1977.
- [38] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer-Verlag, Berlin, 1998.