Many-to-few for non-local branching Markov process

Simon C Harris, Emma Horton, Andreas E Kyprianou, Ellen Powell

Many-to-few for non-local branching Markov process

Simon C. Harris[*], Emma Horton[†], Andreas E. Kyprianou[‡] and Ellen Powell[§]

November 16, 2022

Abstract

We provide a many-to-few formula in the general setting of non-local branching Markov processes. This formula allows one to compute expectations of k-fold sums over functions of the population at k different times. The result generalises [14] to the non-local setting, as introduced in [11] and [8]. As an application, we consider the case when the branching process is critical, and conditioned to survive for a large time. In this setting, we prove a general formula for the limiting law of the death time of the most recent common ancestor of two particles selected uniformly from the population at two different times, as $t \to \infty$. Moreover, we describe the limiting law of the population sizes at two different times, in the same asymptotic regime.

Key words: non-local branching processes, many-to-few, spines.

MSC 2020: 60J80, 60J25

1 Introduction

1.1 Main results

Our main result, the so called *many-to-few formula*, is a way to rewrite the expectation of a general k-fold sum, depending on the entire configuration of a branching Markov process at k different times, as an expectation with respect to the behaviour of k distinguished lines of descent under a tilted measure. We generalise the original and well cited main result of [14], by allowing for non-local branching, and not requiring the k individuals to be sampled at the same time.

The many-to-few formula generalises the role of the classical spine decomposition for spatial branching processes, which converts expectation identities for additive functionals of spatial branching processes to Feynman-Kac formulae for a single Markov particle trajectory. The latter has proved to be an important tool in analysing the growth and spread of a rich variety of branching Markov processes and related models; see for example the monographs [1, 6, 18, 15, 2]

[*]University of Auckland. E-mail: simon.harris@auckland.ac.nz
[†]Inria Research Centre Bordeaux. E-mail: emma.horton@inria.fr
[‡]University of Bath. E-mail: a.kyprianou@bath.ac.uk
[§]University of Durham. E-mail: ellen.g.powell@durham.ac.uk

among a wide base of research literature that is to too extensive to exhaustively list here. The many-to-few formula has already played an important and similar role to the classical spine decomposition as a tool to interrogate various questions pertaining to particle correlation that arise in e.g. genealogical coalescent structure, [12], martingale convergence, [10], maximal displacement of extreme particles, [17], the structure of level sets for branching Brownian motion, [4] and the analysis of certain models from the theory of stochastic genetics, [7]. We further remark that the many-to-few formula is related to recent work pertaining to asymptotic moment convergence in [8].

We refrain from attempting to give a precise statement of the many-to-few formula here, deferring instead to Lemma 6 below, as we will need to introduce several objects in order for the formula to be understood in a meaningful way. We note that, simultaneously to the results we present here a general branching Markov process setting, similar ideas have been developed in [7].

Our main motivating application is to understand the limiting genealogy of a so-called *critical* branching Markov process, when conditioned to survive for an arbitrarily long time. Our second main result, Proposition 8, is a general statement in this direction. More precisely, for a critical non-local branching Markov process conditioned to survive until a large time t, we provide a precise asymptotic for the death time of the most recent common ancestor of two individuals sampled uniformly from the population at two different times. In Proposition 9, we also describe the limiting law of the population sizes at two different times, in the same asymptotic regime.

1.2 Set-up and assumptions

Let E be a Lusin space. Throughout, will write $B(E)$ for the Banach space of bounded measurable functions on E with norm $\|\cdot\|$, $B^+(E)$ for non-negative bounded measurable functions on E and $B_1^+(E)$ for the subset of functions in $B^+(E)$ which are uniformly bounded by unity.

We consider a spatial branching process in which, given their point of creation, particles evolve independently according to a Markov process, $(\xi, \mathbf{P})$, which can be characterised via the semigroup $\mathtt{P}_t[f](x) = \mathbf{E}_x[f(\xi_t)]$, for $x \in E$, $t \geq 0$ and $f \in B_1^+(E)$. In an event which we refer to as 'branching', particles positioned at x die at rate $\beta(x)$ where $\beta \in B^+(E)$ and instantaneously, new particles are created in E according to a point process. The configurations of these offspring are described by the random counting measure

$$\mathcal{Z}(A) = \sum_{i=1}^{N} \delta_{x_i}(A), \tag{1}$$

for Borel A in E. The law of the aforementioned point process may depend on x, the point of death of the parent, and we denote it by $\mathcal{P}_x$, $x \in E$, with associated expectation operator given by $\mathcal{E}_x$, $x \in E$. This information is captured in the so-called branching mechanism

$$\mathtt{G}[f](x) := \beta(x)\mathcal{E}_x\left[\prod_{i=1}^{N} f(x_i) - f(x)\right], \qquad x \in E,\ f \in B_1^+(E). \tag{2}$$

Without loss of generality we can assume that $\mathcal{P}_x(N = 1) = 0$ for all $x \in E$ by viewing a branching event with one offspring as an extra jump in the motion. On the other hand, we do allow for the possibility that $\mathcal{P}_x(N = 0) > 0$ for some or all $x \in E$.

Moreover, we do not need $\mathtt{P}$ to have the Feller property, and it is not necessary that $\mathtt{P}$ is conservative. That said, if so desired, we can append a cemetery state $\{\dagger\}$ to E, which is to be treated as an absorbing state, and regard $\mathtt{P}$ as conservative on the extended space $E \cup \{\dagger\}$, which can also be treated as a Lusin space. Equally, we can extend $\mathtt{G}$ to $E \cup \{\dagger\}$ by defining it to be zero on $\{\dagger\}$, that is no branching activity on the cemetery state.

Henceforth we refer to this spatial branching process as a $(\mathtt{P}, \mathtt{G})$-branching Markov process. It is well known that if we arbitrarily choose an order for the particles at each branching event then we may denote the configuration of particles at time t by $\{x_1(t), \dots, x_{N_t}(t)\}$ (where N_t denotes the number of particles alive at time t) and we can also associate an Ulam-Harris label

$$v_i(t) \in \Omega := \{\emptyset\} \cup \bigcup_{n \geq 1} \mathbb{N}^n$$

to the i-th particle at time t. For $v, w \in \Omega$, we write $v \preceq w$ to mean that v is an ancestor of w, which means there exists $u \in \Omega$ such that $vu = w$. Moreover, we write $v \prec w$ to mean that $v \preceq w$ in the strict sense, that is, the possibility that $v = w$ is excluded. We say that $v, w \in \Omega$ are siblings if there exists $u \in \Omega$ and $i \neq j$ such that $v = ui$ and $w = uj$.

The branching Markov process can be described via the co-ordinate process $X = (X_t, t \geq 0)$ in the space of counting measures on $E \times \Omega$ with non-negative integer total mass, denoted by $M(E \times \Omega)$, where

$$X_t(\cdot) = \sum_{i=1}^{N_t} \delta_{(x_i(t), v_i(t))}(\cdot), \qquad t \geq 0.$$

In particular, X is Markovian in $M(E \times \Omega)$. Its probabilities will be denoted $\mathbb{P} := (\mathbb{P}_{\delta_x}, x \in E)$ where for $x \in E$, $\mathbb{P}_{\delta_x}$ denotes the law of the process starting from $\delta_{(x,\emptyset)} \in M(E \times \Omega)$.

Under the additional assumption that $\sup_{x \in E} \mathcal{E}_x(N) < \infty$, where we recall that N is the (random) number of offspring produced at a branching event, we define the linear semigroup

$$\mathtt{T}_t[f](x) := \mathbb{E}_{\delta_x}[X_t[f]] := \mathbb{E}_{\delta_x}\left[\sum_{i=1}^{N_t} f(x_i(t))\right], \quad f \in B^+(E).$$

Now let us introduce an assumption, that we will use throughout the article unless stated otherwise. We will use the notation

$$\langle f, \mu \rangle := \int_E f(x)\mu(\mathrm{d}x), \quad f \in B(E), \mu \in M(E),$$

where $M(E)$ is the set of finite measures on E.

Assumption 1. *For the Markov process $(\xi, \mathbf{P})$, we assume that*

(a) it admits a càdlàg modification;

(b) there exists an eigenvalue $\lambda \in \mathbb{R}$ *and a corresponding right eigenfunction* $\varphi \in B^+(E)$ *and finite left eigenmeasure* $\tilde{\varphi}$ *such that, for* $f \in B^+(E)$*,*

$$\langle \mathtt{T}_t[\varphi], \mu \rangle = \mathrm{e}^{\lambda t} \langle \varphi, \mu \rangle \text{ and } \langle \mathtt{T}_t[f], \tilde{\varphi} \rangle = \mathrm{e}^{\lambda t} \langle f, \tilde{\varphi} \rangle,$$

for all $\mu \in M(E)$*.*

The first part of the above assumption is a regularity assumption on the Markov process $(\xi, \mathbf{P})$, which ensures that we can use the theory of martingale changes of measure. The second is a Perron Frobenius assumption that ensures the existence of the leading eigenvalue and corresponding eigenfunctions.

1.3 Outline

The many-to-few formula, Lemma 6, will ultimately allow us to express expectations of general k-fold sums depending on the entire branching process under $\mathbb{P}$, in terms of an expectation with respect to k so-called *spine particles* under a different measure $\mathbb{Q}^k$. In Section 2, we define this measure $\mathbb{Q}^k$, introduce the notion of spines, and give an explicit expression for the Radon-Nikodym derivative of $\mathbb{Q}^k$, with respect to the original law of our branching process plus some uniformly chosen marked lines of descent (we call this measure $\mathbb{P}^k$). We then state and prove our main result, the many-to-few formula, in Section 3. We also explain some special cases in which the formula simplifies nicely. Section 4 is devoted to our main application, which is to derive some two-point asymptotics for the geneologies of critical branching processes, when they are conditioned to survive for a large time t. More precisely, we provide an asymptotic (as $t \to \infty$) for the death time of the most recent common ancestor of two particles sampled uniformly from the population at two different times.

2 Spines, martingales and changes of measure

In this section, we introduce two measures under which our class of branching Markov process additionally identifies k distinguished genealogical lines of descent, or *spines*. The first of these two measures is a simple adaptation of the original law. The second measure is identified via a change of measure with respect to a certain multiplicative martingale.

2.1 Definition of the measures $\mathbb{P}^k$ and $\mathbb{Q}^k$

2.1.1 Definition of $\mathbb{P}^k$

We first introduce a measure $\mathbb{P}^k$ on the set of processes taking values in $M(E \times \Omega \times \mathcal{P}(\{1, \ldots, k\}))$, the space of counting measures on $E \times \Omega \times \mathcal{P}(\{1, \ldots, k\})$, where $\mathcal{P}(\{1, \ldots, k\})$ is the set of subsets of $\{1, \ldots, k\}$. For convenience, let $\tilde{X}$ denote the branching process on the space $M(E \times \Omega \times \mathcal{P}(\{1, \ldots, k\}))$, that is

$$\tilde{X}_t = \sum_{i=1}^{N_t} \delta_{(x_i(t), v_i(t), \mathbf{b}_i(t))},$$

where $\mathbf{b}_i(t) \in \mathcal{P}(\{1, \dots, k\})$ denotes the set of marks carried by the i-th particle alive at time t. Whenever $\mathbf{b}_i(t) \neq \emptyset$, we refer to the individual i as a *spine*. In that case, we say that the spine *carries* $|\mathbf{b}_i(t)|$ *marks*. Given $\tilde{X}$, define X to be its projection onto $M(E \times \Omega)$. With this notation in hand, we let $(\mathcal{F}_t, t \geq 0)$ denote the natural filtration generated by X and $(\mathcal{F}_t^k, t \geq 0)$ denote the natural filtration generated by $\tilde{X}$.

Then, we have the following description of the measure $\mathbb{P}^k$.

Definition 1. *The construction of $\tilde{X}$ under $\mathbb{P}^k$ goes as follows.*

1. *We start with a single particle at $x \in E$ which carries $k \geq 1$ marks.*
2. *All particles move according to the semigroup $\mathtt{P}$, independently of each other given their birth times and configurations.*
3. *Let ξ_t^i denote the position of the particle that carries the mark $1 \leq i \leq k$ at time $t \geq 0$.*
4. *A particle at $y \in E$ carrying j marks $b_1 < b_2 < \cdots < b_j$, dies at rate $\beta(y)$ and simultaneously produces a random number of new particles according to $(\mathcal{Z}, \mathcal{P}_y)$. The j marks each choose a particle to follow independently and uniformly from the $N = \langle 1, \mathcal{Z} \rangle$ available particles.*
5. *In the event that a particle carrying $j > 0$ marks dies and is replaced by 0 offspring, it is sent to the cemetery state, along with its marks.*

Note that the above definition of $\mathbb{P}^k$ is such that X has the same law under both $\mathbb{P}^k$ and $\mathbb{P}$.

2.1.2 The measure $\mathbb{Q}^k$

We will now introduce a second measure, $\mathbb{Q}^k$, under which particles not carrying any marks evolve in the same manner as particles under $\mathbb{P}$, while spines (that is, particles carrying marks) evolve differently. In the next section we will show that $\mathbb{Q}^k$ can be defined via a change of measure of $\mathbb{P}^k$. Before describing the process, we need to introduce some more notation.

Suppose that we are given a functional $(\zeta(\cdot, t), t \geq 0)$ such that $\zeta(\xi, t)$ is a non-negative unit-mean martingale with respect to the natural filtration of the Markov process $(\xi_t, t \geq 0)$ with semigroup $(\mathtt{P}_t, t \geq 0)$. We will assume that ζ takes the value 0 whenever $\xi_t = \dagger$. Now for $k, n \in \mathbb{N}$, define

$$\langle \varphi, \mathcal{Z} \rangle_{k,n} = \mathbf{1}_{(n \leq N)} \sum_{[k_1, \dots, k_N]_k^n} \binom{k}{k_1, \dots, k_N} \prod_{i : k_i > 0} \varphi(x_i), \tag{3}$$

where $(x_i, i = 1, \dots, N)$ are as in (1) and $[k_1, \dots, k_N]_k^n$ is the set of non-negative integer N-tuples $(k_1, \dots, k_N)$ such that $k_1 + \cdots + k_N = k$ and exactly n of the k_is are positive. If $\mathcal{Z}$ corresponds to the offspring of a particle carrying k marks, one can think of a single term in the sum $\langle \varphi, \mathcal{Z} \rangle_{k,n}$ as a weight associated to the event that the k marks are distributed among the offspring, by giving exactly k_i marks to the ith offspring particle, $i = 1, \dots, N$ (see below for a more precise interpretation). With this notation in hand, now define

$$\langle \varphi, \mathcal{Z} \rangle_k(x) := \sum_{1 \leq n \leq k} \varphi(x)^{-n} \langle \varphi, \mathcal{Z} \rangle_{k,n} \tag{4}$$

and set $\mathtt{m}_k(x) = \mathcal{E}_x(\langle\varphi, \mathcal{Z}\rangle_k(x))$. Note that in the case of local branching, $\langle\varphi, \mathcal{Z}\rangle_k \equiv N^k$ for any $x \in E$.

Definition 2. *The construction of $\tilde{X}$ under $\mathbb{Q}^k$ goes as follows.*

1. *Again, we begin with one particle at $x \in E$ carrying all the marks $\{1, \ldots, k\}$. In what follows, particles carrying marks are referred to as spines.*

2. *Any spine (that is, any particle carrying any number of marks) moves according to the semigroup*
$$\mathtt{P}_t[g](x) := \frac{1}{\zeta(\xi, 0)} \mathbf{E}_x \left[\zeta(\xi, t) g(\xi_t)\right], \qquad x \in E, \tag{5}$$
where $(\xi_s, s \geq 0)$ denotes the motion.

3. *Suppose a spine carries marks $\mathbf{b} = (b_1, \ldots, b_j)$. Then for each $1 \leq n \leq j$, an independent exponential clock rings at rate $\beta(x)\mathtt{m}_{j,n}(x)$,*
$$\mathtt{m}_{j,n}(x) := \varphi(x)^{-n} \mathcal{E}_x(\langle\varphi, \mathcal{Z}\rangle_{j,n}), \qquad x \in E.$$
When the first of these clocks rings, a branching event occurs, and if it is the nth clock, the j marks carried by the parent will be given to exactly n distinct offspring particles.

   More precisely, if the first clock to ring is the nth one, the positions of the offspring are described by $\mathcal{Z}$ with law $\mathcal{P}_x^{(j,n)}$ defined by
$$\frac{\mathrm{d}\mathcal{P}_x^{(j,n)}}{\mathrm{d}\mathcal{P}_x} := \frac{\langle\varphi, \mathcal{Z}\rangle_{j,n}}{\mathcal{E}_x(\langle\varphi, \mathcal{Z}\rangle_{j,n})}. \tag{6}$$
Then given $\mathcal{Z}$, for each $(k_1, \ldots, k_N) \in [k_1, \ldots, k_N]_j^n$, the probability that the ith offspring particle receives exactly k_i marks for each $1 \leq i \leq N$, is given by
$$\frac{\binom{k}{k_1, \ldots, k_N} \prod_{i: k_i > 0} \varphi(x_i)}{\langle\varphi, \mathcal{Z}\rangle_{j,n}}.$$
On this event, the way that the marks $b_1, \ldots, b_j$ are distributed among the offspring is that such that any valid configuration (that is, satisfying the constraint that exactly k_i marks are given to offspring particle i for each $1 \leq i \leq N$) has the same probability:
$$\frac{1}{\binom{k}{k_1, \ldots, k_N}}.$$

4. *Particles that do not carry marks issue independent copies of $(X, \mathbb{P})$. Marked particles then continue from Step 2.*

Remark 1. An alternative description of the third step above, is in terms of the *total* branching rate. Namely, suppose a spine carries marks $\mathbf{b} = (b_1, \ldots, b_j)$. Then it branches at rate $\beta(x)\mathtt{m}_j(x)$, and on such a branching event, the offspring positions are described by $\mathcal{Z}$

whose law is weighted (with respect to $\mathcal{P}_x$) by $\langle\varphi,\mathcal{Z}\rangle_j/\mathtt{m}_j(x)$. Moreover, given $\mathcal{Z}$ (and the position x of the spine before branching), any particular allocation of the marks $b_1,\dots,b_j$ among the N offspring (that is, a partition $S_1,\dots,S_N$ of $\{b_1,\dots,b_j\}$) has probability equal to $\langle\varphi,\mathcal{Z}\rangle_j^{-1}\prod_{i:|S_i|>0}(\varphi(x_i)/\varphi(x))$. However, we prefer to highlight the decomposition according to n (as given in the description of the measure above), since this will be the one we use in practice.

Remark 2. There are many different variants of the measure $\mathbb{Q}^k$ that we could have described, that would also be related to $\mathbb{P}^k$ by a martingale change of measure, and would also lead to a many-to-few type formula. Our specific choice of $\mathbb{Q}^k$ is motivated by our main application: describing the genealogical structure of the branching process when it is conditioned on survival. In particular, when we use our many-to-few formula for this purpose, we get an extremely simple structure - see (18). Note that in the case of local branching, this measure $\mathbb{Q}^k$ is identical to that which appears in [14]. When $k=1$, and we make a particular choice for ζ, this also agrees with the "spine decomposition" given in [11].

2.2 The martingale change of measure

We will now explain how the measures $\mathbb{P}^k$ and $\mathbb{Q}^k$ are connected via a change of measure. Let us first introduce some further notation.

Given $v\in\Omega$, note that $X_t(E\times\{v\})=0$ except on some unique (possibly empty) interval $[\sigma_v,\tau_v)$, on which $X_t(E\times\{v\})=1$. We will often use the notation τ_v^- for the left limit of τ_v. If $t\in[\sigma_v,\tau_v)$, there exists a unique $X_v(t)\in E$ such that $X_t(X_v(t)\times\{v\})=1$ and a unique $\mathbf{b}_v\in\mathcal{P}(\{1,\dots,k\})$ such that $\tilde{X}_t(E\times\{v\}\times\mathbf{b}_v)=1$ for all $t\in[\sigma_v,\tau_v)$. Heuristically, σ_v and τ_v are the birth and death times of particle v, respectively, $X_v(t)$ represents its position at time t during its lifetime, and $\mathbf{b}_v$ represents the set of marks it carries. We further set $D_v=|\mathbf{b}_v|$ to be the number of marks carried by the particle with label v. For each $v\in\Omega$, let N_v denote the number of offspring produced by $X_v(\tau_v)$.

Set $\mathcal{N}_t:=\{v\in\Omega : t\in[\sigma_v,\tau_v)\}$ so that $N_t=|\mathcal{N}_t|$. For each $t\geq 0$ and $j=1,\dots,k$, let ψ_t^j and ξ_t^j denote the unique elements of Ω and E, respectively, such that there exists $\mathbf{b}\in\mathcal{P}(\{1,\dots,k\})$ with $j\in\mathbf{b}$ and $\tilde{X}_t(\xi_t^j\times\{\psi_t^j\}\times\mathbf{b})=1$. Finally, let us define the skeleton at time t to be $S_k(t):=\{\psi_s^1,\dots,\psi_s^k : s\leq t\}$, so that $S_k(t)$ is a subset of labels in the tree. For $v\in S_k(t)$, let M_v denote the number of distinct offspring of v that are given a mark (that is, the number of distinct spine offspring).

Definition 3. *Define the $\mathcal{F}_t^k$-adapted process $(W_t^k, t\geq 0)$ by*

$$W_t^k := \prod_{v\in S_k(t)}\frac{\zeta(X_v,\tau_v^-\wedge t)}{\zeta(X_v,\sigma_v)}\prod_{v\in S_k(t)}\mathrm{e}^{-\int_{\sigma_v}^{\tau_v\wedge t}\beta(X_v(s))(\mathtt{m}_{D_v}(X_v(s))-1)\mathrm{d}s}$$
$$\times\prod_{v\in S_k(t)\setminus\{\emptyset\}}\varphi(X_v(\sigma_v))\prod_{v\in S_k(t)\setminus\mathcal{N}_t}\frac{N_v^{D_v}}{\varphi(X_v(\tau_v^-))^{M_v}}. \tag{7}$$

Remark 3. Henceforth, for $s\leq t$ and $x,y\in E$, we set $\zeta(x,t)/\zeta(y,s)=1$ whenever $\zeta(x,t)=\zeta(y,s)=0$.

Remark 4. We emphasise that then when a branching event occurs and the spines all choose the same particle to follow, this is included as an element of the skeleton, $S_k(t)$.

Remark 5. Note that we may equivalently write

$$
\begin{aligned}
W_t^k = &\prod_{v\in S_k(t)} \frac{\zeta(X_v, \tau_v^- \wedge t)}{\zeta(X_v,\sigma_v)} \\
&\times \prod_{v\in S_k(t)\setminus\mathcal{N}_t} \mathtt{m}_{D_v}(X_v(\tau_v^-))\mathrm{e}^{-\int_{\sigma_v}^{\tau_v}\beta(X_v(s))(\mathtt{m}_{D_v}(X_v(s))-1)\mathrm{d}s} \prod_{v\in\mathcal{N}_t} \mathrm{e}^{-\int_{\sigma_v}^{t}\beta(X_v(s))(\mathtt{m}_{D_v}(X_v(s))-1)\mathrm{d}s} \\
&\times \prod_{v\in S_k(t)\setminus\mathcal{N}_t} \frac{\langle\varphi, \mathcal{Z}_v\rangle_{D_v}(X_v(\tau_v^-))}{\mathtt{m}_{D_v}(X_v(\tau_v^-))} \\
&\times \prod_{v\in S_k(t)\setminus\{\emptyset\}} \frac{N_{p_v}^{D_{p_v}}}{\langle\varphi, \mathcal{Z}_{p_v}\rangle_{D_{p_v}}(X_v(\tau_{p_v}^-))} \frac{\varphi(X_v(\sigma_v))}{\varphi(X_v(\tau_{p_v}^-))^{M_{p_v}}},
\end{aligned}
$$

where p_v is the label of the parent of particle with label v. The above decomposition holds since

$$
\begin{aligned}
&\prod_{v\in S_k(t)\setminus\mathcal{N}_t} \frac{N_v^{D_v}}{\varphi(X_v(\tau_v^-))^{M_v}\langle\varphi, \mathcal{Z}_v\rangle_{D_v}(X_v(\tau_v^-))} \prod_{v\in S_k(t)\setminus\{\emptyset\}} \varphi(X_v(\sigma_v)) \\
&\quad = \prod_{v\in S_k(t)\setminus\{\emptyset\}} \frac{N_{p_v}^{D_{p_v}}}{\langle\varphi, \mathcal{Z}_{p_v}\rangle_{D_{p_v}}(X_v(\tau_{p_v}^-))} \frac{\varphi(X_v(\sigma_v))}{\varphi(X_v(\tau_{p_v}^-))^{M_{p_v}}},
\end{aligned}
$$

by noting that in the product on the left-hand side of the above equality, each of the terms in the first product is a parent of an element in the second product. Using Remark 1, one can see that each of the terms above describes a change of measure with respect to $\mathbb{P}^k$ for, in order: the motion; branch rate; offspring distribution; and selection of the spine particles immediately after a branching event. These changes of measure account for the differences between the pathwise constructions of the measures $\mathbb{P}^k$ and $\mathbb{Q}^k$. Indeed, we have the following result.

Proposition 4. *For $x \in E$, $(W_t^k, t \geq 0)$ is a martingale. Define*

$$
\left.\frac{\mathrm{d}\widetilde{\mathbb{Q}}_{\delta_x}^k}{\mathrm{d}\mathbb{P}_{\delta_x}^k}\right|_{\mathcal{F}_t^k} = W_t^k, \qquad t \geq 0,\ x \in E. \tag{8}
$$

Then for all $x \in E$, $\mathbb{Q}_{\delta_x}^k = \widetilde{\mathbb{Q}}_{\delta_x}^k$.

Proof. In the spirit of [3], and the proofs of Proposition 11 and Theorem 12 in [9], it suffices to demonstrate that the change of measure holds for the behaviour of the initial particle, up to and including its branching event. Thereafter, the Markov property ensures that the result is true in general.

To this end, let us suppose that T_1 is the first branch time. According to the definition of $\mathbb{Q}_{\delta_x}^k$ we have, for $x \in E$ and any bounded measurable H,

$$
\mathbb{Q}_{\delta_x}^k[H(\xi_s, 0 \leq s \leq t); t < T_1] = \mathbf{E}_x\left[\frac{\zeta(\xi,t)}{\zeta(\xi,0)} H(\xi_s, s \leq t)\mathrm{e}^{-\int_0^t \beta(\xi_s)\mathtt{m}_k(\xi_s)\mathrm{d}s}\right], \qquad t \geq 0,
$$

where, on $\{t < T_1\}$, $\xi = \xi^1 = \cdots = \xi^k$. Noting similarly that

$$\mathbb{P}^k_{\delta_x}[H(\xi_s, 0 \le s \le t); t < T_1] = \mathbf{E}_x\left[H(\xi_s, s \le t)\mathrm{e}^{-\int_0^t \beta(\xi_s)\mathrm{d}s}\right], \qquad t \ge 0,$$

it follows that

$$\mathbb{Q}^k_{\delta_x}[H(\xi_s, 0 \le s \le t); t < T_1] = \mathbb{P}^k_x\left[\frac{\zeta(\xi,t)}{\zeta(\xi,0)} H(\xi_s, s \le t)\mathrm{e}^{-\int_0^t \beta(\xi_s)(\mathtt{m}_k(\xi_s)-1)\mathrm{d}s}\right], \qquad t \ge 0,$$

which agrees with (8) on $\{t < T_1\}$.

Next, we extend this to include what happens at the first branch event. Again referring to the definition of $\mathbb{Q}^k_{\delta_x}$, for $t > 0$, $x \in E$, $n \in \mathbb{N}$ and H and ξ as before, $f \in B(E)$ and $i_1, \dots, i_k \in \mathbb{N}$ such that $|\{i_1, \dots, i_k\}| = n$, $L_{\mathbf{i}} := \{l \in \mathbb{N} \text{ s.t. } i_j = l \text{ for some } 1 \le j \le k\}$

$$\begin{aligned}
&\mathbb{Q}^k_{\delta_x}[H(\xi_s, 0 \le s \le t)\mathrm{e}^{-\langle f, \mathcal{Z}\rangle}\mathbf{1}_{\{T_1 \in \mathrm{d}t\}}\mathbf{1}_{\{i_1,\dots,i_k \in \{1,\dots N\}\}}\mathbf{1}_{\{\psi^1_t = i_1, \dots, \psi^k_t = i_k\}}] \\
&= \mathbb{P}^k_x\Bigg[\frac{\zeta(\xi,t)}{\zeta(\xi,0)} H(\xi_s, s \le t)\mathrm{e}^{-\langle f, \mathcal{Z}\rangle}\mathrm{e}^{-\int_0^t \beta(\xi_s)(\mathtt{m}_k(\xi_s)-1)\mathrm{d}s} \\
&\qquad \times \mathbf{1}_{\{i_1,\dots,i_k \in \{1,\dots N\}\}}\mathbf{1}_{\{\psi^1_t = i_1, \dots, \psi^k_t = i_k\}}\beta(\xi_t)\mathtt{m}_k(\xi_t)\frac{\mathtt{m}_{k,n}(\xi_t)}{\mathtt{m}_k(\xi_t)}\frac{\langle \varphi, \mathcal{Z}\rangle_{k,n}}{\mathcal{E}_{\xi_t}(\langle \varphi, \mathcal{Z}\rangle_{k,n})}\frac{\prod_{l \in L_{\mathbf{i}}}\varphi(x_l)}{\langle \varphi, \mathcal{Z}\rangle_{k,n}}\frac{1}{1/N^k}\Bigg]\mathrm{d}t, \\
&= \mathbb{P}^k_x\Bigg[\frac{\zeta(\xi,t)}{\zeta(\xi,0)} H(\xi_s, s \le t)\mathrm{e}^{-\langle f, \mathcal{Z}\rangle}\mathrm{e}^{-\int_0^t \beta(\xi_s)(\mathtt{m}_k(\xi_s)-1)\mathrm{d}s} \\
&\qquad\qquad \times \mathbf{1}_{\{i_1,\dots,i_k \in \{1,\dots N\}\}}\mathbf{1}_{\{\psi^1_t = i_1, \dots, \psi^k_t = i_k\}}\beta(\xi_t)\varphi(\xi_t)^{-n}\prod_{l \in L_{\mathbf{i}}}\varphi(x_l)\, N^k\Bigg]\mathrm{d}t,
\end{aligned}$$

where the factor $1/N^k$ is removing the selection bias of the k-spines under $\mathbb{P}^k$. Noting that, in the above calculation, on $\{T_1 \in \mathrm{d}t\}$, $S_k(t)\backslash\mathcal{N}_t = \{\emptyset\}$, the set $S_k(t)\backslash\{\emptyset\}$, agrees with the offspring of $\emptyset$, $\varphi(X_\emptyset(\tau_\emptyset^-)) = \varphi(\xi_t)$, $M_\emptyset = n$ and $N_\emptyset = k$. As such we note that on $\{T_1 \in \mathrm{d}t\}$, the change of measure (8) is valid. □

2.3 Spines at different times

We would also like to consider the "skeleton at different times". To this end, fix $k \ge 1$, suppose $0 \le s_k \le \cdots \le s_1$ and write $\mathbf{s} = (s_1, \dots, s_k)$. Let

$$\mathcal{N}_{\mathbf{s}} := \{(v_1, \dots, v_k) : v_i \in \mathcal{N}_{s_i}\ 1 \le i \le k\}$$

and for $\mathbf{v} = (v_1, \dots, v_k) \in \mathcal{N}_{\mathbf{s}}$, let

$$S_k(\mathbf{v}, \mathbf{s}) := \{w \in \Omega : w \preceq v_i \text{ for some } 1 \le i \le k\}$$

be the "skeleton" formed by the ancestors of v_i up to times s_i. We also write

$$S_k(\mathbf{s}) := S_k((\psi^1_{s_1}, \dots, \psi^k_{s_k}), \mathbf{s})$$

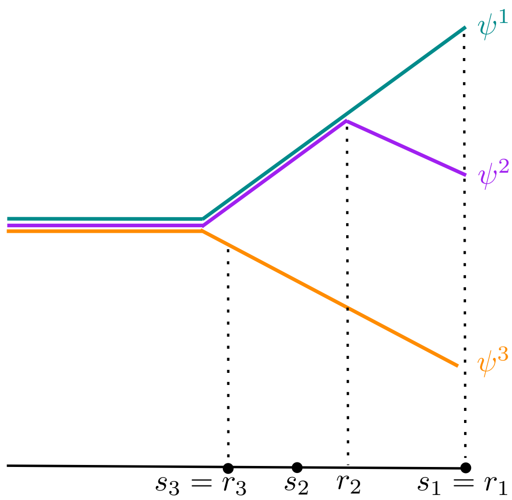

Figure 1: The spines carrying marks 1, 2 and 3 are depicted in cyan, purple and orange respectively. The times s_i and r_i for $i = 1, 2, 3$ are shown (the dotted vertical lines correspond to r_1, r_2 and r_3).

for the skeleton generated by the spines and

$$\partial S_k(\mathbf{s}) = \{v \in S_k(\mathbf{s}) \text{ such that } \nexists\, w \in S_k(\mathbf{s}) \text{ with } v \prec w\}$$

for the "leaves" of this skeleton. If for each $1 \le i \le k$ we set

$$r_i = \sup\{s > 0 : \psi^i_s = \psi^k_s \text{ for some } k \text{ with } s_k \ge s\}.$$

then for each i with $r_i = s_i$, $v = \psi^i_{s_i}$ is an element of $\partial S_k(\mathbf{s})$. We define a further set of labels associated to those i for which $r_i > s_i$:

$$\hat{\partial} S_k(\mathbf{s}) = \{\psi^i_{r_i},\, 1 \le i \le k, r_i > s_i\}.$$

Note that while $\partial S_k(\mathbf{s})$ is a subset of $S_k(\mathbf{s})$, $\hat{\partial} S_k(\mathbf{s})$ is not. Finally, for $v = \psi^i_{r_i} \in \hat{\partial} S_k(\mathbf{s}) \cup \partial S_k(\mathbf{s})$ we define

$$r_v := r_i \tag{9}$$

(so for $v \in \psi^i_{s_i} \in \partial S_k(\mathbf{s})$, we equivalently have $r_v = s_i$). See Figure 1.

Now fix $t \ge s_1$ and let $\mathcal{F}^k_{t,\mathbf{s}}$ denote the σ-algebra generated by:

- $\{\xi^i_s : s \le s_i, 1 \le i \le k\}$ (the motion of the spine with mark i up to time s_i for each i);
- $\{\psi^i_s : s \le s_i, 1 \le i \le k\}$ (the Ulam-Harris labels associated to the the spine with mark i up to time s_i for each i);
- $\{\mathbf{b}_{\psi^i_s} : s \le s_i, 1 \le i \le k\}$ (the collection of marks carried by the spine with mark i up to time s_i for each i);
- the subtree rooted at each $w \in \Omega$ that does not carry any marks and is a sibling of *some* $v \in S_k(\mathbf{s})$[1], considered up until (global) time t.

[1]where by subtree we mean the subprocess started at time σ_w with root label w

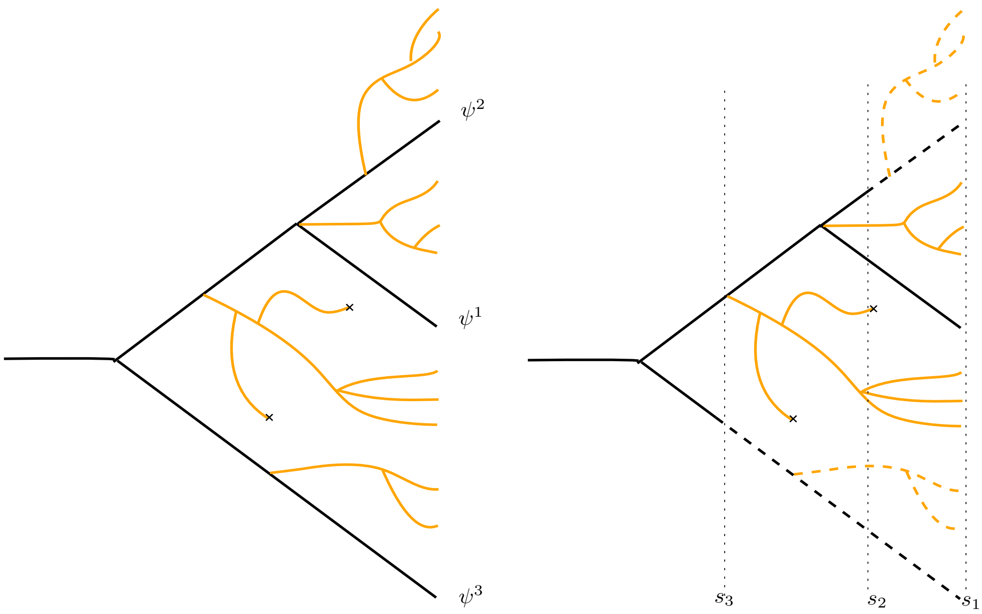

Figure 2: The left-hand side figure shows the tree up to time s_1 with three spines marked in black. The right-hand side shows the information from the filtration $\mathcal{F}^3_{s_1,\mathbf{s}}$, with $\mathbf{s} = (s_1, s_2, s_3)$, with dashed lines denoting information that is not included.

Note that the collection of random variables in the third bullet point above will not always be measurable with respect to the collection in the second. For example, if $k = 2$ and the spine carrying mark 1 also carries mark 2 at time $s \in (s_2, s_1)$, then since $\{\psi^i_s : s \le s_i, i = 1, 2\}$ does not tell us about the labels associated to the spine with mark 2 after time s_2, $\mathbf{b}_{\psi^1_s}$ is not measurable with respect to it.

We will also use the notation

$$\mathcal{F}^k_{\mathbf{s}} := \mathcal{F}^k_{s_1,\mathbf{s}}.$$

For each $\mathbf{s}$, we define an $\mathcal{F}^k_{\mathbf{s}}$-measurable random variable

$$W^k_{\mathbf{s}} := \prod_{v \in S_k(\mathbf{s})} \frac{\zeta(X_v, \tau_v^- \wedge r_v)}{\zeta(X_v, \sigma_v)} \mathrm{e}^{-\int_{\sigma_v}^{\tau_v \wedge r_v} \beta(X_v(s))(\mathtt{m}_{D_v}(X_v(s))-1)\mathrm{d}u}$$
$$\times \prod_{v \in S_k(\mathbf{s}) \setminus \{\emptyset\}} \varphi(X_v(\sigma_v)) \prod_{v \in S_k(\mathbf{s}) \setminus \partial S_k(\mathbf{s})} \frac{N_v^{D_v}}{\varphi(X_v(\tau_v^-))^{M_v}}, \tag{10}$$

where we have used the notation r_v defined in (9) for $v \in \partial S_k(\mathbf{s})$ (and set $r_v = \infty$ otherwise).

Recalling Lemma 4, restricting instead to $\mathcal{F}^k_{\mathbf{s}}$ yields the following result.

Lemma 5. *For each $k \ge 1$, $0 \le s_k \le \ldots \le s_1$ and $x \in E$, we have*

$$\left.\frac{\mathrm{d}\mathbb{Q}^k_{\delta_x}}{\mathrm{d}\mathbb{P}^k_{\delta_x}}\right|_{\mathcal{F}^k_{\mathbf{s}}} = W^k_{\mathbf{s}}. \tag{11}$$

Proof. Fix $k \geq 1$, $0 \leq s_k \leq \cdots \leq s_1 \leq t$ and $x \in E$. Then, due to the structure of W_t^k, we may write

$$
\begin{aligned}
W_t^k = W_{\mathbf{s}}^k \times & \prod_{v \in S_k(t) \setminus S_k(\mathbf{s})} \frac{\zeta(X_v, \tau_v^- \wedge t)}{\zeta(X_v, \sigma_v)} \mathrm{e}^{-\int_{\sigma_v}^{\tau_v \wedge t} \beta(X_v(s))(\mathtt{m}_{D_v}(X_v(s))-1)\mathrm{d}u} \\
& \times \prod_{v \in \partial S_k(\mathbf{s})} \frac{\zeta(X_v, \tau_v^- \wedge t)}{\zeta(X_v, r_v)} \mathrm{e}^{-\int_{r_v}^{\tau_v \wedge t} \beta(X_v(s))(\mathtt{m}_{D_v}(X_v(s))-1)\mathrm{d}u} \\
& \times \prod_{v \in V_{\mathbf{s},t}} \frac{N_v^{D_v}}{\varphi(X_v(\tau_v^-))^{M_v}} \prod_{v \in S_k(t) \setminus S_k(\mathbf{s})} \varphi(X_v(\sigma_v)), \qquad (12)
\end{aligned}
$$

where $V_{\mathbf{s},t} := (S_k(t) \setminus S_k(\mathbf{s}) \cup \mathcal{N}_t) \cup \partial S_k(\mathbf{s})$ (in words, the difference of the skeletons at times t and $\mathbf{s}$, minus the boundary at time t, but plus the boundary at time $\mathbf{s}$). Hence,

$$
\begin{aligned}
\mathbb{P}_{\delta_x}^k[W_t^k | \mathcal{F}_{\mathbf{s}}^k] = W_{\mathbf{s}}^k \times \mathbb{P}_{\delta_x}^k \Bigg[& \prod_{v \in S_k(t) \setminus S_k(\mathbf{s})} \frac{\zeta(X_v, \tau_v^- \wedge t)}{\zeta(X_v, \sigma_v)} \mathrm{e}^{-\int_{\sigma_v}^{\tau_v \wedge t} \beta(X_v(s))(\mathtt{m}_{D_v}(X_v(s))-1)\mathrm{d}u} \\
& \times \prod_{v \in \partial S_k(\mathbf{s})} \frac{\zeta(X_v, \tau_v^- \wedge t)}{\zeta(X_v, r_v)} \mathrm{e}^{-\int_{r_v}^{\tau_v \wedge t} \beta(X_v(s))(\mathtt{m}_{D_v}(X_v(s))-1)\mathrm{d}u} \\
& \times \prod_{v \in V_{\mathbf{s},t}} \frac{N_v^{D_v}}{\varphi(X_v(\tau_v^-))^{M_v}} \prod_{v \in S_k(t) \setminus S_k(\mathbf{s})} \varphi(X_v(\sigma_v)) \Bigg| \mathcal{F}_{\mathbf{s}}^k \Bigg]. \qquad (13)
\end{aligned}
$$

Now consider the collection of subprocesses initiated at times r_v for each $v \in \partial S_k(\mathbf{s}) \cup \hat{\partial} S_k(\mathbf{s})$. The branching Markov property implies that these are independent of each other and of $\mathcal{F}_{\mathbf{s}}^k$. Moreover, we may rewrite the right-hand side of (13) as

$$
W_{\mathbf{s}}^k \times \mathbb{P}_{\delta_x}^k \Bigg[\prod_{v \in \hat{\partial} S_k(\mathbf{s}) \cup \partial S_k(\mathbf{s})} W_{t_v}^{(v)} \Bigg| \mathcal{F}_{\mathbf{s}}^k \Bigg],
$$

where for each $v \in \partial S_k(\mathbf{s}) \cup \hat{\partial} S_k(\mathbf{s})$, $W^{(v)}$ is a copy of the martingale W^{D_v} associated to the subprocess rooted at v and $t_v = t - r_v$. In particular,

$$
\mathbb{P}_{\delta_x}^k \Bigg[\prod_{v \in \hat{\partial} S_k(\mathbf{s}) \cup \partial S_k(\mathbf{s})} W_{t_v}^{(v)} \Bigg| \mathcal{F}_{\mathbf{s}}^k \Bigg] = \prod_{v \in \hat{\partial} S_k(\mathbf{s}) \cup \partial S_k(\mathbf{s})} \mathbb{P}_{\delta_x}^k [W_{t_v}^{(v)}] = 1,
$$

which gives the result. □

3 Many-to-few lemma

3.1 Statement and proof of the many-to-few lemma

We are now ready to formulate and prove our main result, which is a many-to-few lemma for general non-local branching Markov processes at a collection of different times. This generalises the result of [14] in two ways: firstly, it holds for non-local branching mechanisms; and secondly, it allows us to deal with sums over the population at different times.

Lemma 6 (Many-to-few at different times)**.** *Let $x \in E$, $k \geq 1$ and $0 \leq s_k \leq \cdots \leq s_1$ be fixed. Suppose that*

$$Y = \sum_{v_i \in \mathcal{N}_{s_i},\, i=1,\dots,k} Y(v_1,\dots,v_k)\mathbf{1}_{\{\psi^i_{s_i}=v_i,\, 1\leq i\leq k\}}$$

is non-negative and $\mathcal{F}^k_{\mathbf{s}}$-measurable, where $Y(v_1,\dots,v_k)$ is $\mathcal{F}_{s_1}$-measurable[2] *for every $(v_1,\dots,v_k) \in \mathcal{N}_{\mathbf{s}}$. Then it holds that*

$$\mathbb{P}_{\delta_x}\Bigg[\sum_{\substack{v_i\in\mathcal{N}_{s_i}\\ i=1,\dots,k}} Y(v_1,\dots,v_k)\Bigg] = \mathbb{Q}^k_{\delta_x}\Bigg[Y \prod_{v\in S_k(\mathbf{s})} \frac{\zeta(X_v,\sigma_v)}{\zeta(X_v,\tau_v^-\wedge s_v)} \mathrm{e}^{\int_{\sigma_v}^{\tau_v\wedge s_v}\beta(X_v(s))(\mathtt{m}_{D_v}(X_v(s))-1)\mathrm{d}u}$$
$$\times \frac{\prod_{v\in S_k(\mathbf{s})\setminus\partial S_k(\mathbf{s})}\varphi(X_v(\tau_v^-))^{M_v}}{\prod_{v\in S_k(\mathbf{s})\setminus\{\emptyset\}}\varphi(X_v(\sigma_v))} \prod_{v\in S_k(\mathbf{s})\setminus\partial S_k(\mathbf{s})} N_v^{-D_v} \prod_{i=1}^k \prod_{\emptyset\preceq v\prec\psi^i_{s_i}} N_v\Bigg].$$

Remark 6. Recall that

$$S_k(\mathbf{s})\setminus\partial S_k(\mathbf{s}) = \{w\in\Omega : \emptyset\preceq w\prec\psi^i_{s_i} \text{ for some } 1\leq i\leq k\}.$$

Also recall that the mark carried by an individual (spine) v is $\mathbf{b}_v$ with cardinality $D_v = |\mathbf{b}_v|$ (that is the number of marks carried by v) and M_v is the number of offspring of v that inherit a mark from v. As a consequence,

$$|\{i\in\{1,\dots,k\} : \emptyset\preceq v\prec\psi^i_{s_i}\}| \leq D_v. \tag{14}$$

The inequality in (14) can be strict when, for example, $\emptyset\preceq v=\psi^j_{s_j}\prec\psi^i_{s_i}$ for some pair $i\neq j$.

Proof of Lemma 6. Let us start by rewriting the right-hand side of the expression in the lemma. Noting that under $\mathbb{Q}_{\delta_x}$, $W^k_{\mathbf{s}}$ is positive, we have

$$\mathbb{Q}^k_{\delta_x}\Bigg[\frac{Y}{W^k_{\mathbf{s}}}\prod_{i=1}^k\prod_{\emptyset\preceq v\prec\psi^i_{s_i}} N_v\Bigg] = \mathbb{P}^k_{\delta_x}\Bigg[Y\prod_{i=1}^k\prod_{\emptyset\preceq v\prec\psi^i_{s_i}} N_v\Bigg]$$
$$= \mathbb{P}^k_{\delta_x}\Bigg[\sum_{u_i\in\mathcal{N}_{s_i},\,1\leq i\leq k} Y(u_1,\dots,u_k)\mathbf{1}_{\{\psi^i_{s_i}=u_i\ 1\leq i\leq k\}}\prod_{i=1}^k\prod_{\emptyset\preceq v\prec u_i} N_v\Bigg] \tag{15}$$

where the first equality holds thanks to the change of measure (11).

Conditioning on $\mathcal{F}_{s_1}$, we have

$$\mathbb{P}^k_{\delta_x}\left(\psi^i_{s_i}=u_i,\, i=1,\dots,k\,\middle|\,\mathcal{F}_{s_1}\right) = \prod_{i=1}^k\prod_{\emptyset\preceq v\prec u_i} N_v^{-1},$$

and thus, by the properties of conditional expectation, we can rewrite (15) as

$$\mathbb{P}^k_{\delta_x}\Bigg[\sum_{u_i\in\mathcal{N}_{s_i},\,1\leq i\leq k} Y(u_1,\dots,u_k)\prod_{i=1}^k\prod_{\emptyset\preceq v\prec u_i} N_v\prod_{\emptyset\preceq v\prec u_i} N_v^{-1}\Bigg] = \mathbb{P}_{\delta_x}\Bigg[\sum_{v_i\in\mathcal{N}_{s_i},\,1\leq i\leq k} Y(v_1,\dots,v_k)\Bigg],$$

as required. □

[2] By this we mean that we have a collection $(Y(v_1,\dots,v_k), v_1,\dots,v_k\in\Omega)$ of $\mathcal{F}_{s_1}$-measurable random variables such that $Y(v_1,\dots,v_k)=0$ unless $v_i\in\mathcal{N}_{s_i}$ for all $i=1,\dots,k$.

Referring to Remark 6, if we take $(s_1, \dots, s_k) = (t, \dots, t)$ then $v = \psi_t^j \prec \psi_t^i$ cannot occur and the inequality in (14) is an equality. In that case,

$$\prod_{i=1}^{k} \prod_{\emptyset \preceq v \prec \psi_{s_i}^i} N_v = \prod_{v \in S_k(\mathbf{s}) \setminus \partial S_k(\mathbf{s})} N_v^{D_v},$$

which yields the following corollary (as a special case).

Corollary 7 (Many-to-few)**.** *Let $x \in E$, $k \geq 1$ and $t \geq 0$ be fixed. Suppose that*

$$Y = \sum_{v_i \in \mathcal{N}_t,\, i=1,\dots,k} Y(v_1, \dots, v_k) \mathbf{1}_{\{\psi_t^i = v_i,\, 1 \leq i \leq k\}}$$

is $\mathcal{F}_t^k$-measurable with $Y(v_1, \dots, v_k)$ $\mathcal{F}_t$-measurable for $v_i \in \mathcal{N}_t$ for $i = 1, \dots k$. Then

$$\mathbb{P}_{\delta_x}\Bigg[\sum_{\substack{v_i \in \mathcal{N}_t \\ i=1,\dots,k}} Y(v_1, \dots, v_k)\Bigg] = \mathbb{Q}_{\delta_x}^k\Bigg[Y \prod_{v \in S_k(t)} \frac{\zeta(X_v, \sigma_v)}{\zeta(X_v, \tau_v^- \wedge t)} \mathrm{e}^{\int_{\sigma_v}^{\tau_v \wedge t} \beta(X_v(s))(\mathtt{m}_{D_v}(X_v(s)) - 1)\mathrm{d}s}$$
$$\times \frac{\prod_{v \in S_k(t) \setminus \mathcal{N}_t} \varphi(X_v(\tau_v^-))^{M_v}}{\prod_{v \in S_k(t) \setminus \{\emptyset\}} \varphi(X_v(\sigma_v))}\Bigg].$$

3.2 Natural choice of ζ and separated skeletons

We will now consider a specific example of the many-to-few formula given in Lemma 6 by choosing a particular form for the martingale ζ. Using the fact that φ is the right eigenfunction for the MBP, $(X, \mathbb{P})$, with corresponding eigenvalue λ, it isn't too difficult to show that under $\mathbf{P}_x$

$$\zeta(\xi, t) = \frac{\varphi(\xi_t)}{\varphi(x)} \exp\left(\int_0^t \beta(\xi_s)(\mathtt{m}_1(\xi_s) - 1)\mathrm{d}s\right), \quad t \geq 0, \tag{16}$$

is a martingale.

In this case, we have that

$$W_t^k = \frac{1}{\varphi(x)} \prod_{v \in S_k(t)} \mathrm{e}^{-\int_{\sigma_v}^{\tau_v \wedge t} \beta(X_v(s))(\mathtt{m}_{D_v}(X_v(s)) - \mathtt{m}_1(X_v(s)))\mathrm{d}s}$$
$$\times \prod_{v \in S_k(t) \cap \mathcal{N}_t} \varphi(X_v(t)) \prod_{v \in S_k(t) \setminus \mathcal{N}_t} \frac{N_v^{D_v}}{\varphi(X_v(\tau_v^-))^{M_v - 1}}. \tag{17}$$

We now further assume that each of the nodes $v_i \in \mathcal{N}_{s_i}$, $i = 1, \dots, k$, that make up the skeleton $S_k(\mathbf{v}, \mathbf{s})$, are distinct. We refer to a skeleton with this property as a *separated skeleton.* This implies that at time s_i, node v_i only carries one mark for each $i = 1, \dots, k$. Recalling again Remark 6, we thus have the advantage of writing

$$\prod_{v \in S_k(\mathbf{s}) \setminus \mathcal{N}_\mathbf{s}} N_v^{-D_v} = \prod_{i=1}^{k} \prod_{\emptyset \leq v < \psi_{s_i}^i} N_v^{-1}.$$

Applying the many-to-few formula with this martingale, for the special choice of ζ, and in the case of separated skeletons, we get

$$\begin{aligned}
&\mathbb{P}^k_{\delta_x}\Bigg[\sum_{v_i\in\mathcal{N}_t \text{ distinct}} Y(v_1,\dots,v_k)\Bigg]\\
&= \varphi(x)\mathbb{Q}^k_{\delta x}\Bigg[Y\mathbf{1}_{\{\{\psi^i_t\}_{1\le i\le k}\text{ distinct}\}}\prod_{j=1}^k \varphi(\xi^j_t)^{-1}\\
&\qquad\times\prod_{v\in S_k(t)} \mathrm{e}^{\int_{\sigma_v}^{\tau_v\wedge t}\beta(X_v(s))(\mathtt{m}_{D_v}(X_v(s))-\mathtt{m}_1(X_v(s)))\mathrm{d}s}\prod_{v\in S_k(t)\setminus\mathcal{N}_t}\varphi(X_v(\tau_v^-))^{M_v-1}\Bigg]. \qquad (18)
\end{aligned}$$

4 Application to genealogies in the critical case

As an application, we determine the asymptotic law of the death time of the most recent common ancestor, henceforth referred to as *split time*, of two particles sampled uniformly from a critical population at two different times. The limit is taken as $t\to\infty$, when we have conditioned on survival of the process up to time t.

We assume in this section that the measures $\mathbb{Q}^k$ and $\mathbb{P}^k$ are as defined in section 3.2. We remind the reader of the notation $\xi^1,\dots,\xi^k$ for the motion of the k spines under $\mathbb{Q}^k$ and $\mathbb{P}^k$, and $\psi^1,\dots,\psi^k$ for the Ulam-Harris labels that they carry.

We assume throughout the section that our branching process satisfies the following additional criticality requirement.

Assumption 2. *The following criticality assumptions hold.*

1. $\lambda = 0$, *where* λ *was introduced in Assumption 1*
2. *Define* $\Delta_t := \sup_{x\in E, f\in B_1^+(E)} |\varphi(x)^{-1}\mathtt{T}_t[f](x) - \langle f,\tilde\varphi\rangle|$. *Then*
$$\sup_{t\ge 0}\Delta_t<\infty \text{ and } \lim_{t\to\infty}\Delta_t = 0.$$
3. *The number of offspring produced at a branching event is bounded above by* $n_{max}<\infty$.
4. *There exists a constant* $C>0$ *such that for all* $g\in B^+(E)$,
$$\Sigma := \langle\beta\mathbb{V}[g],\tilde\varphi\rangle \ge C\langle g,\tilde\varphi\rangle^2,$$
where $\mathbb{V}[g](x) := \mathcal{E}_x[\langle g,\mathcal{Z}\rangle_{2,2}]$ *and where the notation* $\langle g,\mathcal{Z}\rangle_{k,n}$ *was defined in* (3).
5. *For all* $x\in E$, $\mathbb{P}_{\delta_x}(\exists t>0 \text{ such that } N_t = 0) = 1$.

Remark 7. We note that the assumptions above are inherited from [11, 8], where asymptotic results concerning criticality and moment growth were considered. Whereas the Assumption 2.1 is clearly a standard criticality assumption, the remaining assumptions 2.2 - 2.5 can be interpreted as follows. Roughly speaking, Assumption 2.2 describes the uniform stability of the mean semigroup. In particular, by taking $f \equiv 1$, Assumption 2.2 ensures that the first moment of the process settles down to a stationary value. Assumption 2.3 rather obviously requires the random number of offspring to be deterministically bounded. Assumption 2.4 can be thought of as an irreducibility condition written in terms of the two-point correlation (or variance) functional $\mathbb{V}[g]$ and ensures a minimal level of spatial mixing occurring for second order effects associated to the semigroup assumption in Assumption 2.2. Finally Assumption 2.5 ensures that, even though at criticality in Assumption 2.1, we can guarantee there is extinction almost surely. Assumption 2.5 is automatically satisfied e.g. for a branching Brownian motion in a compact domain with killing on the boundary, or the category of neutron branching process considered in [16, 10, 5].

Let us now present our main application. Here and in the rest of this section $\Rightarrow$ means convergence in distribution.

Proposition 8. *Let $0 < a < 1$ and let $x \in E$ be fixed. Let T_t have the $\mathbb{P}_{\delta_x}(\cdot|N_t > 0)$ law of the split time of two particles: one chosen uniformly from those alive at time t and one chosen uniformly from those alive at time at. Then*

$$\frac{T_t}{t} \Rightarrow T \qquad \text{as } t \to \infty$$

where T has density

$$f_a(u) := \frac{2a}{1-a} \frac{2(a-u)\log(1-\frac{u}{a}) - (2-u-\frac{u}{a})\log(1-u)}{u^3}$$

with respect to Lebesgue measure $\mathrm{d}u$ *on* $[0, a]$.

Remark 8. Note that

$$\lim_{a \nearrow 1} \frac{2a}{1-a} \frac{2(a-u)\log(1-\frac{u}{a}) - (2-u-\frac{u}{a})\log(1-u)}{u^3} = 2(-2u + (u-2)\log(1-u))$$

This agrees with the density in the case $a = 1$ (for critical GW processes) calculated in [13].

Remark 9. *It is not obvious a priori that f_a is the density function of a random variable. However, this is indeed the case, since we can calculate that*

$$\begin{aligned}
&\frac{2a}{1-a} \int_0^a \frac{2(a-u)\log(1-\frac{u}{a}) - (2-u-\frac{u}{a})\log(1-u)}{u^3} du \\
=&\frac{2}{a(1-a)} \int_0^1 \frac{2a(1-u)\log(1-u) - (2-ua-u)\log(1-au)}{u^3} du \\
=&\frac{2}{a(1-a)} \left[\frac{(1-u)}{u^2} \left(a(u-1)\log(1-u) + (1-au)\log(1-au) \right) \right]_0^1 \\
=&-\frac{2}{a(1-a)} \lim_{u \to 0} \frac{a(u-1)(-u-\frac{u^2}{2} + \mathrm{o}(u^2)) + (1-au)(-au - \frac{a^2u^2}{2} + \mathrm{o}(u^2))}{u^2} \\
=&1.
\end{aligned}$$

We also have the following result concerning the joint convergence of the (normalised) population size at two different times under $\mathbb{Q}^1$. Recall from Assumption 2.3 that $\Sigma := \langle \beta \mathbb{V}[\varphi], \tilde{\varphi} \rangle$.

Proposition 9. *Under $\mathbb{Q}^1_{\delta_x}$,*

$$\left(\frac{N_{at}}{t}, \frac{N_t}{t} \right) \Rightarrow (Z, \hat{Z}), \qquad \text{as } t \to \infty,$$

where Z is equal in law to a $\texttt{Gamma}(2, (a\Sigma\langle 1, \tilde{\varphi}\rangle/2)^{-1})$ random variable and, conditionally on Z, the law of $\hat{Z}$ is that of a $\texttt{Gamma}(2 + K, (\Sigma(1-a)\langle 1, \tilde{\varphi}\rangle/2)^{-1})$ random variable where $K \sim \texttt{Poisson}(((1-a)\Sigma\langle 1, \tilde{\varphi}\rangle/2)^{-1} Z)$.

Equivalently, under $\mathbb{P}_{\delta_x}(\cdot | N_t > 0)$, we have the joint convergence of $(N_{at}/t, N_t/t)$ to $(Y, \hat{Y})$, where the joint law of $(Y, \hat{Y})$ is that of $(Z, \hat{Z})$ weighted by *$1/\hat{Z}$.*

Before moving on to the proofs of the above two propositions, we first state a lemma that will be used throughout the aforementioned proofs.

Lemma 10. *In what follows, we suppose that $(g_t, t \geq 0)$ are a collection of functions with $g_t \in B_1^+(E)$ for each $t > 0$, and such that $g_t \to g \in B_1^+(E)$ pointwise as $t \to \infty$. For any $x \in E$, the following hold.*

(a) $t\mathbb{P}_{\delta_x}(N_t > 0) \to \dfrac{2\varphi(x)}{\Sigma}$ *as $t \to \infty$,* $\sup_{t,x} |t\mathbb{P}_{\delta_x}(N_t > 0)| < \infty$ *and* $\inf_t |t\mathbb{P}_{\delta_x}(N_t > 0)| > 0$.

(b) The joint law of

$$\left(\frac{X_t[g_t]}{t}, \frac{N_t}{t} \right) \text{ under } \mathbb{P}_{\delta_x}(\cdot \,|\, N_t > 0)$$

converges to that of $(\langle g, \tilde{\varphi}\rangle Z, \langle 1, \tilde{\varphi}\rangle Z)$ as $t \to \infty$, where $Z \sim \texttt{Exponential}(2/\Sigma)$.

(c) The joint law of

$$\left(\frac{X_t[g_t]}{t}, \frac{N_t}{t}, \xi_t \right) \text{ under } \mathbb{Q}^1_{\delta_x} \tag{19}$$

converges to that of $(\langle g, \tilde{\varphi}\rangle Z, \langle 1, \tilde{\varphi}\rangle Z, \bar{\xi})$ as $t \to \infty$, where $Z \sim \texttt{Gamma}(2, 2/\Sigma)$ and $\bar{\xi}$ is independent of Z, with law given by

$$P(\bar{\xi} \in A) = \langle \mathbf{1}_A \varphi, \tilde{\varphi} \rangle \tag{20}$$

Proof. (a) follows from [11, Theorem 1.2], [11, Lemma 7.4] and [11, Lemma 7.2]. For the proof of (b), first note that the joint convergence of $(X_t[g]/t, N_t/t)$ under $\mathbb{P}_{\delta_x}(\cdot \,|\, N_t > 0)$ follows from [11, Theorem 1.3], together with the fact that $t^{-1}[X_t[f] - \langle f, \tilde{\varphi}\rangle X_t[\varphi]] \to 0$ in probability as $t \to \infty$ for any bounded f (see the proof of [11, Theorem 1.3]). Writing

$$X_t[g_t] = X_t[g] + X_t[g_t - g],$$

it thus suffices to show that $X_t[g_t - g]/t \to 0$ in $\mathbb{P}_{\delta_x}(\cdot|N_t > 0)$-probability. For this, we will show that we actually have L^1 convergence. Setting $f_t = |g_t - g|$ (which is bounded by 2), we have

$$\begin{aligned}\frac{1}{t}\mathbb{E}_{\delta_x}[|X_t[g_t - g]||N_t > 0] &\le \frac{1}{t\mathbb{P}_{\delta_x}(N_t > 0)}\mathbb{E}_{\delta_x}[X_t[f_t]] \\ &\le \frac{1}{t\mathbb{P}_{\delta_x}(N_t > 0)}(|\mathbb{E}_{\delta_x}[X_t[f_t]] - \varphi(x)\langle f_t, \tilde{\varphi}\rangle| + \varphi(x)\langle f_t, \tilde{\varphi}\rangle).\end{aligned}$$

From Lemma 10(a), it follows that $(t\mathbb{P}_{\delta_x}(N_t > 0))^{-1}$ is uniformly bounded. By the first part of Assumption 2, the first term in the parentheses on the right-hand side above converges to 0 uniformly. Finally, by dominated convergence (since $\tilde{\varphi}$ is a finite measure), the second term in the parentheses also converges to 0.

For (c), note that if W_t^1 is the martingale from (11) (in the classical case of one spine),

$$\mathbb{P}_{\delta_x}[W_t^1|\mathcal{F}_t] = X_t[\varphi], \tag{21}$$

so that for any bounded, continuous function F,

$$\mathbb{Q}^1_{\delta_x}\left[F\left(\frac{X_t[g_t]}{t}, \frac{N_t}{t}\right)\right] = \frac{t\mathbb{P}_{\delta_x}(N_t > 0)}{\varphi(x)}\mathbb{P}_{\delta_x}\left[F\left(\frac{X_t[g_t]}{t}, \frac{N_t}{t}\right)\frac{X_t[\varphi]}{t}\,\middle|\, N_t > 0\right]. \tag{22}$$

Taking F to be of the form $F(x, y) = \mathrm{e}^{-\theta x - \mu y}$, for $\theta, \mu \ge 0$ and using (a) and (b) yields the convergence of the first two components of the triple under $\mathbb{Q}^1_{\delta_x}$. The marginal convergence in law of ξ_t to ξ under $\mathbb{Q}_{\delta_x}$ follows from [11, section 5].

To see the joint convergence in law of the triple in (19), note that due to the aforementioned marginal convergence of ξ and the first two components, we immediately have tightness. Moreover, any subsequential limit has the form $(\langle g, \tilde{\varphi}\rangle Z, \langle 1, \tilde{\varphi}\rangle Z, \bar{\xi})$, where the marginals of Z and $\bar{\xi}$ are as desired. Thus it remains to show that for any such subsequential limit, Z and $\bar{\xi}$ (or equivalently $\langle 1, \tilde{\varphi}\rangle Z$ and $\bar{\xi}$) are independent.

For this, define N_t^* to be the contribution to N_t of all descendants branching off the (single) spine particle *before* time $t - t^{1/3}$. Then $N_t - N_t^*$ behaves like $N_{t^{1/3}}$ under $\mathbb{Q}^1$. Applying the Markov property at time $t - t^{1/3}$, it follows that $t^{-1/3}(N_t - N_t^*)$ converges in law to $\langle 1, \tilde{\varphi}\rangle Z$, where $Z \sim \mathtt{Gamma}(2, 2/\Sigma)$, thanks to the discussion following (22). Thus, for $\varepsilon > 0$, we have that $\mathbb{Q}^1_{\delta_x}(N_t - N_t^* \ge t^{1-\varepsilon}) \to 0$ as $t \to \infty$. We also have that $\mathbb{Q}^1_{\delta_x}(N_t \ge t^{1-\varepsilon/2}) \to 0$, uniformly in x, thanks to the Markov inequality and Assumption 2.2.

Now note that, on the one hand, $N_t^*/N_t \le 1$. On the other hand, $1 - (N_t^*/N_t) = (N_t - N_t^*)/N_t = t^{-1}(N_t - N_t^*)/(t^{-1}N_t)$. In the final equality, the numerator tends to zero in $\mathbb{Q}^1$-probability, and the denominator converges under $\mathbb{Q}^1$ to a Gamma distributed random variable. It follows that $N_t^*/N_t \to 1$ under $\mathbb{Q}^1_{\delta_x}$ as $t \to \infty$. Next note that the part of the spatial branching tree consisting of all descendants branching off the spine particle *before* time $t - t^{1/3}$ is conditionally independent (given the position of the spine at time $t - t^{1/3}$) of descendants branching off the spine particle *after* time $t - t^{1/3}$. This implies that any subsequential distributional limit of $(N_t/t, \xi_t)$ as $t \to \infty$ under $\mathbb{Q}^1_{\delta_x}$, say $(\langle 1, \tilde{\varphi}\rangle Z, \bar{\xi})$, can be extended to a subsequential limit $(Y^*, \langle 1, \tilde{\varphi}\rangle Z, \bar{\xi})$ of $(N_t^*/t, N_t/t, \bar{\xi})$ satisfying $Y^* = \langle 1, \tilde{\varphi}\rangle Z$ almost surely. On the other hand, ergodicity of the spine motion [11, section 5] implies that Y^* and $\bar{\xi}$ are independent. That is to say, Z and $\bar{\xi}$ are independent. □

Remark 10. Observe that a slight variant of the proof of (c) given above, implies that for any $c \in (0,1]$, the joint law of $(X_t[g_t]/t, N_t/t, \xi_{ct}, \xi_t)$ under $\mathbb{Q}^1_{\delta_x}$ also converges to that of $(\langle g, \tilde{\varphi}\rangle Z, \langle 1, \tilde{\varphi}\rangle Z, \bar{\xi}, \bar{\xi}')$ as $t \to \infty$, where $(\bar{\xi}, \bar{\xi}', Z)$ are mutually independent and $\bar{\xi}, \bar{\xi}'$ both have law (20).

We are now ready to prove Proposition 8. However, let us first give a sketch proof, which (roughly) describes the main steps of the argument (although some details look slightly different in the final version, in order to deal with technicalities that arise).

Sketch proof of Proposition 8. For $v, w \in \Omega$, let $\tau_{v,w}$ be the split time of v and w. That is, the death time of the most recent common ancestor of v and w.

- We first show that the asymptotic probability of selecting $w \in \mathcal{N}_{at}$ and $v \in \mathcal{N}_t$ with $w \preceq v$ is zero. Roughly speaking, this is because the probability that any specific individual at time at has *any* descendants at time t, tends to 0 as $t \to \infty$. This means that to obtain the asymptotic law of the split time, it is enough to provide an asymptotic for

$$\mathbb{P}_{\delta_x}\left[\sum_{w \in \mathcal{N}_{at}, v \in \mathcal{N}_t, w \not\preceq v} \frac{1}{N_{at}\hat{N}_t^w} F\Big(\frac{\tau_{v,w}}{t}\Big) \,\Big|\, N_t > 0\right]$$

  when F is an arbitrary bounded continuous function, and where for $w \in \mathcal{N}_{at}$, $\hat{N}_t^w$ is the size of the population at time t *without* counting the descendants of w. *(This roughly corresponds to Step 1 in the full proof below).*

- The many-to-two lemma at times at, t allows the above expectation to be written as

$$\frac{\varphi(x)}{\mathbb{P}_{\delta_x}(N_t > 0)}\mathbb{Q}^2_{\delta_x}\left[F\Big(\frac{\tau}{t}\Big)\mathbf{1}_{\{\tau \le at\}}\frac{\varphi(\xi^1_{\tau-})}{\varphi(\xi^1_t)\varphi(\xi^2_{at})}\frac{1}{N_{at}\hat{N}_t}\mathrm{e}^{\int_0^\tau \beta(\xi^1_s)(\mathtt{m}_2(\xi^1_s) - \mathtt{m}_1(\xi^1_s))\,\mathrm{d}s}\right]$$

  where τ is the split time of the two spines under $\mathbb{Q}^2_{\delta_x}$ and $\hat{N}_t$ is the population size at time t, not counting descendants of the second spine at time at. *(This roughly corresponds to Step 3 in the full proof below).*

- Next we consider $\hat{\mathbb{Q}}^2_{\delta_x}$ obtained by reweighting $\mathbb{Q}^2_{\delta_x}$ by

$$(\beta(\xi^1_{\tau-})(\mathtt{m}_2(\xi^1_{\tau-}) - \mathtt{m}_1(\xi^1_{\tau-})))^{-1}\mathrm{e}^{\int_0^\tau \beta(\xi^1_s)(\mathtt{m}_2(\xi^1_s) - \mathtt{m}_1(\xi^1_s))\mathrm{d}s - \tau}.$$

  This change of measure alters the rate at which the spine particles split into two distinct spines (from rate $\beta(\mathtt{m}_2 - \mathtt{m}_1)$ to rate 1) but doesn't affect the rate at which branching events occur that don't result in the spines splitting. Combining this with a change of variables and conditioning on τ (which has an exponential 1 distribution under $\hat{\mathbb{Q}}^2_{\delta_x}$), we rewrite our expectation again, as

$$\frac{\varphi(x)}{t\mathbb{P}_{\delta_x}(N_t > 0)}\int_0^a \mathrm{d}u F(u)\hat{\mathbb{Q}}^2_{\delta_x}\left[\frac{\beta(\xi^1_{ut})\varphi(\mathtt{m}_2(\xi^1_{ut}) - \mathtt{m}_1(\xi^1_{ut}))}{\varphi(\xi^1_t)\varphi(\xi^2_{at})}\frac{t^2}{N_{at}\hat{N}_t}\,\Big|\,\tau = ut\right].$$

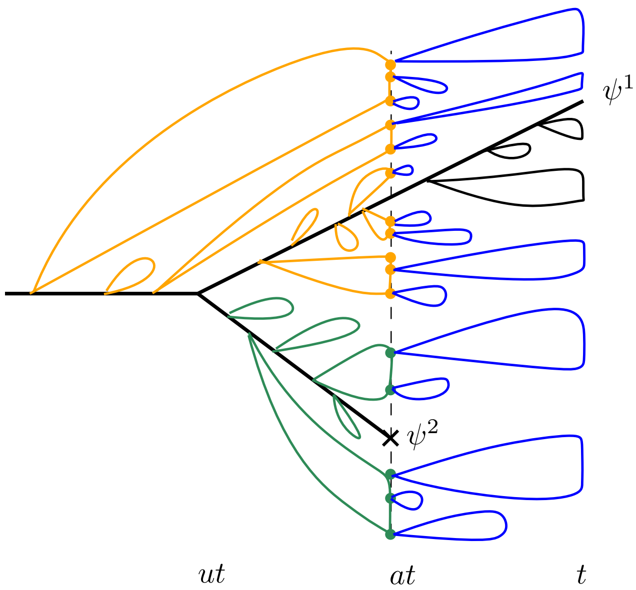

Figure 3: Suppose the two spines split from each other at time $\tau = ut$. The population at time at can be broken up into those individuals that have branched off the first spine before time at (depicted in orange) and those individuals that have branched off the second spine between times ut and at (depicted in green). Given the population at time at, the size of the population at time t (without the descendants of the second spine, that is, $\hat{N}_t$) can again be broken up into two subpopulations: those that branch off the first spine between times at and t (depicted in black) and those that are descendants of non-spine particles at time at (depicted in blue).

Here the law $\hat{\mathbb{Q}}^2_{\delta_x}[\cdot|\tau = ut]$ makes rigorous sense: the system has a single spine and in fact evolves as under $\mathbb{Q}^1_{\delta_x}$ until time ut, where some (biased) branching event occurs, two spines are selected, and each of these initiates an independent $\mathbb{Q}^1$ process. *(This roughly corresponds to Step 4 in the full proof below).*

- Now, we know by Lemma 10 that $\varphi(x)(t\mathbb{P}_{\delta_x}(N_t > 0))^{-1} \to \Sigma/2$ as $t \to \infty$. Moreover, under $\hat{\mathbb{Q}}^2_{\delta_x}$, similar arguments to those in the proof of Lemma 10 (in particular, due to ergodicity of the spine motion) imply that the positions $\xi^1_{ut}, \xi^1_t, \xi^2_{at}$ of the spines are asymptotically independent of each other and of $N_{at}, \hat{N}_t$ as $t \to \infty$, with limiting laws described by $P(\xi \in A) = \langle \mathbf{1}_A\varphi, \tilde{\varphi}\rangle$ for $A \subset E$.

  Furthermore, the limiting law of N_{at}/t as $t \to \infty$ is described by $aZ + (a-u)Z'$, where (Z, Z') are a pair of independent $\texttt{Gamma}(2,\, 2/\Sigma\langle 1, \tilde{\varphi}\rangle)$ random variables; this is because of the explicit description of the process under $\hat{\mathbb{Q}}^2_{\delta_x}(\cdot|\tau = ut)$ and item (c) of Lemma 10. In Figure 3, aZ and $(a-u)Z'$ correspond to the sizes of the orange and green populations respectively (after rescaling by t).

  Finally, the conditional limiting law of $\hat{N}_t/t$ given N_{at}/t is that of a Gamma random variable with parameter $(2+K,\, 2/\Sigma(1-a)\langle 1, \tilde{\varphi}\rangle)$, where $K \sim \texttt{Poisson}(2N/(1-a)\Sigma\langle 1, \tilde{\varphi}\rangle N)$ is itself random. This is because, given the collection of particles alive at time at, the first spine particle will have a number of offspring at time t which is asymptotically like

t times a $\texttt{Gamma}(2, 2/\Sigma(1-a)\langle 1, \tilde{\varphi}\rangle)$ random variable (Lemma 10(c) again; this corresponds to the population depicted in black in Figure 3). Then, independently, each of the non-spine particles alive will have some descendant alive at time t with probability asymptotically proportional to t^{-1} times φ of their positions. Using (essentially) the Poisson approximation of the binomial distribution, this results in a total number of non-spine particles with some descendant alive at time t having asymptotic conditional distribution given by a $\texttt{Poisson}(2N/(1-a)\Sigma\langle 1, \tilde{\varphi}\rangle N)$ random variable. By Lemma 10 (b), the number of offspring of each of these will approximately t times an independent $\texttt{Exponential}(2/\Sigma(1-a)\langle 1, \tilde{\varphi}\rangle)$, that is, a $\texttt{Gamma}(1, 2/\Sigma(1-a)\langle 1, \tilde{\varphi}\rangle)$ random variable. This corresponds to the population depicted in blue in Figure 3. The additivity property of independent Gamma distributions completes the argument. *(This roughly corresponds to Step 5 in the full proof below).*

- Plugging these asymptotics into the $\hat{\mathbb{Q}}^2_{\delta_x}$ expectation, and performing some simple explicit computations, we obtain the desired formula. *(This roughly corresponds to Steps 6 and 7 in the full proof below).*

□

Proof of Proposition 8. Fix $0 < a < 1$ and for $w \in \mathcal{N}_{at}$ and $v \in \mathcal{N}_t$, write $\tau_{v,w}$ for the split time of v and w (as in the sketch proof). It suffices to show that, for each continuous $F : [0, \infty) \to [0, 1]$,

$$\mathbb{P}_{\delta_x}\left[\frac{1}{N_{at}N_t}\sum_{v\in\mathcal{N}_t, w\in\mathcal{N}_{at}} F(\tau_{v,w}/t)\,\middle|\, N_t > 0\right] \to \int_0^a F(u) f_a(u)\mathrm{d}u \tag{23}$$

as $t \to \infty$.

Step 1 In order to apply the many-to-two lemma, we first need to write the left-hand side of (23) in a slightly different form (that is asymptotically equivalent). We claim that

$$\mathbb{P}_{\delta_x}\left[\frac{1}{N_{at}N_t}\sum_{v\in\mathcal{N}_t, w\in\mathcal{N}_{at}} F(\tau_{v,w}/t)\,\middle|\, N_t > 0\right] \sim \mathbb{P}_{\delta_x}\left[\sum_{\substack{v\in\mathcal{N}_t, w\in\mathcal{N}_{at}\\ w\not\preceq v}} \frac{1}{N_{at}\hat{N}^w_t} F(\tau_{v,w}/t)\,\middle|\, N_t > 0\right] \tag{24}$$

as $t \to \infty$ where $\hat{N}^w_t = N_t - N^w_t$ and $N^w_t = |\{v \in \mathcal{N}_t : w \preceq v\}|$.

To prove this claim, let us consider the difference of the two quantities in the above asymptotic. Recalling that $F \in [0,1]$, we have

$$\Big|\mathbb{P}_{\delta_x}\Big[\frac{1}{N_{at}N_t}\sum_{v\in\mathcal{N}_t, w\in\mathcal{N}_{at}} F(\tau_{v,w}/t) \,\big|\, N_t>0\Big] - \mathbb{P}_{\delta_x}\Big[\sum_{\substack{v\in\mathcal{N}_t, w\in\mathcal{N}_{at}\\ w\not\preceq v}} \frac{1}{N_{at}\hat{N}_t^w} F(\tau_{v,w}/t) \,\big|\, N_t>0\Big]\Big|$$

$$\le \Big|\mathbb{P}_{\delta_x}\Big[\sum_{\substack{v\in\mathcal{N}_t, w\in\mathcal{N}_{at}\\ w\preceq v}} \frac{F(\tau_{v,w}/t)}{N_tN_{at}}\big|N_t>0\Big]\Big| + \Big|\mathbb{P}_{\delta_x}\Big[\sum_{\substack{v\in\mathcal{N}_t, w\in\mathcal{N}_{at}\\ w\not\preceq v}} \frac{F(\tau_{v,w}/t)}{N_{at}}\big(\frac{1}{\hat{N}_t^w}-\frac{1}{N_t}\big)\big|N_t>0\Big]\Big|$$

$$\le \mathbb{P}_{\delta_x}\Big[\sum_{v\in\mathcal{N}_t}\frac{1}{N_tN_{at}}\big|N_t>0\Big] + \mathbb{P}_{\delta_x}\Big[\frac{1}{N_{at}}\sum_{w\in\mathcal{N}_{at}}\frac{N_t^w}{N_t}\big|N_t>0\Big]$$

$$= 2\mathbb{P}_{\delta_x}\Big[\frac{1}{N_{at}}\Big|N_t>0\Big],$$

where we have used the simple decomposition $\sum_{w\in\mathcal{N}_{at}} N_t^w = N_t$ in the final equality. Note that, for arbitrarily large $M>0$,

$$\mathbb{P}_{\delta_x}\Big[\frac{1}{N_{at}}\Big|N_t>0\Big] \le \frac{1}{M}\mathbb{P}_{\delta_x}\Big[N_{at}\ge M\Big|N_t>0\Big] + \mathbb{P}_{\delta_x}\Big[N_{at}\le M\Big|N_t>0\Big]. \tag{25}$$

The first term on the right-hand side of (25) can be made arbitrarily small by taking M sufficiently large. Moreover, the second term on the right-hand side of (25) converges to zero by Lemma 10.

It therefore suffices to prove that

$$\mathbb{P}_{\delta_x}\Big[\sum_{\substack{v\in\mathcal{N}_t, w\in\mathcal{N}_{at}\\ w\not\preceq v}} \frac{1}{N_{at}\hat{N}_t^w}F(\tau_{v,w}/t)\,\Big|\,N_t>0\Big] \to \int_0^a F(u)f_a(u)\mathrm{d}u, \quad \text{as } t\to\infty \tag{26}$$

This will be the new goal for the remainder of the proof.

Step 2 In order to apply some bounded convergence results, it is convenient to define the following event for $\delta>0$. Namely, we write $A^\delta_{v,w}$ for the event that

$$\varphi(X_v(t))\ge\delta\ ,\ \varphi(X_w(at))\ge\delta\ ,\ \varphi(X_v(\tau^-_{v,w})) = \varphi(X_w(\tau^-_{v,w}))\ge\delta\ ,\ \frac{N_{at}}{t}\ge\delta \text{ and } \frac{\hat{N}_t^w}{t}\ge\delta.$$

We claim that it suffices to show that for each $\delta>0$ and continuous $F:[0,\infty)\to[0,1]$,

$$\mathbb{P}_{\delta_x}\Big[\sum_{\substack{v\in\mathcal{N}_t, w\in\mathcal{N}_{at}\\ w\not\preceq v}} \frac{\mathbf{1}_{A^\delta_{v,w}}}{N_{at}\hat{N}_t^w}F(\tau_{v,w}/t)\,\Big|\,N_t>0\Big] \to \frac{c_\delta}{\langle 1,\tilde\varphi\rangle^2\Sigma}\int_0^a F(u)f_a^\delta(u)\mathrm{d}u \tag{27}$$

as $t\to\infty$, where

$$c_\delta := \langle\mathbf{1}_{\{\varphi\ge\delta\}},\tilde\varphi\rangle^2\langle\beta\varphi^2(\mathtt{m}_2-\mathtt{m}_1)\mathbf{1}_{\{\varphi\ge\delta\}},\tilde\varphi\rangle \tag{28}$$

and for some $f_a^\delta(u)\nearrow f_a(u)$ as $\delta\searrow 0$, pointwise on $[0,a]$. To see why (27) suffices, note that using the definitions of $\mathbb{V}$ and $\mathtt{m}_k$,

$$\langle\beta\varphi^2(\mathtt{m}_2-\mathtt{m}_1)\mathbf{1}_{\{\varphi\ge\delta\}},\tilde\varphi\rangle = \langle\beta\mathbb{V}[\varphi]\mathbf{1}_{\{\varphi\ge\delta\}},\tilde\varphi\rangle.$$

Thus, using the boundedness of $\tilde{\varphi}$, φ, β and Assumption 2.2, it follows that $c_\delta \uparrow \langle 1, \tilde{\varphi}\rangle^2\Sigma$ as $\delta \downarrow 0$. Moreover, since we know by Remark 8 that $f_a(u)$ integrates to 1 over $[0, a]$, we can take $F \equiv 1$ in (27) to see that

$$\lim_{\delta\to 0}\lim_{t\to\infty} \mathbb{P}_{\delta_x}\left[\sum_{\substack{v\in\mathcal{N}_t, w\in\mathcal{N}_{at}\\ w\not\leq v}} \frac{\mathbf{1}_{(A^\delta_{v,w})^c}}{N_{at}\hat{N}^w_t}\,\middle|\, N_t>0\right] = \lim_{\delta\to 0}\lim_{t\to\infty}\left(1-\mathbb{P}_{\delta_x}\left[\sum_{\substack{v\in\mathcal{N}_t, w\in\mathcal{N}_{at}\\ w\not\leq v}} \frac{\mathbf{1}_{A^\delta_{v,w}}}{N_{at}\hat{N}^w_t}\,\middle|\, N_t>0\right]\right) = 0.$$

Thus, given (27), we can take $t\to\infty$ and then $\delta\downarrow 0$ to deduce that the right-hand side of (24) converges to the right-hand side of (23) as $t\to\infty$.

The remaining steps will focus on the proof of (27) for an appropriate choice of f^δ_a and with c_δ defined in (28).

Step 3 We may apply the many-to-two formula, Lemma 7, with $s_1 = t, s_2 = at$, to write

$$\begin{aligned}
&\mathbb{P}_{\delta_x}\left[\sum_{v\in\mathcal{N}_t, w\in\mathcal{N}_{at}, w\not\leq v} \frac{\mathbf{1}_{A^\delta_{v,w}}}{N_{at}\hat{N}^w_t} F(\tau_{v,w}/t)\,\middle|\, N_t>0\right]\\
&=\frac{\varphi(x)}{\mathbb{P}_{\delta_x}(N_t>0)}\mathbb{Q}^2_{\delta_x}\left[F(\tau/t)\mathbf{1}_{\{\tau\leq at\}}\frac{\varphi(\xi^1_{\tau-})}{\varphi(\xi^1_t)\varphi(\xi^2_{at})}\frac{\mathbf{1}_{A^\delta_t}}{N_{at}\hat{N}_t}\mathrm{e}^{\int_0^\tau \beta(\xi^1_s)(\mathtt{m}_2(\xi^1_s)-\mathtt{m}_1(\xi^1_s))\,\mathrm{d}s}\right] \qquad (29)
\end{aligned}$$

where $\tau = \tau_{\psi^1_t,\psi^2_{at}}$, $\hat{N}_t = \hat{N}^{\psi^2_{at}}_t$, and $A^\delta_t = A^\delta_{\psi^1_t,\psi^2_{at}}$. The application of Lemma 6 is justified since

$$Y = F\big(\tau/t\big)\mathbf{1}_{\{\tau\leq at\}}\frac{\mathbf{1}_{A^\delta_t}}{N_{at}\hat{N}_t}$$

is $\mathcal{F}^2_{(at,t)}$- measurable.

Step 4 Next we consider $\hat{\mathbb{Q}}^2_{\delta_x}$ obtained by reweighting $\mathbb{Q}^2_{\delta_x}$ by

$$\frac{1}{\beta(\xi^1_{\tau-})(\mathtt{m}_2(\xi^1_{\tau-})-\mathtt{m}_1(\xi^1_{\tau-}))}\mathrm{e}^{\int_0^\tau \beta(\xi^1_s)(\mathtt{m}_2(\xi^1_s)-\mathtt{m}_1(\xi^1_s))\mathrm{d}s-\tau}.$$

This change of measure alters the rate at which the spine particles split into two distinct spines (from rate $\beta(\mathtt{m}_2-\mathtt{m}_1)$ to rate 1). Note, however, that it doesn't affect the rate at which branching events occur that don't result in the spines splitting. Combining this with a change of variables and conditioning on τ, it follows that the right-hand side of (29) is equal to

$$\frac{\varphi(x)}{t\mathbb{P}_{\delta_x}(N_t>0)}\int_0^a \mathrm{d}u F(u)\hat{\mathbb{Q}}^2_{\delta_x}\left[\frac{\beta(\xi^1_{ut})\varphi(\xi^1_{ut})(\mathtt{m}_2(\xi^1_{ut})-\mathtt{m}_1(\xi^1_{ut}))\mathbf{1}_{A^\delta_t}}{\varphi(\xi^1_t)\varphi(\xi^2_{at})}\frac{t^2}{N_{at}\hat{N}_t}\,\middle|\,\tau=ut\right]. \qquad (30)$$

Note that under $\hat{\mathbb{Q}}^2_{\delta_x}(\cdot|\tau=ut)$, the process behaves as follows.

- Until time ut, the process moves with biased motion as in (5) with ζ given by (16). At rate $\beta\mathtt{m}_1$ branching events occur, at which point, the offspring distribution is given by $\mathcal{P}^{2,1}$ (as in (6)) and the i-th particle is chosen to be the spine with probability proportional to $\varphi(x_i)$.

- At time ut a branching event occurs, where the law of the offspring is given by $\mathcal{P}^{2,2}$ and particles i, j with $i \neq j$ are chosen as the two (distinct) spines with probability proportional to $\varphi(x_i)\varphi(x_j)$.
- After time ut, the processes issued from the two spine particles evolve under $\mathbb{Q}^1$ and those issued from the non-spine particles evolve under $\mathbb{P}$.

Note also that by Lemma 10, $\varphi(x)/(t\mathbb{P}_{\delta_x}(N_t > 0)) \to \Sigma/2$ as $t \to \infty$. Thus, if we write $\hat{\mathbb{Q}}^2_{\delta_x,ut}(\cdot) := \hat{\mathbb{Q}}^2_{\delta_x}(\cdot \,|\, \tau = ut)$, we have that

$$\mathbb{P}_{\delta_x}\left[\sum_{\substack{v\in\mathcal{N}_t,\, w\in\mathcal{N}_{at}\\ w\not\preceq v}} \frac{\mathbf{1}_{A^\delta_{v,w}}}{N_{at}\hat{N}^w_t} F(\tau_{v,w}/t) \,\middle|\, N_t > 0\right]$$
$$\sim \frac{\Sigma}{2}\int_0^a \mathrm{d}u F(u) \hat{\mathbb{Q}}^2_{\delta_x,ut}\left[\frac{\beta(\xi^1_{ut})\varphi(\xi^1_{ut})(\mathtt{m}_2(\xi^1_{ut}) - \mathtt{m}_1(\xi^1_{ut}))}{\varphi(\xi^1_t)\varphi(\xi^2_{at})} \frac{t^2\mathbf{1}_{A^\delta_t}}{N_{at}\hat{N}_t}\right], \tag{31}$$

as $t \to \infty$.

Step 5 Our next goal is to describe the limit in law of $(\xi^1_{ut}, \xi^1_t, \xi^2_{at}, N_{at}/t, \hat{N}_t/t)$ under $\hat{\mathbb{Q}}^2_{x,ut}$ as $t \to \infty$. More precisely, we claim that it converges to $(\bar{\xi}, \bar{\xi}', \bar{\xi}'', N, \hat{N})$ where

- $\bar{\xi}, \bar{\xi}', \bar{\xi}''$ are independent of eachother and of $(N, \hat{N})$, each with law given by (20);
- the law of N is that of $aZ+(a-u)Z'$, where (Z, Z') are a pair of independent $\mathtt{Gamma}(2,\, 2/\Sigma\langle 1, \tilde{\varphi}\rangle)$ random variables;
- conditionally on N, the law of $\hat{N}$ is that of a $\mathtt{Gamma}(2 + K,\, 2/\Sigma(1 - a)\langle 1, \tilde{\varphi}\rangle)$ random variable with random $K \sim \mathtt{Poisson}(2N/(1 - a)\Sigma\langle 1, \tilde{\varphi}\rangle N)$.

To justify this claim, we identify the limiting Laplace transform of $(\xi^1_{ut}, \xi^1_t, \xi^2_{at}, N_{at}/t, \hat{N}_t/t)$. To this end, for arbitrary $\theta, \mu, \eta, \rho, \chi \geq 0$, let us consider

$$\hat{\mathbb{Q}}^2_{x,ut}\big[\mathrm{e}^{-\theta\xi^1_{ut}}\mathrm{e}^{-\mu\xi^1_t}\mathrm{e}^{-\eta\xi^2_{at}}\mathrm{e}^{-\rho N_{at}/t}\mathrm{e}^{-\chi\hat{N}_t/t}\big]$$
$$= \hat{\mathbb{Q}}^2_{x,ut}\big[\mathrm{e}^{-\theta\xi^1_{ut}-\eta\xi^2_{at}-\rho N_{at}/t}\,\hat{\mathbb{Q}}^2_{x,ut}\big[\mathrm{e}^{-\mu\xi^1_t}\mathrm{e}^{-\chi\hat{N}_t/t} \,|\, \mathcal{F}^2_{at}\big]\big], \tag{32}$$

where $\mathcal{F}^2_{at}$ is the σ-algebra containing all the information about the process, including the spines, up to time at.

Recalling the description of the process under $\hat{\mathbb{Q}}^2_{x,ut}$, we see that

$$\hat{\mathbb{Q}}^2_{\delta_x,ut}\big[\mathrm{e}^{-\mu\xi^1_t}\mathrm{e}^{-\chi\hat{N}_t/t} \,|\, \mathcal{F}^2_{at}\big] = \mathbb{Q}^1_{\delta_{\xi^1_{at}}}\big[\mathrm{e}^{-\mu\xi_{(1-a)t}}\mathrm{e}^{-\chi N_{(1-a)t}/t}\big] \prod_{\substack{v\in\mathcal{N}_{at}\\ v\neq\psi^1_{at},\psi^2_{at}}} \mathbb{P}_{\delta_{X_v(at)}}\big[\mathrm{e}^{-\chi N_{(1-a)t}/t}\big]$$
$$= \mathbb{Q}^1_{\delta_{\xi^1_{at}}}\big[\mathrm{e}^{-\mu\xi_{(1-a)t}}\mathrm{e}^{-\chi N_{(1-a)t}/t}\big] \exp\Bigg(\sum_{\substack{v\in\mathcal{N}_{at}\\ v\neq\psi^1_{at},\psi^2_{at}}} \log\big(1 - (1 - \mathbb{P}_{\delta_{X_v(at)}}[\mathrm{e}^{-\chi N_{(1-a)t}/t}])\big)\Bigg)$$
$$= \mathbb{Q}^1_{\delta_{\xi^1_{at}}}\big[\mathrm{e}^{-\mu\xi_{(1-a)t}}\mathrm{e}^{-\chi N_{(1-a)t}/t}\big](1 - E_t) \exp\Bigg(\sum_{\substack{v\in\mathcal{N}_{at}\\ v\neq\psi^1_{at},\psi^2_{at}}} -(1 - \mathbb{P}_{\delta_{X_v(at)}}[\mathrm{e}^{-\chi N_{(1-a)t}/t}])\Bigg) \tag{33}$$

where

$$E_t := 1 - \exp\Bigg(\sum_{\substack{v\in\mathcal{N}_{at}\\ v\neq\psi^1_{at},\psi^2_{at}}} \log\big(1-(1-\mathbb{P}_{\delta_{X_v(at)}}[\mathrm{e}^{-\chi N_{(1-a)t}/t}])\big) + (1-\mathbb{P}_{\delta_{X_v(at)}}[\mathrm{e}^{-\chi N_{(1-a)t}/t}])\Bigg).$$

We claim that E_t belongs to $[0,(cN_{at}/t^2)\wedge 1]$ for some absolute deterministic constant c, by Lemma 10(a). The lower bound of zero and the upper bound of 1 are obvious. To see where the upper bound of cN_{at}/t^2comes from, note that there are at most N_{at} elements of the sum defining E_t, $\log(1-x)+x \le -x^2/2$ and $1-\mathbb{P}_{\delta_{X_v(at)}}[\mathrm{e}^{-\chi N_{(1-a)t}/t}] \le \chi \sup_{x\in E}\mathbb{P}_{\delta_x}[N_{(1-a)t}]/t \le K/t$, for some appropriate $K>0$ (where this final inequality follows from criticality and Assumption 2).

Note also that by Lemma 10(c),

$$\mathbb{Q}^1_{\delta_{\xi^1_{at}}}\big[\mathrm{e}^{-\mu\xi_{(1-a)t}}\mathrm{e}^{-\chi N_{(1-a)t}/t}\big] = s^2\langle\varphi \mathrm{e}^{-\mu\cdot},\tilde\varphi\rangle(1+e_t(\xi^1_{at}))$$

where $s := (1+\chi\Sigma(1-a)\langle 1,\tilde\varphi\rangle/2)^{-1} < 1$ and $e_t(x)$ is such that $e_t(x)\to 0$ in the pointwise sense on E and

$$\sup_{t,x}|e_t(x)| < \infty.$$

Let us further denote for $x\in E, t\ge 0$,

$$g_t(x) := t\big(1-\mathbb{P}_{\delta_x}[\mathrm{e}^{-\chi N_{(1-a)t}/t}]\big) > 0$$

so that by Lemma 10, $\sup_{t,x} g_t(x)<\infty$ and

$$g_t(x) = t\mathbb{P}_{\delta_x}(N_{(1-a)t}>0)\big(1-\mathbb{P}_{\delta_x}(\mathrm{e}^{-\chi N_{(1-a)t}/t}\,|\,N_{(1-a)t}>0)\big) \to \frac{2(1-s)}{\Sigma(1-a)}\varphi(x)$$

pointwise on E as $t\to\infty$. Using this notation in the right hand side of (33), we see that

$$\hat{\mathbb{Q}}^2_{\delta_x,ut}\big[\mathrm{e}^{-\mu\xi^1_t}\mathrm{e}^{-\chi\hat N_t/t}\,|\,\mathcal{F}^2_{at}\big] = s^2\langle\varphi\mathrm{e}^{-\mu\cdot},\tilde\varphi\rangle \mathrm{e}^{-\frac{X_{at}[g_t]}{t}}(1+e_t(\xi^1_{at}))(1-E_t)$$

so that

$$\begin{aligned}\hat{\mathbb{Q}}^2_{x,ut}&\big[\mathrm{e}^{-\theta\xi^1_{ut-}}\mathrm{e}^{-\mu\xi^1_t}\mathrm{e}^{-\eta\xi^2_{at}}\mathrm{e}^{-\rho N_{at}/t}\mathrm{e}^{-\chi\hat N_t/t}\big]\\ &= s^2\langle\varphi\mathrm{e}^{-\mu\cdot},\tilde\varphi\rangle\hat{\mathbb{Q}}^2_{x,ut}\big[\mathrm{e}^{-\theta\xi^1_{ut-}-\eta\xi^2_{at}-\rho N_{at}/t}\mathrm{e}^{-X_{at}[g_t]/t}(1+e_t(\xi^1_{at}))(1-E_t)\big]. \quad (34)\end{aligned}$$

Now, we claim that under $\hat{\mathbb{Q}}^2_{x,ut}$,

$$(\xi^1_{ut},\xi^2_{at},e_t(\xi^1_{at}),N_{at}/t,X_t[g_t]/t,E_t) \Rightarrow (\bar\xi,\bar\xi'',0,N,\tfrac{2(1-s)}{\Sigma(1-a)\langle 1,\tilde\varphi\rangle}N,0) \quad (35)$$

as $t\to\infty$, where $(\bar\xi,\bar\xi'',N)$ have joint law as described in the bullet points at the start of Step 5. Since everything inside the expectation on right-hand side of (34) is deterministically bounded, (35) implies that

$$\begin{aligned}\hat{\mathbb{Q}}^2_{x,ut}&\big[\mathrm{e}^{-\theta\xi^1_{ut-}}\mathrm{e}^{-\mu\xi^1_t}\mathrm{e}^{-\eta\xi^2_{at}}\mathrm{e}^{-\rho N_{at}/t}\mathrm{e}^{-\chi\hat N_t/t}\big] \to\\ &\langle\varphi\mathrm{e}^{-\theta\cdot},\tilde\varphi\rangle\langle\varphi\mathrm{e}^{-\mu\cdot},\tilde\varphi\rangle\langle\varphi\mathrm{e}^{-\lambda\cdot},\tilde\varphi\rangle s^2\int_0^\infty p_{a,u}(x)\mathrm{e}^{-\rho x}\mathrm{e}^{\frac{2(s-1)}{(1-a)\Sigma\langle 1,\tilde\varphi\rangle}x}\mathrm{d}x,\end{aligned}$$

where $p_{a,u}$ is the density of N (with law as described in the second bullet point). It is easy to check using the explicit expressions for the Laplace transforms of Poisson and Gamma random variables, that the right-hand side above is exactly the joint Laplace transform of our desired limit $(\xi, \xi', \xi'', N, \hat{N})$. Thus, it only remains to justify (35).

To this end, let us write $\mathcal{N}^1_{at}$ for the collection of particles alive at time at that have branched off the first spine between times 0 and at (depicted in orange in Figure 3). We also write $\mathcal{N}^2_{at}$ for those particles alive at time at that have branched off the second spine between times ut and at (depicted in green in Figure 3). Then, using Lemma 10 (and its extension Remark 10), it follows that

$$(\xi^1_{ut}, e_t(\xi^1_{at}), \frac{1}{t}|\mathcal{N}^1_{at}|, \frac{1}{t}\sum_{v\in\mathcal{N}^1_{at}} g_t(X_v(at))) \Rightarrow (\xi, 0, N', \tfrac{2(1-s)}{\Sigma(1-a)\langle 1,\tilde{\varphi}\rangle}N') \tag{36}$$

say, where (ξ, N') are independent, ξ has law given in (20) and N' has the law of a times a $\texttt{Gamma}(2, 2/\Sigma\langle 1,\tilde{\varphi}\rangle)$ random variable (recall that $e_t(x)$ is deterministically uniformly bounded over t, x and converges pointwise to 0 on E). Now, given all the information in the quadruple displayed on the left-hand side of (36), the (conditional) joint law of

$$(\xi^2_{at}, \frac{1}{t}|\mathcal{N}^2_{at}|, \frac{1}{t}\sum_{v\in\mathcal{N}^2_{at}} g_t(X_v(at)))$$

is given by the $\mathbb{Q}^1_{\delta_X}$ law of $(\xi^2_{(a-u)t}, N_{(a-u)t}/t, X_{(a-u)t}[g_t]/t)$, where the (conditional) law of X is explicit but not required here. Again, by Lemma 10, in particular part (c),

$$(\xi^2_{(a-u)t}, N_{(a-u)t}/t, X_{(a-u)t}[g_t]/t) \Rightarrow (\hat{\xi}, \hat{N}, \tfrac{2(1-s)}{\Sigma(1-a)\langle 1,\tilde{\varphi}\rangle}\hat{N})$$

as $t\to\infty$ under $\mathbb{Q}^1_{\delta_X}$, where $\hat{\xi}, \hat{N}$ are independent, $\hat{\xi}$ has law given by (20), and $\hat{N}$ has the law of $(a-u)$ times a $\texttt{Gamma}(2, 2/\Sigma\langle 1,\tilde{\varphi}\rangle))$ random variable, independently of the value of X. Putting these observations together, plus the fact that $E_t \in [0, cN_{at}/t^2]$ for some deterministic c, gives us (35).

Step 6 Using the convergence in law (and associated notation for limiting variables) from Step 5, plus boundedness of the functionals in question, we deduce that for each $0 \le u \le a$,

$$\begin{aligned}&\frac{\Sigma}{2}\int_0^a \mathrm{d}u F(u)\hat{\mathbb{Q}}^2_{\delta_x,ut}\left[\frac{\beta(\xi^1_{ut})\varphi(\xi^1_{ut})(\mathtt{m}_2(\xi^1_{ut}) - \mathtt{m}_1(\xi^1_{ut}))}{\varphi(\xi^1_t)\varphi(\xi^2_{at})}\frac{t^2\mathbf{1}_{A^\delta_t}}{N_{at}\hat{N}_t}\right]\\ &\to\frac{\Sigma}{2}\int_0^a \mathrm{d}u F(u)\mathbb{E}\left[\frac{\beta(\bar{\xi})\varphi(\bar{\xi})(\mathtt{m}_2(\bar{\xi}) - \mathtt{m}_1(\bar{\xi}))}{\varphi(\bar{\xi}')\varphi(\bar{\xi}'')}\mathbf{1}_{\{\varphi(\bar{\xi})\ge\delta,\varphi(\bar{\xi}')\ge\delta,\varphi(\bar{\xi}'')\ge\delta\}}\frac{\mathbf{1}_{\{N\ge\delta,\hat{N}\ge\delta\}}}{N\hat{N}}\right],\end{aligned} \tag{37}$$

as $t \to \infty$. The expectation in the integrand on the right-hand side of (37) is equal to

$$
\begin{aligned}
& c_\delta \mathbb{E}\left[\frac{\mathbf{1}_{\{N\geq\delta\}}}{N}\mathbb{E}\left[\frac{\mathbf{1}_{\{\hat{N}\geq\delta\}}}{\hat{N}}\,\middle|\, N\right]\right] \\
&= c_\delta \mathbb{E}\left[\frac{\mathbf{1}_{\{N\geq\delta\}}}{N}\left(\mathbb{E}\left[\frac{1}{\hat{N}}\,\middle|\, N\right] - \mathbb{E}\left[\frac{\mathbf{1}_{\{\hat{N}<\delta\}}}{\hat{N}}\,\middle|\, N\right]\right)\right] \\
&= c_\delta \mathbb{E}\left[\frac{\mathbf{1}_{\{N\geq\delta\}}}{N}\frac{1-\exp(-\frac{1}{1-a}(\frac{2}{\Sigma\langle 1,\tilde{\varphi}\rangle}N))}{N}\right] - c_\delta \mathbb{E}\left[\frac{\mathbf{1}_{\{N\geq\delta\}}}{N}\mathbb{E}\left[\frac{\mathbf{1}_{\{\hat{N}<\delta\}}}{\hat{N}}\,\middle|\, N\right]\right] \\
&= c_\delta \mathbb{E}\left[\frac{1-\exp(-\frac{1}{1-a}(\frac{2}{\Sigma\langle 1,\tilde{\varphi}\rangle}N))}{N^2}\right] - c_\delta \mathbb{E}\left[\frac{\mathbf{1}_{\{N<\delta\}}}{N}\frac{1-\exp(-\frac{1}{1-a}(\frac{2}{\Sigma\langle 1,\tilde{\varphi}\rangle}N))}{N}\right] \\
&\qquad\qquad - c_\delta \mathbb{E}\left[\frac{\mathbf{1}_{\{N\geq\delta\}}}{N}\mathbb{E}\left[\frac{\mathbf{1}_{\{\hat{N}<\delta\}}}{\hat{N}}\,\middle|\, N\right]\right] \\
&=: \frac{4c_\delta}{a^2\Sigma^2\langle 1,\tilde{\varphi}\rangle^2}\mathbb{E}\left[\frac{1-\exp(-\frac{a}{1-a}(\frac{2}{a\Sigma\langle 1,\tilde{\varphi}\rangle}N))}{(\frac{2}{a\Sigma\langle 1,\tilde{\varphi}\rangle}N)^2}\right] - h(\delta) \qquad (38)
\end{aligned}
$$

as $t \to \infty$, where under $\mathbb{E}$, $(\bar{\xi}, \bar{\xi}', \bar{\xi}'', N, \hat{N})$ are described in Step 5 (recalling in particular the law of $\hat{N}$ given N) and $h(\delta) \geq 0$ and $c_\delta := \langle \mathbf{1}_{\{\varphi\geq\delta\}}, \tilde{\varphi}\rangle^2 \langle \beta\varphi^2(\mathtt{m}_2 - \mathtt{m}_1)\mathbf{1}_{\{\varphi\geq\delta\}}, \tilde{\varphi}\rangle$.

Step 7 Recall that by (27) and (31), our aim is to prove that

$$
\frac{\Sigma}{2}\int_0^a \mathrm{d}u F(u)\hat{\mathbb{Q}}^2_{\delta_x,ut}\left[\frac{\beta(\xi^1_{ut})\varphi(\xi^1_{ut})(\mathtt{m}_2(\xi^1_{ut}) - \mathtt{m}_1(\xi^1_{ut}))}{\varphi(\xi^1_t)\varphi(\xi^2_{at})}\frac{t^2\mathbf{1}_{A^\delta_t}}{N_{at}\hat{N}_t}\right] \to \frac{c_\delta}{\langle 1,\tilde{\varphi}\rangle^2\Sigma}\int_0^a F(u)f^\delta_a(u)\mathrm{d}u
$$

as $t \to \infty$, for some $f^\delta_a(u) \nearrow f_a(u)$ as $\delta \searrow 0$, pointwise on $[0,a]$.

First notice that $h(\delta)$ in (38) converges to 0 as $\delta \searrow 0$, since the law of $N\hat{N}$ has negative moments of all orders. Then (38), (37) imply the result, since writing $Y = 2N/a\Sigma(1,\tilde{\varphi})$ (so that $Y \sim Y' + (1-\frac{u}{a})Y''$ for independent $Y', Y'' \sim \mathtt{Gamma}(2,1)$) we have:

$$
\begin{aligned}
& \frac{2}{a^2}\mathbb{E}\left[\frac{1-\exp(-\frac{a}{1-a}(\frac{2}{a\Sigma(1,\tilde{\varphi})}N))}{(\frac{2}{a\Sigma(1,\tilde{\varphi})}N)^2}\right] \\
&= \frac{2}{a^2}\int_0^\infty \mathbb{E}\left[\theta\exp(-\theta Y) - \theta\exp(-(\theta + \tfrac{a}{1-a})Y)\right] d\theta \\
&= \frac{2}{a^2}\int_0^\infty \left(\frac{\theta}{(1+\theta)^2(1+(1-\frac{u}{a})\theta)^2} - \frac{\theta}{(1+\theta+\frac{a}{1-a})^2(1+(1-\frac{u}{a})(\theta+\frac{a}{1-a}))^2}\right)\mathrm{d}\theta \\
&= f_a(u).
\end{aligned}
$$

To calculate the integral in the penultimate line, we have used the change of variables $x = \theta + \dfrac{a}{1-a}$ for the second integrand, the fact that the anti-derivative of $\dfrac{y}{(1+y)^2(1+\gamma y)^2}$ is given by

$$
\frac{1}{(\gamma-1)^3}\left(\frac{(\gamma-1)(\gamma y + y + 2)}{(y+1)(\gamma y+1)} - (\gamma+1)\log\left(\frac{y+1}{\gamma y+1}\right)\right),
$$

and that the anti-derivative of $\dfrac{1}{(1+y)^2(1+\gamma y)^2}$ is given by

$$\frac{1}{(\gamma-1)^3}\left(\frac{-(\gamma-1)(2\gamma y+\gamma+1)}{(y+1)(\gamma y+1)}+2\gamma\log\big(\frac{y+1}{\gamma y+1}\big)\right).$$

The proof is now complete. □

Proof of Proposition 9. The proof of this proposition is contained in the proof of Step 5 above, ignoring the contribution from the second spine. □

Acknowledgements This work was partially supported by by EPSRC grant EP/W026899/1, UKRI Future Leader's Fellowship MR/W008513/1, and the "PHC Alliance" programme (project number: 47867UJ), funded by the French Ministry for Europe and Foreign Affairs, the French Ministry for Higher Education and Research and the UK Department for Business, Energy and Industrial Strategy.

References

[1] S. Asmussen and H. Hering. *Branching processes*, volume 3 of *Progress in Probability and Statistics*. Birkhäuser Boston, Inc., Boston, MA, 1983.

[2] J. Bertoin. *Random fragmentation and coagulation processes*, volume 102 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.

[3] B. Chauvin and A. Rouault. KPP equation and supercritical branching Brownian motion in the subcritical speed area. application to spatial trees. *Probability theory and related fields*, 80(2):299–314, 1988.

[4] A. Cortines, L. Hartung, and O. Louidor. The structure of extreme level sets in branching Brownian motion. *Ann. Probab.*, 47(4):2257–2302, 2019.

[5] A. M. Cox, E. Horton, A. E. Kyprianou, and D. Villemonais. Stochastic methods for neutron transport equation III: Generational many-to-one and $k_{\texttt{eff}}$. *SIAM J. Appl. Math.*, 81(3):982–1001, 2021.

[6] J. Engländer. *Spatial branching in random environments and with interaction*, volume 20 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2015.

[7] F. Foutel-Rodier and E. Schertze. Convergence of genealogies through spinal decomposition with an application to population genetics.

[8] I. Gonzalez, E. Horton, and A. E. Kyprianou. Asymptotic moments of spatial branching processes. *Probability Theory and Relatied Fields*, to appear.

[9] S. C. Harris, M. Hesse, and A. E. Kyprianou. Branching Brownian motion in a strip: survival near criticality. *Ann. Probab.*, 44(1):235–275, 2016.

[10] S. C. Harris, E. Horton, and A. E. Kyprianou. Stochastic methods for the neutron transport equation II: almost sure growth. *Ann. Appl. Probab.*, 30(6):2815–2845, 2020.

[11] S. C. Harris, E. Horton, A. E. Kyprianou, and M. Wang. Yaglom limit for critical non-local branching Markov processes. *Annals of Probability*, 50(6):2373–2408, 2022.

[12] S. C. Harris, S. G. G. Johnston, and M. I. Roberts. The coalescent structure of continuous-time Galton-Watson trees. *Ann. Appl. Probab.*, 30(3):1368–1414, 2020.

[13] S. C. Harris, S. G. G. Johnston, and M. I. Roberts. The coalescent structure of continuous-time Galton-Watson trees. *Ann. Appl. Probab.*, 30(3):1368–1414, 2020.

[14] S. C. Harris and M. I. Roberts. The many-to-few lemma and multiple spines. *Ann. Inst. Henri Poincaré Probab. Stat.*, 53(1):226–242, 2017.

[15] E. Horton and A. E. Kyprianou. *Stochastic neutron transport and non-local branching Markov processes*. Probability and its Applications. Birkhäuser, 2023.

[16] E. Horton, A. E. Kyprianou, and D. Villemonais. Stochastic methods for the neutron transport equation I: linear semigroup asymptotics. *Ann. Appl. Probab.*, 30(6):2573–2612, 2020.

[17] E. Lubetzky, C. Thornett, and O. Zeitouni. Maximum of branching Brownian motion in a periodic environment. *Ann. Inst. Henri Poincaré Probab. Stat.*, 58(4):2065–2093, 2022.

[18] Z. Shi. *Branching random walks*, volume 2151 of *Lecture Notes in Mathematics*. Springer, Cham, 2015. Lecture notes from the 42nd Probability Summer School held in Saint Flour, 2012.