



**HAL**  
open science

# Spinal constructions for continuous type-space branching processes with interactions

Charles Medous

► **To cite this version:**

Charles Medous. Spinal constructions for continuous type-space branching processes with interactions. 2023. hal-04179190v2

**HAL Id: hal-04179190**

**<https://hal.science/hal-04179190v2>**

Preprint submitted on 6 Nov 2024

**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.

Public Domain

# SPINAL CONSTRUCTIONS FOR CONTINUOUS TYPE-SPACE BRANCHING PROCESSES WITH INTERACTIONS

CHARLES MEDOUS

ABSTRACT. We consider branching processes describing structured, interacting populations in continuous time. Dynamics of each individual's characteristics and branching properties can be influenced by the entire population. We propose a Girsanov-type result based on a spinal construction, and establish a many-to-one formula. By combining this result with the spinal decomposition, we derive a generalized continuous-time version of the Kesten-Stigum theorem that incorporates interactions. Additionally, we propose an alternative approach of the spine construction for exact simulations of stochastic size-dependent populations.

## INTRODUCTION

Spine techniques and spinal trees are classical tools in the general context of branching processes since the work of Kallenberg [36], Chauvin and Rouault [12, 13], and later Kurtz, Lyons, Pemantle and Peres [45, 44, 38]. Spinal trees are constructed based on the original branching process by distinguishing a lineage, called the spine. Its only living representative- the spinal individual- follows a biased reproduction law compared to the other individuals in the process, ensuring that the spine does not die out. In the specific yet widely studied case of size-biased trees, the reproductive law  $(\hat{p}_n, n \geq 0)$  of the spinal individual is defined by

$$\hat{p}_n = \frac{np_n}{m}, \quad \text{for } n \geq 0,$$

where  $m$  is the mean value of the law of reproduction  $(p_n, n \geq 0)$  in the branching process. This size-biased reproductive law of the spinal individual is closely related to the biased ancestral reproduction [13, 27]. In fact the spine was used to characterize the process of the trait of a uniformly sampled individual, in a large population approximation see *e.g.* [46]. The sampling of  $k \geq 2$  distinct individuals from those living in a population at a time  $t$  is associated with a  $k$ -spines construction. For further literature on multiple-particles sampling we refer the reader to [31, 30, 35, 14] and the references herein.

Many-to-One formulas [32, 27] are prominent among classical spine results. Such formulas are derived from a Girsanov-type result on the change of probability measure associated with the spine, which can be regarded as a Doob's  $h$  transform with  $h = 1$ , as described in [15, 2] for instance. However for non harmonic functions  $h$ , these formulas express expectations of sums over particles in the branching process in terms of a Feynman-Kac path integral expectation related to the spinal individual. Consequently, the spinal individual is often referred to as a "typical individual" within the population. The connection to Feynman-Kac path integrals implies a shared foundation between these concepts. For a comprehensive overview on this subject, we refer to [19].

Another interesting property of spinal constructions is the "spinal decomposition" [12, 45]. It is a representation of the spinal tree as a branching process with immigration, where the immigration events are governed by a biased reproductive law specific to the spine, while the distribution of offspring remains consistent with that of the original process. In recent decades, the spinal decomposition has emerged as a highly valuable tool for investigating branching processes. One of its notable contributions is providing a new proof of the  $L \log L$  criteria, which were originally proved by Kesten and Stigum for Galton-Watson (GW) processes [37] and by Biggins

for continuous-time branching processes [9] using analytical methods. These results give specific conditions on the reproductive law, ensuring the non-degeneracy of the martingale involved in the spinal change of measure at infinity. By combining the spinal decomposition with a previously known result from measure theory, Lyons, Pemantle, and Peres [45] provided a probabilistic "conceptual" proof of Kesten and Stigum's theorem for single-type GW processes. This method proved to be easily generalizable to continuous-time structured branching processes [2, 27, 8] as well as branching Brownian motions [39, 23, 28]. More recently Bertoin and Mallein extended this proof to general branching Levy processes [7]. For similar results on superprocesses we refer to [41, 42, 43, 22, 54]. Finally, we mention Hardy and Harris [29] who adapted the spinal decomposition to prove the  $\mathcal{L}^p$ -convergence of some key martingales, which was later used to establish strong laws of large numbers [24].

The fundamental assumption in the aforementioned works is the branching property, which assumes that the behavior of all particles in the process during their life is independent of one another. However, in various systems of population dynamics such as genetics, epidemiology, chemistry, and even queueing systems, interactions between individuals do occur and this fundamental hypothesis falls apart. Recently, Bansaye [3] established a spine construction and Many-to-One formulas for interacting branching populations, where the branching rates and reproductive laws depend on the traits of all individuals in the population. These traits belong to a finite set and are fixed during the life of the individuals. Using the spinal decomposition, Bansaye has found an  $L \log L$  criterion for a single-type, density-dependent population. We also mention a recent work on spine processes for density-dependent individual-based models in a large population approximation [34].

In this article we consider a wide class of continuous-time structured branching processes with general interactions. These processes are used to model structured populations, where the behavior of each individual is influenced by the overall population state. Every individual in the population is characterized by its trait, taking values in a subset of  $\mathbb{R}^d$ . The lifespan of each individual is exponentially distributed with a rate that depends on the traits of all individuals. Upon an individual's death, a random number of children are generated, each inheriting random traits at birth, that are influenced by the traits of all individuals in the population. Between these branching events, the evolution of the traits of all individuals in the population is deterministic and also influenced by the entire population's state, extending the work of Bansaye in [3] to include many physical models. Notice that the branching parameters are determined by the traits of all individuals, thereby the branching property no longer holds in this framework. We introduce a comprehensive  $\psi$ -spinal construction for those processes, using a change of measure associated with a positive weight function  $\psi$  that depends on the trait of the spinal individual and of those of all other individuals. For a fixed function  $\psi$  both the spinal individual and those outside the spine are subject to a bias. We derive a Girsanov-type formula associated with this change of measure, taking the form of a path-integral formulation that involves a non-linear operator. A classical approach to establishing limiting results, such as the central limit theorem or large deviations, involves determining the eigenfunctions of such operators [8, 23, 16]. However, due to the presence of interactions, this operator is contingent on the entire population, necessitating the eigenfunctions to be dependent on the traits of all individuals. Thus, the weight function  $\psi$  must rely on both the trait of the spinal individual and the traits of all individuals in order to account for this dependency. Under certain non-explosion assumptions regarding the branching parameters and the set of weight functions, we obtain a modified Many-to-One formula. Unlike the classical Many-to-One formula- that describes the behavior of the branching process using only the behavior of the spine- our formula relies on the whole spinal population.

Subsequently, we use this result in conjunction with the spinal decomposition to establish  $L \log L$  type necessary and  $L \log L$  type sufficient conditions for processes with interactions, generalizing the work of [3] for non constant functions  $\psi$ . More precisely, we exhibit both a sufficient

condition and a necessary condition for the non-degeneracy at infinity of the additive martingale associated with the  $\psi$ -spinal change of measure, for a large set of  $\psi$  functions.

Finally we propose an alternative use of the spinal construction, where the function  $\psi$  is not chosen as an eigenfunction but to simplify the dynamics of the associated spine process. As an example we study a particular case of structured Yule process with mass loss events happening at size-dependent rates. Yule processes are pure birth processes that are widely used in population genetics to model and reconstruct phylogenetic trees, see *e.g.* Aldous' review [1]. We use the spinal construction with a multiplicative weight function to retrieve a conditional branching property in the associated spine process, and propose an efficient algorithmic construction based on this property.

**Notation.** In the sequel  $\mathbb{N}^* = \{1, 2, \dots\}$  will denote the set of positive integers,  $\mathbb{R}_+ := [0, +\infty)$  the non-negative real line,  $\overline{\mathbb{R}}_+ := \mathbb{R}_+ \cup \{+\infty\}$  and  $\mathbb{R}_+^* := (0, +\infty)$ . We will denote respectively by  $\mathfrak{B}(A, B)$  (resp.  $\mathcal{C}^i(A, B)$ ) the set of measurable (resp.  $i$  times continuously differentiable)  $B$ -valued functions on a set  $A$ . For every couple  $(f, g)$  of real-valued measurable functions on a set  $A$ , we denote for all  $x$ ,  $fg(x)$  the product  $f(x)g(x)$ .

The set of trait  $\mathcal{X}$  is a subset of  $\mathbb{R}^d$  and for all  $(x, y) \in \mathcal{X}^2$ , we will denote  $x \cdot y$  and  $|x|$  respectively the canonical scalar product and the  $\ell_1$ -norm on  $\mathbb{R}^d$ .

We introduce the Ulam-Harris-Neveu notations [50] to label the individuals of the population:

$$\mathcal{U} := \{\emptyset\} \cup \bigcup_{k \geq 0} (\mathbb{N}^*)^{k+1}.$$

We consider branching processes starting from multiple initial individuals, thus the root  $\emptyset$  will be treated as a phantom individual and its direct descendants will be the ancestor generation. For two elements  $u, v$  of  $\mathcal{U} \setminus \{\emptyset\}$ , there exist two positive integers  $n, p$  such that  $u = (u_1, \dots, u_n)$  and  $v = (v_1, \dots, v_p)$  and we write  $uv := (u_1, \dots, u_n, v_1, \dots, v_p)$  the concatenation of  $u$  and  $v$ . We identify both  $\emptyset u$  and  $u\emptyset$  with  $u$ . An individual  $v \in \mathcal{U}$  is a descendant of  $u$  if there exists  $w \in \mathcal{U}$  such that  $v = uw$ . In this case we denote  $u \preceq v$ .

Let us introduce  $\overline{\mathbb{V}}$  the set composed of all finite point measures on  $\mathcal{U} \times \mathcal{X}$

$$\overline{\mathbb{V}} := \left\{ \sum_{i=1}^N \delta_{(u^i, x^i)}, N \in \mathbb{N}, (u^i, 1 \leq i \leq N) \in \mathcal{U}^N, (x^i, 1 \leq i \leq N) \in \mathcal{X}^N \right\}.$$

We also define the set of marginal population measures, that is

$$\mathbb{V} := \left\{ \sum_{i=1}^N \delta_{x^i}, N \in \mathbb{N}, (x^i, 1 \leq i \leq N) \in \mathcal{X}^N \right\}.$$

For any measure  $\bar{\nu} = \sum_{i=1}^N \delta_{(u^i, x^i)}$  in  $\overline{\mathbb{V}}$ , we will write  $\nu := \sum_{i=1}^N \delta_{x^i}$  its projection on  $\mathbb{V}$ . By convention, if the number of points in the measure is  $N = 0$ ,  $\bar{\nu}$  and  $\nu$  are the trivial zero measures on  $\mathcal{U} \times \mathcal{X}$  and  $\mathcal{X}$ . We introduce for every  $\bar{\nu} \in \overline{\mathbb{V}}$ , every  $g \in \mathfrak{B}(\mathcal{U} \times \mathcal{X}, \mathbb{R})$  and every  $f \in \mathfrak{B}(\mathcal{X}, \mathbb{R})$

$$\langle \bar{\nu}, g \rangle := \int_{\mathcal{U} \times \mathcal{X}} g(u, x) \bar{\nu}(du, dx), \quad \text{and} \quad \langle \nu, f \rangle := \int_{\mathcal{X}} f(x) \nu(dx).$$

Finally, we denote by  $\mathbb{D}(A, B)$  the Skorohod space of  $B$ -valued càdlàg functions on a subset  $A$  of  $\mathbb{R}_+$ . For every process  $(\nu_t, t \in A) \in \mathbb{D}(A, B)$  and  $x \in B$ , we will denote

$$\mathbb{E}_x[f(\nu_t)] := \mathbb{E}[f(\nu_t) | X_0 = x].$$

## 1. DEFINITION OF THE POPULATION

In this section we describe informally the population process. In Section 5.1 we give a rigorous definition as a strong solution of a stochastic differential equation (SDE).

The population is described at any time  $t \in \mathbb{R}_+$  by the finite point measure  $\bar{\nu}_t \in \bar{\mathbb{V}}$  giving the label and the trait of every individual living in the population at this time. We introduce

$$\mathbb{G}(t) := \left\{ u \in \mathcal{U} : \int_{\mathcal{U} \times \mathcal{X}} \mathbb{1}_{\{v=u\}} \bar{\nu}_t(dv, dx) \geq 1 \right\}$$

the set of labels of living individuals at time  $t$ . Each individual is assigned a unique label  $u \in \mathbb{G}(t)$ , and is given a trait that we denote by  $X_t^u$ , and with a slight abuse of notation,  $X_s^u$  will denote the trait of its unique ancestor living at time  $s \in [0, t]$ . Using these notations, we can write

$$\bar{\nu}_t = \sum_{u \in \mathbb{G}(t)} \delta_{(u, X_t^u)} \quad \text{and} \quad \nu_t := \sum_{u \in \mathbb{G}(t)} \delta_{X_t^u}.$$

The initial population is given by a measure  $\bar{z} = \sum_{i=1}^N \delta_{(u^i, x^i)}$ . During their lives, the traits of the individuals in the population evolve according to population-dependent dynamics. We introduce  $\mu \in \mathfrak{B}(\mathcal{X} \times \mathbb{V}, \mathcal{X})$ , such that, in a population  $\bar{\nu}_t$  at time  $t$ , for all  $u \in \mathbb{G}(t)$

$$\frac{dX_t^u}{dt} = \mu(X_t^u, \nu_t).$$

The branching rate is given by a function  $B \in \mathfrak{B}(\mathcal{X} \times \mathbb{V}, \mathbb{R}_+)$ . Thus, an individual with trait  $x$  in a population  $\nu$  at time  $t$  dies at an instantaneous rate  $B(x, \nu_t)$ .

It produces an offspring of  $n$  individuals, where  $n$  is randomly chosen with distribution  $(p_k(x, \nu_t), k \in \mathbb{N})$  of first moment  $m(x, \nu_t)$ . Thus branching events that lead to  $n$  children happen at rate  $B_n(\cdot) := B(\cdot)p_n(\cdot)$ . If  $n \geq 0$ , we denote by  $\mathbf{y} = (y^1, \dots, y^n) \in \mathcal{X}^n$  the offspring traits at birth, randomly chosen according to the law  $K_n(x, \nu_t, \cdot) \mathcal{M}_n(d\cdot)$ , where  $K_n(x, \nu_t, \cdot) \in \mathfrak{B}(\mathcal{X}^n, \mathbb{R}_+)$  and  $\mathcal{M}_n(\cdot)$  is a finite positive measure on  $\mathcal{X}^{\mathbb{N}}$  such that, for all  $A = (A^1, \dots, A^n, A^{n+1}, \dots) \in \mathcal{X}^{\mathbb{N}}$  with  $A^1 = \dots = A^n = 0$ ,  $\mathcal{M}_n(A) = 0$ . For example,  $\mathcal{M}_n(\cdot)$  could be the product measure of the Lebesgue measure on  $\mathcal{X}^n$  and the null measure  $z(\cdot)$  for the other components of the vector  $\mathbf{y} \in \mathcal{X}^{\mathbb{N}}$ , such that

$$\mathcal{M}_n(d\mathbf{y}) := \bigotimes_{i=1}^n dy^i \bigotimes_{i \geq n+1} z(dy^i).$$

The labeling choice for these children is arbitrary yet necessary to uniquely define a stochastic point process in  $\mathbb{D}(\mathbb{R}_+, \bar{\mathbb{V}})$ . Here, for a parent individual of label  $u$ , for all  $1 \leq i \leq n$ , the  $i$ -th child is labeled  $ui$  and its trait is  $y^i$ , the  $i$ -th coordinate of the vector  $\mathbf{y}$ .

We denote by  $T_{\text{Exp}} \in \bar{\mathbb{R}}_+$  the explosion time of the process  $(\bar{\nu}_t, t \geq 0)$ , defined as the limit of its jumps times  $(T_k, k \geq 0)$ . In order to ensure the non-explosion of this process in finite time we introduce the following set of hypotheses.

**Assumption 1.** *We consider the following assumptions:*

(1) *There exists  $\mu_0 \in \mathbb{R}_+$ , such that for all  $(x, \nu) \in \mathcal{X} \times \mathbb{V}$*

$$|\mu(x, \nu)| \leq \mu_0 \left( 1 + |x| + \left| \frac{\int_{\mathcal{X}} y \nu(dy)}{\langle \nu, 1 \rangle} \right| \right).$$

(2) *For all  $(x, \nu) \in \mathcal{X} \times \mathbb{V}$ ,  $B_1(x, \nu) < +\infty$ .*

(3) *There exists  $b_0 \in \mathbb{R}_+$ , such that for all  $(x, \nu) \in \mathcal{X} \times \mathbb{V}$*

$$\sum_{n \neq 1} n B_n(x, \nu) \leq b_0 (1 + |x|).$$

(4) *For all  $(x, \nu, n) \in \mathcal{X} \times \mathbb{V} \times \mathbb{N}^*$ ,*

$$K_n(x, \nu, \mathcal{A}_n(x)) = 0, \quad \text{where } \mathcal{A}_n(x) := \left\{ (y^i, 1 \leq i \leq n) \in \mathcal{X}^n : \sum_{i=1}^n |y^i| > |x| \right\}.$$

The growth rate of the traits of individuals is bounded in the first hypothesis by an exponential growth rate controlled by the trait of each individual and the mean trait in the population. The second point ensures that events that do not change the number of individuals do not accumulate in finite time. The third hypothesis uniformly controls the minimum lifetime of an individual and is used in the proof of Proposition 1.1 to control the first moment of the total mass. This hypothesis, together with the first one, ensures that the lifespan of each individual decreases exponentially at most with its trait. Note that this assumption does not constrain the function  $B_1(\cdot)$ . The last hypothesis restricts the framework under consideration to fragmentation processes that do not create matter or energy. This set of hypotheses is sufficiently large to cover a large portion of models in physics [52] and ecology, subexponential growth being a classical assumption in many stochastic models in ecology and evolution, see *e.g.* [53, 17]. The results presented in this work can be easily extended in a time-inhomogeneous framework, see [48].

**Proposition 1.1.** *Under Assumption 1, the sequence  $(T_k, k \geq 0)$  of jumps times of the process  $(\bar{\nu}_t, t \geq 0)$  tends to infinity almost surely.*

This proposition will be later proved in Section 5.1. We can thus conclude, following the proof of Theorem 2.1 in [46], that under this set of hypotheses, the process  $(\bar{\nu}_t, t \geq 0)$  is uniquely defined on  $\mathbb{R}_+$ . Notice that the spinal construction introduced in Section 2 can be established with less restrictive hypotheses. In this case, accumulation of jumps times may happen in finite time and the spinal construction holds until this explosion time. However, accurately quantifying the explosion time may be impractical, necessitating the identification of such conditions that guarantee the non-explosion of the process. Furthermore, to study the long-term behavior of branching processes, it is crucial that the process does not explode in finite time.

## 2. RESULTS

In this section, we consider the law of a randomly sampled individual in the general branching population described in Section 1. Our main result gives the appropriate change of measure linking this distribution at time  $t$  to the trajectory of an auxiliary process until this time. We then explicitly construct this auxiliary process as a spinal decomposition.

The spinal construction generates a  $\bar{\mathbb{V}}$ -valued process along with the label of a distinguished individual that can change with time. When there is no possible confusion, the label of the spinal individual at any time will be denoted  $e$ , and  $x_e$  its trait. For convenience, we will denote by  $\bar{\mathbb{W}}$  and  $\mathbb{W}$  the sets such that

$$\bar{\mathbb{W}} := \{(u, \bar{\nu}) \in \mathcal{U} \times \bar{\mathbb{V}} : \langle \bar{\nu}, \mathbb{1}_{\{u\} \times \mathcal{X}} \rangle \geq 1\}, \quad \text{and} \quad \mathbb{W} := \{(x, \nu) \in \mathcal{X} \times \mathbb{V} : \langle \nu, \mathbb{1}_{\{x\}} \rangle \geq 1\}.$$

Thus, the spine process is a  $\bar{\mathbb{W}}$ -valued branching process and its marginal is a  $\mathbb{W}$ -valued branching process. We propose here a generalized spinal construction, where branching rates are biased with weight functions chosen in a set  $\mathcal{D}$ , defined by

$$\mathcal{D} := \{F_f \in \mathfrak{B}(\bar{\mathbb{W}}, \mathbb{R}_+^*) \text{ s.t. } (f, F) \in \mathcal{C}^1(\mathcal{X}, \mathbb{R}) \times \mathcal{C}^1(\mathcal{X} \times \mathbb{R}, \mathbb{R}_+^*)\}, \quad (2.1)$$

where for every  $(x, \nu) \in \mathbb{W}$ ,  $F_f(x, \nu) := F(x, \langle \nu, f \rangle)$ . In order to alleviate the notations, we will omit the subscript  $f$  when there is no ambiguity.

In the following, for every  $(u, \bar{\nu}) \in \bar{\mathbb{W}}$  we denote by  $x_u$  the trait of the individual of label  $u$  in the population  $\bar{\nu}$ , and for every  $n \geq 0$  and every  $\mathbf{y} = (y^i, 1 \leq i \leq n) \in \mathcal{X}^n$  we write

$$\bar{\nu}_+(u, \mathbf{y}) := \bar{\nu} - \delta_{(u, x_u)} + \sum_{i=1}^n \delta_{(u_i, y^i)}, \quad \text{and} \quad \nu^+(x, \mathbf{y}) := \nu - \delta_x + \sum_{i=1}^n \delta_{y^i}. \quad (2.2)$$

We introduce the key operator  $\mathcal{G}$  involved in the spinal construction. It is defined for all  $F \in \mathcal{D}$  and  $(x, \nu) \in \mathbb{W}$  by

$$\begin{aligned} \mathcal{G}F(x, \nu) &:= GF(x, \nu) \\ &+ \sum_{n \geq 0} \left\{ B_n(x, \nu) \int_{\mathcal{X}^n} \left[ \sum_{i=1}^n F(y^i, \nu^+(x, \mathbf{y})) - F(x, \nu^+(x, \mathbf{y})) \right] K_n(x, \nu, \mathbf{y}) \mathcal{M}_n(d\mathbf{y}) \right. \\ &\quad \left. + \int_{\mathcal{X}} B_n(\mathbf{x}, \nu) \int_{\mathcal{X}^n} [F(x, \nu^+(\mathbf{x}, \mathbf{y})) - F(x, \nu)] K_n(\mathbf{x}, \nu, \mathbf{y}) \mathcal{M}_n(d\mathbf{y}) \nu(d\mathbf{x}) \right\}, \end{aligned} \quad (2.3)$$

where the operator  $G$  is the generator of the deterministic evolution of the traits between branching events, given for every  $F \in \mathcal{D}$ , and  $(x, \nu) \in \mathbb{W} \times \mathbb{R}_+$  by

$$GF(x, \nu) := D_1 F(x, \langle \nu, f \rangle) \cdot \mu(x, \nu) + D_2 F(x, \langle \nu, f \rangle) \int_{\mathcal{X}} f'(\mathbf{x}) \cdot \mu(\mathbf{x}, \nu) \nu(d\mathbf{x}),$$

where  $D_i F(\cdot, \cdot)$  denotes the derivative of the function  $F \in \mathcal{C}^1(\mathcal{X} \times \mathbb{R}, \mathbb{R}_+^*)$  with respect to  $i$ -th variable. We recall that the dot product between two elements of  $\mathbb{R}^d$  denotes the canonical scalar product. Rigorously, the previously introduced objects are families of operators  $(\mathcal{G}_{x, \nu}, (x, \nu) \in \mathcal{X} \times \mathbb{V})$  and  $(G_{x, \nu}, (x, \nu) \in \mathcal{X} \times \mathbb{V})$ . This abuse of notation will be used for the operators subsequently introduced in this work.

Note that the operator  $\mathcal{G}$  is generally not the generator of a conservative Markov process on  $\mathbb{W}$ . Indeed  $\mathcal{G}1(x, \nu) = B(x, \nu)(m(x, \nu) - 1)$ , which is non-zero if there exists  $(x, \nu) \in \mathcal{X} \times \mathbb{V}$  such that the mean number of children  $m(x, \nu) \neq 1$ . As pointed out in [16], these operators can be seen as Schrödinger operators of the system. In particular, martingales of the system with eigenfunctions  $\psi$  satisfy a similar equation as solutions of the Schrödinger equation, giving a conceptual interpretation of eigenvalues as energy eigen states of the system and eigenfunctions as stationary configurations or conserved quantities.

To ensure that the functions  $\mathcal{G}F$  are finite on  $\mathbb{W}$ , we make an assumption on the functions  $F \in \mathcal{D}$ , that will be necessary to establish the spine Dynamics 2.2 and 2.3.

**Assumption 2.** *Let  $(p_n, n \in \mathbb{N})$  and  $(K_n, n \in \mathbb{N})$  be the reproduction parameters of the original branching process and  $F \in \mathcal{D}$  a weight-function. For all  $(x, x_e, \nu) \in \mathcal{X} \times \mathbb{W}$ ,*

$$\begin{aligned} \sum_{n \in \mathbb{N}} p_n(x, \nu) \int_{\mathcal{X}^n} F(x_e, \nu^+(x, \mathbf{y})) K_n(x, \nu, \mathbf{y}) \mathcal{M}_n(d\mathbf{y}) \\ + \sum_{n \in \mathbb{N}} p_n(x, \nu) \int_{\mathcal{X}^n} \sum_{i=1}^n F(y^i, \nu^+(x, \mathbf{y})) K_n(x, \nu, \mathbf{y}) \mathcal{M}_n(d\mathbf{y}) < +\infty. \end{aligned}$$

Notice that for a chosen set of parameters of the original branching process  $(\nu_s, s \geq 0)$ , this assumption restricts the set of suitable weight functions.

Now we can introduce the  $\overline{\mathbb{W}}$ -valued spine process associated with a function  $\psi \in \mathcal{D}$  satisfying Assumption 2, that will be called the  $\psi$ -spine process in order to keep track of the function used in its construction. For such functions  $\psi$ , the generator  $\widehat{\mathcal{L}}_\psi$  of the  $\psi$ - spine process is defined for

all  $h \in \mathcal{C}^1(\mathcal{U} \times \mathcal{X}, \mathbb{R})$ ,  $H \in \mathcal{C}^1(\mathcal{U} \times \mathbb{R}, \mathbb{R}_+^*)$ , and for all  $(e, \bar{\nu}) \in \overline{\mathbb{W}}$  by

$$\begin{aligned} \widehat{\mathcal{L}}_\psi H(e, \langle \bar{\nu}, h \rangle) &:= \widehat{G}H(e, \langle \bar{\nu}, h \rangle) + \sum_{n \geq 0} \int_{\mathcal{U} \times \mathcal{X}} B_n(x, \nu) \\ &\times \int_{\mathcal{X}^n} K_n(x, \nu, \mathbf{y}) \left\{ \mathbb{1}_{\{u=e\}} \sum_{i=1}^n [H(ei, \langle \bar{\nu}_+(e, \mathbf{y}), h \rangle) - H(e, \langle \bar{\nu}, h \rangle)] \frac{\psi(y^i, \nu^+(x_e, \mathbf{y}))}{\psi(x_e, \nu)} \right. \\ &\left. + \mathbb{1}_{\{u \neq e\}} [H(e, \langle \bar{\nu}_+(u, \mathbf{y}), h \rangle) - H(e, \langle \bar{\nu}, h \rangle)] \frac{\psi(x_e, \nu^+(x_u, \mathbf{y}))}{\psi(x_e, \nu)} \right\} \mathcal{M}_n(d\mathbf{y}) \bar{\nu}(du, dx), \quad (2.4) \end{aligned}$$

where

$$\widehat{G}H(e, \langle \bar{\nu}, f \rangle) := D_2H(e, \langle \bar{\nu}, h(\cdot) \rangle) \int_{\mathcal{U} \times \mathcal{X}} \frac{\partial h}{\partial x}(u, x) \cdot \mu(x, \nu) \bar{\nu}(du, dx).$$

Notice that the branching rates of both spinal and non-spinal individuals are biased by the function  $\psi$ . Assumption 2 ensures that the total branching rate is finite from every state of the spinal process, however it is not sufficient to avoid explosion of this process in finite time. Dynamics 2.2 and 2.3 below will provide a more detailed explanation of the spine process associated with this generator.

Finally, we introduce for all  $t \geq 0$ , the  $\mathcal{U}$ -valued random variable  $U_t$  that picks an individual alive at time  $t$ . Its law is characterized by the function  $p_u(\bar{\nu}_t)$  which yields the probability to choose the individual of label  $u$  in the set  $\mathbb{G}(t)$ . We can now state our main result, that is a Girsanov-type formula for the spinal change of measure. It characterizes the joint probability distribution of  $(U_t, (\bar{\nu}_s, s \leq t))$ - that is the randomly sampled individual in the population  $\bar{\nu}_t$  at time  $t$  and the whole trajectory of the population until this time- and links it to the law of the spine process through a path-integral formula.

**Theorem 2.1.** *Let  $t \geq 0$ , and  $\bar{z} \in \overline{\mathbb{W}}$ . Let  $\psi \in \mathcal{D}$  satisfying Assumption 2. Let  $((E_t, \bar{\chi}_t), t \geq 0)$  be the time-inhomogeneous  $\overline{\mathbb{W}}$ -valued branching process with interactions defined by the infinitesimal generator  $\widehat{\mathcal{L}}_\psi$  introduced in (2.4). Let  $\widehat{T}_{Exp}$  denote its explosion time and  $((Y_t, \chi_t), t \geq 0)$  its projection on  $\mathbb{W}$ .*

*For every non-negative measurable function  $H$  on  $\mathcal{U} \times \mathbb{D}([0, t], \mathbb{V})$ :*

$$\begin{aligned} \mathbb{E}_{\bar{z}} \left[ \mathbb{1}_{\{T_{Exp} > t, \mathbb{G}(t) \neq \emptyset\}} H(U_t, (\bar{\nu}_s, s \leq t)) \right] = \\ \langle z, \psi(\cdot, z) \rangle \mathbb{E}_{\bar{z}} \left[ \mathbb{1}_{\{\widehat{T}_{Exp} > t\}} \xi(E_t, (\bar{\chi}_s, s \leq t)) H(E_t, (\bar{\chi}_s, s \leq t)) \right], \end{aligned}$$

where:

$$\xi(E_t, (\bar{\chi}_s, s \leq t)) := \frac{p_{E_t}(\bar{\chi}_t)}{\psi(Y_t, \chi_t)} \exp \left( \int_0^t \frac{\mathcal{G}\psi(Y_s, \chi_s)}{\psi(Y_s, \chi_s)} ds \right).$$

We take inspiration from the work of Bansaye [3] for the proof. The idea is to decompose both processes on their possible trajectories, then establish by induction on the successive jumps times the equality in law for the trajectories between these times.

The process  $((E_t, \bar{\chi}_t), t \geq 0)$  gives at any time  $t \geq 0$  the label of the spinal individual- that encodes the whole spine lineage- and the spinal population. Our result thus links, for every  $\psi$ , the sampling of an individual and the trajectory of the population to the trajectory of the spine process. The path integral term that links these two terms is difficult to handle in general and finding eigenfunctions of  $\mathcal{G}$  may greatly simplify the expression [16, 2, 3]. Finding such functions for single type, density-dependent populations is possible in models with simple interactions [3][Section 3]. Nevertheless, this becomes a challenging issue in the majority of scenarios. Subsequent sections of this work will explore applications of this formula where the path-integral component is tractable.



We first introduce additional notations concerning the dynamics of the spine process given by the generators introduced in (2.4). Following notations of Section 1, and disregarding the dependency on the chosen function  $\psi \in \mathcal{D}$  satisfying Assumption 2 in the subsequent branching parameters, we introduce the dynamics of the traits in the spine process. As previously discussed, the construction distinguishes dynamics of the spine from the rest of the individuals. We first introduce the branching parameters of the individuals outside the spine in a spinal population.

**Dynamics 2.2** (Individuals outside the spine). *For all  $(n, x, (x_e, \nu))$  in  $\mathbb{N}^* \times \mathcal{X} \times \mathbb{W}$ ,*

- (1)  $\widehat{K}_n(x_e, x, \nu, \cdot) \in \mathfrak{B}(\mathcal{X}^n)$  is the density of the traits at birth of the  $n$  children generated by a non-spinal individual of trait  $x$  in a spinal population  $(x_e, \nu)$ .

$$\widehat{K}_n(x_e, x, \nu, \cdot) := \frac{1}{\widehat{\Gamma}_n(x_e, x, \nu)} \psi(x_e, \nu^+(x, \cdot)) K_n(x, \nu, \cdot), \quad (2.5)$$

where  $\widehat{\Gamma}_n(\cdot)$  is the normalization function, defined as

$$\widehat{\Gamma}_n(x_e, x, \nu) := \int_{\mathcal{X}^n} \psi(x_e, \nu^+(x, \mathbf{y})) K_n(x, \nu, \mathbf{y}) \mathcal{M}_n(d\mathbf{y}).$$

We recall the definition of  $\nu^+$  in (2.2).

- (2) The law  $(\widehat{p}_n(x_e, x, \nu), n \in \mathbb{N})$  of the number of children of a non-spinal individual of trait  $x$  branching in a spinal population  $(x_e, \nu)$ , is defined for all  $n \in \mathbb{N}$  as

$$\widehat{p}_n(x_e, x, \nu) := \frac{1}{\sum_{k \in \mathbb{N}} \widehat{\Gamma}_k(x_e, x, \nu) p_k(x, \nu)} \widehat{\Gamma}_n(x_e, x, \nu) p_n(x, \nu).$$

Note that Assumption 2 ensures that the sum in the denominator is finite.

- (3) Each individual of trait  $x$  outside the spine of trait  $x_e$  in a population  $\nu$ , branches to  $n$  children at rate  $\widehat{B}_n(x_e, x, \nu)$ , defined as

$$\widehat{B}_n(x_e, x, \nu) := \frac{\widehat{\Gamma}_n(x_e, x, \nu)}{\psi(x_e, \nu)} B_n(x, \nu).$$

The total branching rate outside the spine is defined, for all  $(x_e, \nu) \in \mathbb{W}$  by

$$\widehat{\tau}(x_e, \nu) := \int_{\mathcal{X}} \sum_{n \geq 0} \widehat{B}_n(x_e, x, \nu) \nu(dx) - \sum_{n \geq 0} \widehat{B}_n(x_e, x_e, \nu).$$

We now introduce the branching parameters of the spine in a  $\psi$ -spinal construction.

**Dynamics 2.3** (Spinal individual). *For all  $(n, (x_e, \nu))$  in  $\mathbb{N}^* \times \mathbb{W}$ ,*

- (1)  $\widehat{K}_n^*(x_e, \nu, \cdot) \in \mathfrak{B}(\mathcal{X}^n)$  is the density of the traits at birth of the  $n$  children generated by the spinal individual of trait  $x_e$  in a population  $\nu$ . For all  $\mathbf{y} \in \mathcal{X}^n$ ,

$$\widehat{K}_n^*(x_e, \nu, \mathbf{y}) := \frac{1}{\widehat{\Gamma}_n^*(x_e, \nu)} \sum_{i=1}^n \psi(y^i, \nu^+(x_e, \mathbf{y})) K_n(x_e, \nu, \mathbf{y}), \quad (2.6)$$

where  $\widehat{\Gamma}_n^*(\cdot)$  is the normalization function, defined as

$$\widehat{\Gamma}_n^*(x_e, \nu) := \int_{\mathcal{X}^n} \sum_{i=1}^n \psi(y^i, \nu^+(x_e, \mathbf{y})) K_n(x_e, \nu, \mathbf{y}) \mathcal{M}_n(d\mathbf{y}).$$

- (2) The law  $(\widehat{p}_n^*(x_e, \nu), n \in \mathbb{N})$  of the number of children of the spinal individual of trait  $x_e$  branching in a population  $\nu$ , is defined for all  $k \in \mathbb{N}$  as

$$\widehat{p}_n^*(x_e, \nu) := \frac{1}{\sum_{k \in \mathbb{N}} \widehat{\Gamma}_k^*(x_e, \nu) p_k(x_e, \nu)} \widehat{\Gamma}_n^*(x_e, \nu) p_n(x_e, \nu). \quad (2.7)$$

- (3) The spinal individual of trait  $x_e$  in a population  $\nu$  at time  $t$ , branches to  $n$  children at a rate  $\widehat{B}_n^*(x_e, \nu)$ , defined as

$$\widehat{B}_n^*(x_e, \nu) := \frac{\widehat{\Gamma}_n^*(x_e, \nu)}{\psi(x_e, \nu)} B_n(x_e, \nu). \quad (2.8)$$

- (4) When the spinal individual of trait  $x_e$  branches in a population  $\nu$  and is replaced by  $n$  children with trait  $\mathbf{y}$ , the integer-valued random variable  $J(x_e, \nu, \mathbf{y})$  choosing the new spinal individual after a spinal branching event is given, for all  $1 \leq j \leq n$  by

$$\mathbb{P}(J(x_e, \nu, \mathbf{y}) = j) = \frac{\psi(y^j, \nu^+(x_e, \mathbf{y}))}{\sum_{i=1}^n \psi(y^i, \nu^+(x_e, \mathbf{y}))}. \quad (2.9)$$

The first spine in an initial population  $z$  is chosen according to the same law.

The total branching rate from every spinal state  $(x_e, \nu) \in \mathbb{W}$  is denoted by

$$\widehat{\tau}_{\text{tot}}(x_e, \nu) := \sum_{n \geq 0} \widehat{B}_n^*(x_e, \nu) + \widehat{\tau}(x_e, \nu). \quad (2.10)$$

The finiteness of the series  $\sum_{k \geq 0} \widehat{\Gamma}_k^*$ , and  $\sum_{k \geq 0} \widehat{\Gamma}_k$  is ensured by Assumption 2. Notice that  $\widehat{K}_0^* = 0$ , therefore the spinal individual cannot die without children and the spinal population never goes extinct. Notice that for  $\psi \equiv 1$ , individuals outside the spine follow the same dynamics as the individuals in the population ( $\nu_t \geq 0$ ). In this case, the spinal individual of trait  $x_e$  branches in a population  $\nu$  with rate  $m(x_e, \nu)B(x_e, \nu)$ , where  $m(\cdot)$  is the mean number of children that is finite under Assumption 2. The random number of children at a branching event thus follows the size-biased law  $np_n(\cdot)/m(\cdot)$  and the new spinal individual is chosen uniformly among the offspring.

The assertion of Theorem 2.1 is proved to be valid until the first explosion time of both processes. We established in Proposition 1.1 that under Assumption 1 the branching process does not explode in finite time. To ensure the non explosion of the spine process, we have to consider an additional assumption on the weight function  $\psi$  used for the construction.

**Assumption 3.** *There exists a positive continuous function  $\widehat{b}_0$  on  $\mathbb{R}_+$ , such that for all  $(x, \nu) \in \mathcal{X} \times \mathbb{V}$*

$$\sum_{n \neq 1} n \left( \widehat{B}_n(x, \nu) + \widehat{B}_n^*(x, \nu) \right) \leq \widehat{b}_0(t) (1 + |x|).$$

This assumption, involving both the branching parameters and the function  $\psi$ , is stronger than Assumption 2. The set of weight functions that can be used to construct a spine process that does not explode in finite time may differ from one model to another. However, one may rather use more restrictive conditions that are sufficient for every branching process under Assumption 1. For example, in mass-conservative models, taking  $\psi(x, \nu) = x$  ensures the non-explosion of the spine process regardless the initial branching process that satisfies Assumption 1.

**Proposition 2.4.** *Under Assumption 1, for every  $\psi \in \mathcal{D}$  satisfying Assumptions 2 and 3, the  $\psi$ -spine process  $((E_t, \bar{\chi}_t) \geq 0)$  defined by the infinitesimal generator  $\widehat{\mathcal{L}}_\psi$  introduced in (2.4) does not explode in finite time. Furthermore the generator  $\widehat{L}_\psi$  of the marginal spine process  $((Y_t, \chi_t), t \geq 0)$  given by Dynamics 2.2 and 2.3, is defined for all function  $F \in \mathcal{D}$  and all  $(x, \nu) \in \mathcal{X} \times \mathbb{V}$  by*

$$\widehat{L}_\psi F(x, \nu) := \frac{\mathcal{G}[\psi F](x, \nu)}{\psi(x, \nu)} - \frac{\mathcal{G}\psi(x, \nu)}{\psi(x, \nu)} F(x, \nu).$$

The proof of non-explosion is shown following the proof of Proposition 1.1, and the expression of the generator of the marginal spine process is purely computational. The detailed proof is presented in Appendix A.

It follows that the marginal law of the spine process is characterized by the operator  $\mathcal{G}$  and the weight function  $\psi$ .

**Corollary 2.5.** *Let  $t \geq 0$  and  $\bar{z} \in \bar{\mathbb{V}}$ . Let  $\psi \in \mathcal{D}$  satisfying Assumption 3. Under Assumption 1, for any  $f$  non-negative measurable function on  $\mathcal{U} \times \mathbb{D}([0, t], \mathcal{X} \times \bar{\mathbb{V}})$ ,*

$$\begin{aligned} \mathbb{E}_{\bar{z}} \left[ \sum_{u \in \mathbb{G}(t)} \psi(X_t^u, \bar{\nu}_t) f(u, (X_s^u, \bar{\nu}_s), 0 \leq s \leq t) \right] \\ = \langle z, \psi(\cdot, z, 0) \rangle \mathbb{E}_{\bar{z}} \left[ \exp \left( \int_0^t \frac{\mathcal{G}\psi(Y_s, \chi_s)}{\psi(Y_s, \chi_s)} ds \right) f(E_t, (Y_s, \bar{\chi}_s), 0 \leq s \leq t) \right]. \end{aligned}$$

*Proof.* We use Assumptions 1 and 3 to ensure that  $T_{\text{Exp}}$  and  $\widehat{T}_{\text{Exp}}$  are almost surely infinite. Let  $f$  be non-negative measurable function on  $\mathcal{U} \times \mathbb{D}([0, t], \mathcal{X} \times \bar{\mathbb{V}})$ . We introduce  $H$  the non-negative measurable function defined for all  $(u, \bar{z}_s, s \leq t) \in \mathcal{U} \times \mathbb{D}([0, t], \mathbb{V})$  by

$$H(u, \bar{z}_s, s \leq t) := \psi(X_t^u, z_t) f(u, (X_s^u, \bar{z}_s), s \leq t) \langle z_t, 1 \rangle.$$

The corollary is thus a direct application of Theorem 2.1 to the function  $H$  with a uniformly sampled individual.  $\square$

This formula gives a change of probability that involves the function  $\mathcal{G}\psi/\psi$  with a path-integral formula. This study is related to Feynman-Kac path measures and semigroups. We refer to [19] for an overview on this subject. In the case of a branching process with interactions, the integral term depends on the trajectory of the whole spinal population. In general cases with interactions, the branching property is not satisfied and the so-called Many-to-One formula- see *e.g.* Proposition 9.3 in [4]- falls apart. However, if  $\psi$  is an eigenfunction of the operator  $\mathcal{G}$ , then we have the following Many-to-One formula for any non-negative measurable function  $g$  on  $\mathbb{D}([0, t], \mathcal{X})$

$$\mathbb{E}_{\bar{z}} \left[ \sum_{u \in \mathbb{G}(t)} \psi(X_t^u, \bar{\nu}_t) g(X_s^u, s \leq t) \right] = C_t \mathbb{E}_{\bar{z}} [g(Y_s, s \leq t)],$$

where  $C_t$  is a time-dependent positive constant. This formula, established in [16] for branching processes without interactions, reduces the empirical measure of the trajectories of all the individuals until time  $t$  to the law of the trajectory of a unique individual in the spinal construction, the spinal individual. The spinal individual in this case can be considered as a typical individual, reflecting the average behavior of the whole population.

**Remark 2.6.** *If we assume for all  $(x, \nu) \in \mathcal{X} \times \mathbb{V}$ , that  $B(x, \nu)$  and for all  $k \geq 0$ ,  $p_k(x, \nu) = p_k$  in the considered branching process, taking  $\psi \equiv 1$  gives the classical Many-to-One formula [5]:*

$$\mathbb{E}_{\bar{z}} \left[ \sum_{u \in \mathbb{G}(t)} g(X_s^u, s \leq t) \right] = \mathbb{E}_{\bar{z}} [\langle \nu_t, 1 \rangle] \mathbb{E}_{\bar{z}} [g(Y_s, s \leq t)],$$

where the average number of individuals in the population is given by

$$\mathbb{E}_{\bar{z}} [\langle \nu_t, 1 \rangle] = \langle z, 1 \rangle \exp(B(m-1)t).$$

### 3. KESTEN-STIGUM CRITERION

In this section, we present a generalized Kesten-Stigum theorem for structured processes with interactions. For every functions  $\psi \in \mathcal{D}$ , we exhibit a non-negative martingale of the original branching process with interactions ( $\nu_t \geq 0$ ) associated with the  $\psi$ -spinal construction and, under some assumptions on the set of functions  $\psi$ , we present a result on the degeneracy of its limiting martingale. For this section, we will suppose that Assumption 1 holds to prevent explosion of the process in finite time.

**Proposition 3.1.** *Under Assumption 1, for every  $\psi \in \mathcal{D}$  satisfying Assumption 2,*

$$W_t(\psi) := \sum_{u \in \mathbb{G}(t)} \exp \left( - \int_0^t \frac{\mathcal{G}\psi(X_s^u, \nu_s)}{\psi(X_s^u, \nu_s)} ds \right) \psi(X_t^u, \nu_t) \quad (3.1)$$

*is a non-negative martingale with respect to the filtration  $(\mathcal{F}_t, t \geq 0)$  generated by the original process. It almost surely converges to a random variable  $W(\psi) \in [0, \infty)$ .*

Notice that if the process  $(\nu_t, t \geq 0)$  goes extinct almost surely, then  $W(\psi) = 0$  almost surely. However, even on the survival event, the martingale  $(W_t(\psi), t \geq 0)$  may also almost surely converge to 0.

We will establish necessary and sufficient  $L \log L$  criteria for the non-degeneracy of the limiting martingale. First we present the result for bias function  $\psi \equiv 1$  and retrieve classical results. Then we propose an extension to a large class of functions  $\psi$ , to establish result on various martingales of the system.

**3.1. The particular case  $\psi \equiv 1$ .** When  $\psi \equiv 1$  or any positive constant value, we have for all  $x, \nu \in \mathbb{W} \times \mathbb{R}_+$

$$\frac{\mathcal{G}\psi(x, \nu)}{\psi(x, \nu)} = B(x, \nu)(m(x, \nu) - 1), \quad \widehat{K}_n^*(x, \nu) = K_n(x, \nu), \quad \text{and} \quad \widehat{p}_n^*(x, \nu) = \frac{np_n(x, \nu)}{m(x, \nu)}.$$

To prove a Kesten-Stigum theorem, we need some assumptions on the considered branching processes.

**Assumption 4.** *Assumptions 1 and 3 hold true and there exist  $c, C, \underline{B}, \overline{B} \in \mathbb{R}_+^*$  such that, for all  $x, \nu \in \mathbb{W}$*

$$c \leq B(x, \nu)(m(x, \nu) - 1) \leq C \quad \text{and} \quad \underline{B} \leq B(x, \nu) \leq \overline{B}.$$

This assumption ensures that the process is strongly supercritical in the sense that the uniform lower bound of the reproductive law is strictly greater than 1 and the only absorbing state for the branching process is the null measure  $\nu \equiv 0$ . This uniform hypothesis could be partially relaxed under strong positivity assumptions on the generator of the branching process with interactions, see [38] in the discrete case and [33] Chapter 3 for continuous time. It also restricts the setting to branching processes with bounded branching rates.

**Proposition 3.2.** *We assume that  $\psi \equiv 1$  and that Assumption 4 holds true. Let*

$$\overline{p}_n := \sup_{(x, \nu) \in \mathbb{W}} p_n(x, \nu) \quad \text{and} \quad \underline{p}_n := \inf_{(x, \nu) \in \mathbb{W}} p_n(x, \nu).$$

*If  $\sum_{n \geq 1} \overline{p}_n < +\infty$  and  $\sum_{n \geq 1} \underline{p}_n > 0$ , we can introduce  $\overline{L}$  and  $\underline{L}$  the  $\mathbb{N}^*$ -valued random variables of law given respectively by  $(\overline{p}_n / \sum_{k \geq 1} \overline{p}_k, n \geq 0)$  and  $(\underline{p}_n / \sum_{k \geq 1} \underline{p}_k, n \geq 0)$ .*

*In this case, we have the following results on the limiting martingale:*

- *If  $\mathbb{E}[\overline{L} \log(\overline{L})] < +\infty$ , then for all initial measure  $z \in \mathbb{V}$ ,  $\mathbb{E}_z[W(1)] = \langle z, 1 \rangle$ .*
- *If  $\mathbb{E}[\underline{L} \log(\underline{L})] = +\infty$ , then  $W(1) = 0$  almost surely.*

*Proof.* This proposition is a corollary of Theorem 3.4, with  $\psi \equiv 1$ . □

Notice that these conditions are similar to the conditions (16b) and (18b) in [2]. In the case of constant reproductive law ( $p_n, n \geq 0$ ), these two conditions become a dichotomy [37]. Athreya [2] showed that this dichotomy remains valid for multitype Galton-Watson processes with a finite set of traits and no interactions.

**Remark 3.3.** *We recall that for non-structured branching processes the size  $N_t := \langle \nu_t, 1 \rangle$  of the population, is such that  $\mathbb{E}[N_t] = N_0 \exp \left( \int_0^t B(m-1) ds \right)$ . One may ask under what conditions this exponential growth accurately reflects the rate of increase in the population size. In this particular case,  $W_t(1)/N_0 = N_t/\mathbb{E}[N_t]$ , thus finding conditions for the non-degeneracy of this*

martingale gives a direct answer to the question. In the general case with interactions, Corollary 2.5 gives

$$W_t(1) \frac{\mathbb{E}[N_t]}{N_0} = \sum_{u \in \mathbb{G}(t)} e^{-\int_0^t B(X_s^u, \nu_s)(m(X_s^u, \nu_s) - 1) ds} \mathbb{E}_{\bar{z}} \left[ e^{\int_0^t B(Y_s, \chi_s)(m(Y_s, \chi_s) - 1) ds} \right],$$

that is close to  $N_t$  if the mean behavior of the spinal individual is not far from the averaged behavior of all the individuals in the branching process with interactions.

**3.2. The general case  $\psi \in \mathcal{D}$ .** Note that Proposition 3.1 holds true for every function  $\psi \in \mathcal{D}$ . However, in order to prove the Kesten-Stigum result on the limiting martingale, we have to restrict the set of functions  $\psi$ .

**Assumption 5.** We recall the notation  $\nu^+(x, \mathbf{y}) := \nu + \sum_{i=1}^n \delta_{y^i} - \delta_x$ , introduced in (2.2). For all  $(x_e, (x, \nu)) \in \mathcal{X} \times \mathbb{W}$ , for all  $k \geq 0$  and for all  $\mathbf{y}$  in a subset  $\mathcal{A} \subset \mathcal{X}^n$  with  $K_n(x, \nu, \mathbf{y}) \mathcal{M}_n(\mathcal{A}) > 0$ :

$$\psi(x_e, \nu^+(x, \mathbf{y})) = \psi(x_e, \nu).$$

Note that this assumption is weaker than assuming that the function  $\psi$  is independent of the population. In fact the function  $\psi$  is allowed to depend on a quantity that is conserved at non-spinal jump events, for example the total population mass if the jumps are mass conservative. Another interesting case is when  $\psi$  is a function of a coupling variable that derives from the population, like the quantity of available resources for the individuals. This technical assumption is necessary to establish a spinal decomposition.

**Assumption 6.** There exists  $M \in \mathbb{R}_+^*$  such that, for all  $((x, \nu), n) \in \mathbb{W} \times \mathbb{N}^*$  and for all  $1 \leq j \leq n$ :

$$\begin{aligned} \int_{\mathcal{X}^n} \psi(y^j, \nu^+(x, \mathbf{y})) \log(\psi(y^j, \nu^+(x, \mathbf{y}))) K_n(x, \nu, \mathbf{y}) \mathcal{M}_n(d\mathbf{y}) \\ \leq M \int_{\mathcal{X}^n} \sum_{i=1}^n \psi(y^i, \nu^+(x, \mathbf{y})) K_n(x, \nu, \mathbf{y}) \mathcal{M}_n(d\mathbf{y}). \end{aligned}$$

This technical assumption ensures that the function  $\psi$  does not assign too much load to any specific child at a branching event. Notice that bounded functions satisfy this hypothesis. This hypothesis is sufficient to ensure the degeneracy of the limiting martingale.

**Assumption 7.** Assumptions 1 and 3 hold true and there exist  $c, C, \bar{\tau}, \underline{\tau} \in \mathbb{R}_+^*$  such that, for all  $(x, \nu) \in \mathbb{W}$

$$c \leq \frac{\mathcal{G}\psi(x, \nu)}{\psi(x, \nu)} \leq C \quad \text{and} \quad \underline{\tau} \leq \sum_{n \geq 0} \widehat{B}_n^*(x, \nu) \leq \bar{\tau},$$

where  $\mathcal{G}$  and  $\widehat{B}_n^*$  are respectively defined in (2.3) and (2.8).

The hypothesis on  $\mathcal{G}\psi/\psi$  controls the range of variations of the exponential term in the martingale  $W_t(\psi)$ . Note that all eigenfunctions of the operator  $\mathcal{G}$  satisfy this hypothesis. This assumption is also satisfied in all cellular models with exponential growth and mass conservative branching mechanisms, when  $\psi(x) = x$  and the branching rates are bounded functions. The second point ensures that the branching events of the spinal individual in the  $\psi$ -spinal process do not stop nor accumulate too fast.

As an example to illustrate Assumption 7, we consider a toy model, used to describe cellular growth and division mechanisms [6], where the trait space is  $\mathcal{X} = (0, x_M)$  with  $x_M > 0$ , and the branching events are mass conservative, *i.e.*

$$K_n(x, \nu, \mathbf{y}) = \mathbb{1}_{\{\sum_{i=1}^n y^i = x\}}.$$

We also assume that in this model, the trait grows at a logistic rate, such that  $\mu(x, \nu) = \mu_0(\nu)x(x_M - x)$ . We consider a bias function  $\psi(x) = x$ . In this case Assumption 6 is satisfied and for all  $(x, \nu) \in \mathbb{W}$

$$\frac{\mathcal{G}\psi(x, \nu)}{\psi(x)} = \mu_0(\nu)(x_M - x) - B_0(x, \nu).$$

In this case, a sufficient condition for the upper bounds in Assumption 7 to be satisfied is the boundness of the growth rate and the branching rates. For instance logistic functions of the trait and population size satisfy this condition.

We can now express the main result of this section.

**Theorem 3.4.** *Let  $\psi \in \mathcal{D}$  satisfies Assumption 7 and introduce for all  $n \in \mathbb{N}$*

$$\widehat{p}_n^* := \sup_{(x, \nu) \in \mathbb{W}} \widehat{p}_n^*(x, \nu), \quad \underline{\widehat{p}}_n^* := \inf_{(x, \nu) \in \mathbb{W}} \widehat{p}_n^*(x, \nu),$$

$$\widehat{K}_n^*(\cdot) := \sup_{(x, \nu) \in \mathbb{W}} \widehat{K}_n^*(x, \nu, \cdot) \quad \text{and} \quad \underline{\widehat{K}}_n^*(\cdot) := \inf_{(x, \nu) \in \mathbb{W}} \widehat{K}_n^*(x, \nu, \cdot),$$

where  $(\widehat{p}_n^*, k \geq 0)$  and  $\widehat{K}_n^*$  are respectively the law of the number of children and the measure giving the offspring traits for the spine, defined respectively in (2.7) and (2.6).

(1) Under Assumption 5, if

$$\sum_{n \geq 1} \widehat{p}_n^* < +\infty, \quad \sup_{n \geq 1} \int_{\mathcal{X}^n} \widehat{K}_n^*(\mathbf{y}) \mathcal{M}_n(d\mathbf{y}) < +\infty$$

and

$$\sum_{n \geq 1} \widehat{p}_n^* \int_{\mathcal{X}^n} \sup_{(x, \nu) \in \mathbb{W}} \left[ \log \left( \sum_{i=1}^n \psi(y^i, \nu^+(x, \mathbf{y})) \right) \right] \widehat{K}_n^*(\mathbf{y}) \mathcal{M}_n(d\mathbf{y}) < +\infty \quad (3.2)$$

then, for all initial measure  $z \in \mathbb{V}$ , the limit of the martingale  $W_t(\psi)$  satisfies

$$\mathbb{E}_z [W(\psi)] = \langle z, \psi(\cdot, z) \rangle.$$

(2) Under Assumption 6,

$$\sum_{n \geq 1} \widehat{p}_n^* \int_{\mathcal{X}^n} \inf_{(x, \nu) \in \mathbb{W}} \left[ \log \left( \sum_{i=1}^n \psi(y^i, \nu^+(x, \mathbf{y})) \right) \right] \widehat{K}_n^*(\mathbf{y}) \mathcal{M}_n(d\mathbf{y}) = +\infty \quad (3.3)$$

implies that  $W(\psi) = 0$  almost surely.

The idea of the proof, based on the conceptual proofs established in [45, 27, 3], is to use Theorem 2.1 to consider the dual problem associated with the  $\psi$ -process. Then consider the spinal process as a process with immigration where the spinal individual provides new individuals at a  $\psi$ -biased rate, and the new spine is the new source of immigration. However as the function  $\psi$  is not constant, the spinal construction also changes the dynamics of individuals outside the spine, that do not behave as those in the original process ( $\nu_t, t \geq 0$ ). Assumptions 5 and 6 are used to control the behavior of the individuals outside the spine. In fact only the upper bounds in Assumption 7 are necessary to ensure the non-degeneracy of the limiting martingale. In our previous toy model, with bounded branching and growth rates, this condition is satisfied with the bias function  $\psi(x) = x$ . The 'L log L' condition (3.2), is also satisfied in this framework. Thus the martingale

$$W_t(\psi) := \sum_{u \in \mathbb{G}(t)} X_t^u \exp \left\{ \int_0^t (\mu_0(\nu_s)(x_M - X_s^u) - B_0(X_s^u, \nu_s)) ds \right\}$$

satisfies, for all  $z = \sum_{i=1}^n \delta_{x^i}$ , the asymptotic equality

$$\mathbb{E}_z \left[ \lim_{t \rightarrow \infty} \sum_{u \in \mathbb{G}(t)} X_t^u \exp \left\{ \int_0^t (\mu_0(\nu_s)(x_M - X_s^u) - B_0(X_s^u, \nu_s)) ds \right\} \right] = \sum_{i=1}^n x^i.$$

This equality gives an information about the asymptotic growth rate of the biomass  $\sum_{u \in \mathbb{G}(t)} X_t^u$  of the population. It grows at an exponential rate, that is the difference between the growth rate and the death rate of each individual in the population.

#### 4. A YULE PROCESS WITH COMPETITION

In this section, we take interest into simulation methods based on the spinal construction. When an eigenfunction of the operator  $\mathcal{G}$  is available, relevant information on the original process can be captured by simulating only the spine, thereby reducing the need to simulate the entire population, which can be computationally expensive. We propose here an alternative application of the spinal construction method, when an eigenfunction of the operator  $\mathcal{G}$  is out of reach. In this case, as the path integral term does not simplify, the entire spinal population needs to be simulated to obtain relevant information. Nevertheless, it is possible to use the change of measure associated with the spine construction to simplify the interactions and considerably enhance the simulation efficiency.

As a toy model to illustrate this idea, we consider a time-inhomogeneous Yule process with competitive interactions between the individuals, affecting their traits. The individuals in the population are characterized by their trait  $x \in \mathbb{R}_+^*$  that can be for example a mass or a size. An individual with trait  $x$  divides at an instantaneous rate  $rx$  where  $r > 0$ , into two children of size  $\Lambda x$  and  $(1 - \Lambda)x$ , where  $\Lambda$  is a  $[0, 1]$ -valued random variable with probability density function (p.d.f)  $q$ . We assume that:

$$m_{\text{div}} := \mathbb{E}[\Lambda] \in (0, 1), \quad \text{and} \quad K_{\text{div}} := \mathbb{E} \left[ \frac{1}{\Lambda(1 - \Lambda)} \right] < +\infty. \quad (4.1)$$

This mass-conservative mechanism of division is classical in cell modeling [26]. Moreover, each individual experiences the influence of the whole population, leading to a reduction of their trait. Consequently, at an instantaneous rate  $dN_t$  - where  $N_t$  is the population size at time  $t$ , and  $d > 0$  - each individual loses a fraction  $(1 - \Theta)$  of its size.  $\Theta$  is a  $[0, 1]$ -valued random variable with  $p$  its p.d.f and we assume that

$$m_{\text{loss}} := \mathbb{E}[\Theta] \in (0, 1) \quad \text{and} \quad K_{\text{loss}} := \mathbb{E} \left[ \frac{1}{\Theta} \right] < +\infty. \quad (4.2)$$

These events can be interpreted as an inhibition of reproductive material in cells due to competitive interactions within the population. Finally, we consider that the trait of each individual grows exponentially at an instantaneous rate  $\mu(t)$ . The dependency in time of these parameters can represent an external control or the effect of a deterministic environment. This defines a branching process with interactions  $(\nu_t, t \geq 0)$ , whose law is characterized by the infinitesimal generators  $(\mathcal{J}^t, t \geq 0)$ . For all  $t \geq 0$ ,  $\mathcal{J}^t$  is defined on the set  $\mathcal{D}_J$ , where

$$\mathcal{D}_J := \{H_h \in \mathfrak{B}(\mathbb{R}_+), \exists (h, H) \in \mathcal{C}^1(\mathbb{R}_+, \mathbb{R}) \times \mathcal{C}^1(\mathbb{R}_+, \mathbb{R}_+^*) : \forall \nu \in \mathbb{W} \ H_h(\nu) = H(\langle \nu, h \rangle)\}$$

by

$$\begin{aligned} \mathcal{J}^t H_h(\nu) &= H'(\langle \nu, h \rangle) \int_{\mathbb{R}_+} h'(y) \mu(t) y \nu(dy) \\ &+ \int_{\mathbb{R}_+} r y \int_0^1 [H_h(\nu - \delta_y + \delta_{\lambda y} + \delta_{(1-\lambda)y}) - H_h(\nu)] q(\lambda) d\lambda \nu(dy) \\ &+ d\langle \nu, 1 \rangle \int_{\mathbb{R}_+} \int_0^1 [H_h(\nu - \delta_y + \delta_{\theta y}) - H_h(\nu)] p(\theta) d\theta \nu(dy). \end{aligned}$$

We notice that Assumption 1 is verified by the parameters of this branching process, and thus it does not explode in finite time almost surely. Note that in this population, the dynamics are correlated such that an increase in population size accelerates the rate of loss, while loss events slow the rate of division. Even for such a simple model, finding analytic expression of eigenfunctions for the non-local operator  $\mathcal{G}$  for all  $t$  is a complex task, and the existence of such eigenfunctions is not guaranteed in general, see [10, 18]. Furthermore, the  $\psi$ -spinal process associated with an eigenfunction  $\psi$  might have highly intricate dynamics.

Here we propose a new approach to spinal constructions: we use the change of measure associated with the spinal construction in order to simplify the dynamics within the spine process. For this model, we exhibit an appropriate function  $\psi$  for which the  $\psi$ -spinal process is a Markov process indexed by an independent binary tree where every branch lives for an exponential time of mean 1. We believe that this method can be generalized to different models.

We choose  $\psi \in \mathcal{C}^1(\mathbb{R}_+^* \times \mathbb{R}, \mathbb{R}_+^*)$  and  $\phi \in \mathcal{C}^1(\mathbb{R}_+^*, \mathbb{R})$  such that for all  $(x, y) \in \mathbb{R}_+^* \times \mathbb{R} \times \mathbb{R}_+$

$$\psi(x, y) := x e^{-y} \quad \text{and} \quad \phi(x) := \ln(x r K_{\text{div}}). \quad (4.3)$$

Applied to a spinal state  $(x_e, \chi) \in \mathbb{W}$  where  $\chi := \sum_{u \in \mathbb{G}(t)} \delta_{x^u}$ , this weight function verifies

$$\psi(x_e, \chi) = \frac{x_e}{\prod_{u \in \mathbb{G}(t)} (r K_{\text{div}} x^u)}$$

and ensures Assumption 2. We can now determine the parameters of the spinal process using this function  $\psi$ . The behavior of the traits between branching events remains unchanged compared to the Yule process under consideration. The division events occur at rate 1 for both the spinal and the non-spinal individuals. The random variable  $\widehat{\Lambda}$ , which determines the distribution of mass during division in the spinal construction, has a density function  $\widehat{q}$  given by:

$$\widehat{q}(\lambda) := \frac{1}{K_{\text{div}}} \frac{q(\lambda)}{\lambda(1-\lambda)}. \quad (4.4)$$

As a result, in this  $\psi$ -spinal process, division events are no longer dependent on the size of the individuals.

Individuals outside the spine lose a random fraction  $1 - \widehat{\Theta}$  of their mass at a rate  $\widehat{B}_1(x, \nu) := K_{\text{loss}} d\langle \nu, 1 \rangle$  where  $\widehat{\Theta}$  follows a probability density function  $\widehat{p}$  given by:

$$\widehat{p}(\theta) = \frac{1}{K_{\text{loss}}} \frac{p(\theta)}{\theta}. \quad (4.5)$$

The loss events for the spinal individual follow the same dynamics as those in the Yule process being considered. It is worth noting that Assumption 3 holds in this case, leading to  $\widehat{T}_{\text{Exp}} = \infty$  almost surely. Therefore, by using appropriate function  $\psi$  in the spinal construction, we can make division events independent of loss events. In this case, the spinal process verifies the branching property, moreover it is a piece-wise deterministic Markov process with a distinguished individual indexed on a binary tree with unit branching rate. This  $\psi$ -spinal process falls within the scope of [5], and the limits theorems within it apply to this process. However, concluding on the long time behavior of the original branching process with interactions is not direct as Theorem 3.4 does not apply for this chosen function  $\psi$



In the following we show that we can use this property to enhance the simulation complexity of the branching process with interactions. A classical exact method to simulate non-homogeneous Poisson processes is the thinning algorithm, introduced by Lewis and Shedler [40]. It is used to simulate Poisson processes of intensity  $c(t)$  on a window  $[0, T]$  for a fixed  $T > 0$ . The idea is to generate possible jump times  $(t_i, 1 \leq i \leq n)$  at a rate  $\bar{c} := \sup_{[0, T]} c(t)$  and accept them with probability  $c(t_i)/\bar{c}$ . When the intensity  $c(\cdot)$  depends not only on time  $t$  but also on the entire past of the point process, one can use Ogata's modified thinning algorithm [51]. Given the information of the first  $k$  points,  $(t_i, 1 \leq i \leq k)$  the intensity  $c(\cdot)$  is deterministic on  $[t_k, T_{k+1}]$  with  $T_{k+1}$  the next time of jump. As a result, generating the next point in such processes can be considered as generating the first point in an inhomogeneous Poisson process. This idea has been more recently adapted for branching process, see *e.g.* [25, 11, 26]. The main limitation of this method is that  $\bar{c}$  can become excessively large even for small simulation windows  $T$ . This results in the rejection of most of the generated points. Another exact method, based on inverse transform sampling, consists in generating the arrival times of the process by sampling a uniform random variable  $U$  in  $[0, 1]$ , see *e.g.* [20]. The arrival times  $t_k$  are thus generated by inversion of the cumulative distribution function of the jumps times, such that  $1 - \exp(-\int_0^{t_k} c(s) ds) = U$ . However, an exact inversion is inaccessible in general cases, and in this model in particular.

Here, we propose a new simulation method, based on the spinal construction, that can be much faster than the Ogata's algorithm. The idea is to use the fact that the division events are independent from the mass in the spine process. Thus, we can first generate a binary tree with unit rate, then draw a realization of the distribution of masses at division and choose the spinal individual, and finally spawn a Poisson point process indexed on this tree that distributes the loss events. These three steps are illustrated in Figure 1. At last, the trait of each individual at every time is computed using the deterministic growth and the encountered loss events. A general detailed algorithm can be found in Appendix C.

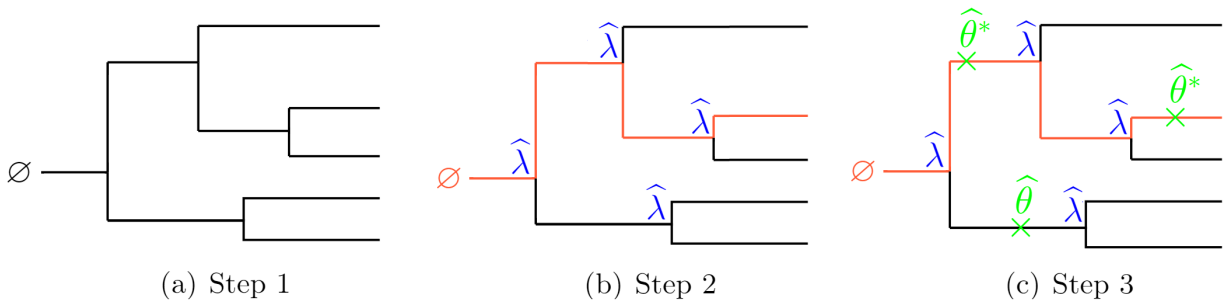


FIGURE 1. Algorithm description

We then use Theorem 2.1 to compute various statistics of the branching process based on the statistics obtained from the spine process. For all  $(x, \nu) \in \mathbb{R}_+^* \times \mathbb{V}$ , the expression of the operator  $\mathcal{G}$  applied to the chosen  $\psi$ , defined in (4.3), is given by

$$\frac{\mathcal{G}\psi(x, \nu)}{\psi(x, \nu)} = \mu + (1 - \mu + d(N - 1)(K_{\text{loss}} - 1))N - rB,$$

where for all  $\nu \in \mathbb{V}$ ,  $N := \langle \nu, 1 \rangle$  and  $B := \int_{\mathbb{R}_+^*} x\nu(dx)$  denote respectively the size of the population  $\nu$  and its total biomass.

**Algorithm efficiency.** In order to compare Ogata's method and our spinal method we fix the parameters  $r, d$  and  $\mu$ . The random variables  $\Theta$  and  $\Lambda$  follow beta laws, respectively  $\beta(a, b)$  and  $\beta(\alpha, \alpha)$ . We will take the values  $a = 10, b = 2$  that correspond to small losses and  $\alpha = 20$  to get a narrow symmetrical distribution at division. With these distributions, the parameters

introduced in (4.1) and (4.2) become

$$m_{\text{div}} = \frac{1}{2}, \quad K_{\text{div}} = 4 + \frac{2}{\alpha - 1}, \quad m_{\text{loss}} = \frac{a}{a + b} \quad \text{and} \quad K_{\text{loss}} = \frac{a + b - 1}{a - 1}.$$

We construct two spinal processes using  $\psi_1$  and  $\psi_2$  such that, for every state  $(x_e, \chi) \in \mathbb{W}$  where  $\chi := \sum_{u \in \widehat{\mathbb{G}}(t)} \delta_{x^u}$ ,

$$\psi_1(x_e, \chi) := \frac{x_e}{\prod_{u \in \widehat{\mathbb{G}}(t)} (r K_{\text{div}} x^u)}, \quad \text{and} \quad \psi_2(x_e, \chi) := \frac{x_e}{\prod_{u \in \widehat{\mathbb{G}}(t)} (K_{\text{div}} x^u)}.$$

The dynamics of both spinal processes constructed with these functions  $\psi$  are summarized in Figure 2 where we use  $r_\psi := \mathbb{1}_{\{\psi=\psi_1\}} + r \mathbb{1}_{\{\psi=\psi_2\}}$ .

	division rate	loss rate	distribution at division	distribution at loss
Spine	$r_\psi$	$d\widehat{N}_t$	$\widehat{\Lambda}^* \sim \beta(\alpha - 1, \alpha - 1)$	$\widehat{\Theta}^* \sim \beta(a, b)$
Non-spinal	$r_\psi$	$K_{\text{loss}} d\widehat{N}_t$	$\widehat{\Lambda} \sim \beta(\alpha - 1, \alpha - 1)$	$\widehat{\Theta} \sim \beta(a - 1, b)$

FIGURE 2. Dynamics of the  $\psi$ -spinal processes

In order to benchmark these two spinal methods and the Ogata's method, we estimate the mean size  $\bar{x}$  of an individual picked at random in the population at time  $t$ , starting from a population  $\bar{z}$  of size  $N_0 := \langle z, 1 \rangle$  and total mass  $B_0 := \int xz(dx)$ . For  $i \in \{1, 2\}$ , the  $\psi_i$ -spinal processes at any time  $t$  are denoted by  $\chi_t^i$  and we denote by  $\widehat{N}_t^i := \langle \chi_t^i, 1 \rangle$  and  $\widehat{B}_t^i := \int x\chi_t^i(dx)$  the size and total mass of the population  $\widehat{\mathbb{G}}_t^i$  at any time  $t$ . We will also use the following notation for any  $t$

$$\Pi_t^i := \prod_{u \in \widehat{\mathbb{G}}_t^i} X_t^u.$$

Theorem 2.2 applied with the  $\psi$ -spinal constructions gives

$$\bar{x} := \frac{B_0 e^{\mu t}}{(r_{\psi_i} K_{\text{div}})^{N_0}} \mathbb{E}_{\bar{z}} \left[ \frac{(r_{\psi_i} K_{\text{div}})^{\widehat{N}_t^i}}{\widehat{N}_t^i} \exp \left( C_1^i \int_0^t \widehat{N}_s^i ds + C_2 \int_0^t (\widehat{N}_s^i)^2 ds - r \int_0^t \widehat{B}_s^i ds \right) \frac{\Pi_t^i}{\Pi_0} \right]$$

where

$$C_1^i := r_{\psi_i} - \mu - d(K_{\text{loss}} - 1), \quad \text{and} \quad C_2 := d(K_{\text{loss}} - 1).$$

We developed python functions, based on the algorithm detailed in Appendix C, to generate these trajectories. In this case, the difficulty in the implementation lies in optimizing the computation of the integral terms. We present in Appendix C the formula used to compute these integral terms. Figure 3 compares the running times  $T_S$  of the two spinal methods with the running time  $T_O$  of the Ogata's method for the estimation of  $\bar{x}$  with 100.000 sample paths for different sets of parameters. At each line of Figure 3, we change one parameter at a time and the remaining parameters are fixed at a base value that changes at each row. The code used to generate this figure is available in a github repository [47].

Notice that the spine methods complexity does not depend on the individual traits, unlike Ogata's complexity that grows exponentially with the biomass. This point explains the exponential difference in complexity with the parameters  $x_0$  and  $\mu$ . The evolution of the efficiency with the initial size is due to the fact that the division rate in the spine method does not depend on the size of the population. Also the division rate in the first spinal method is independent of  $r$ , explaining the evolution of the efficiency with the division parameter  $r$ . However this is not the case for the second spinal method that has a more intricate dependency on  $r$ . Finally, notice that

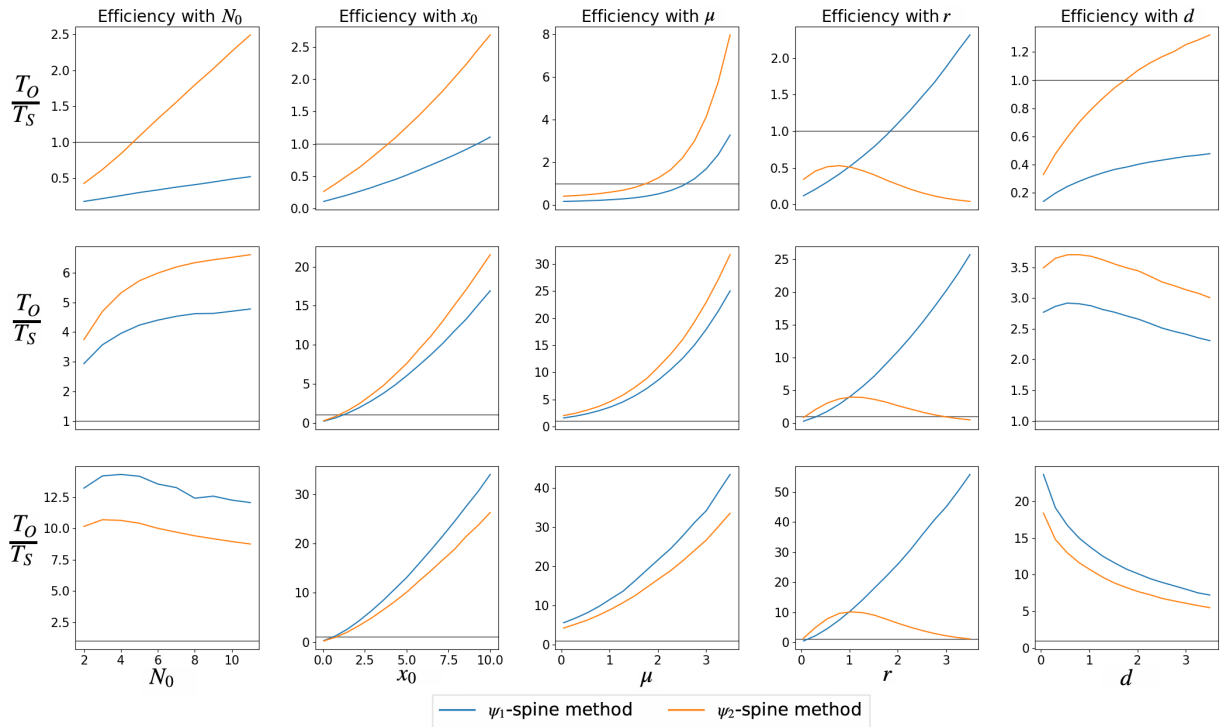


FIGURE 3. Running time comparison

Running time for generating 50000 exact trajectories of the population with the spinal method, normalized by the running time of Ogata's method. At each column, one parameter is modified and at each line, we change the parameters base values: from the first to the third line,  $(N_0, x_0, \mu, r, d)$  are respectively  $(2, 0.5, 0.4, 0.4, 0.4)$ ,  $(2, 3, 0.8, 0.8, 0.8)$  and  $(2, 6, 1.2, 1.2, 1.2)$ . The grey lines indicate a comparable efficiency between spinal and Ogata's methods.

the complexities of the spine methods with the loss parameter  $d$  are linear, unlike the complexity of the Ogata's method. In fact, the number of generated Poisson points increases with  $d$  but the probability of rejection of those points decreases with  $d$ , resulting in a complex dependency with the parameter  $d$ . Depending on the parameter values, using a spine method to sampling paths instead of the classical Ogata's one, can be exponentially faster.

## 5. PROOFS

**5.1. Proof of Section 1.** In this section, we derive the result on the existence and uniqueness of the considered branching process.

*Proof of Proposition 1.1.* The description of the process  $(\bar{\nu}_t, t \geq 0)$  introduced in Section 1 leads to a canonical SDE driven by the dynamics of the trait and a multivariate point process. This SDE is then used to show that the mean number of individuals in the population and the trait are bounded at any time  $t \geq 0$ . Using Assumption 1.2 we conclude that there is no explosion in finite time of the population, and thus, following the proof in [25], it is sufficient to establish uniqueness and existence of a solution of this SDE by construction.

We recall that the offspring traits at birth  $\mathbf{y} = (y^1, \dots, y^n)$  of an individual of trait  $x$  in a population  $\nu$  is given by the law  $K_n(x, \nu, \mathbf{y})\mathcal{M}_n(d\mathbf{y})$ . A classical way to define the associated SDE is to assume that the random traits depend only on a uniform random variable on  $[0, 1]$ , see [5, 46]. However this method correlates the trait at birth between every child. To solve this

issue, we will introduce the measure  $\mathcal{M}$  on  $\mathbb{N} \times \mathcal{X}^{\mathbb{N}}$  such that

$$\mathcal{M}(dn, d\mathbf{y}) := \sum_{i \geq 1} \delta_i(dn) \mathcal{M}_i(d\mathbf{y}).$$

Let  $E = \mathbb{R}_+ \times \mathcal{U} \times \mathbb{R}_+ \times \mathbb{N} \times \mathcal{X}^{\mathbb{N}}$  and  $Q(ds, du, dr, dn, d\mathbf{y})$  be a Poisson point measure on  $E$  with intensity  $q$  such that

$$q(ds, du, dr, dn, d\mathbf{y}) = ds \left( \sum_{v \in \mathcal{U}} \delta_v(du) \right) dr \mathcal{M}(dn, d\mathbf{y}).$$

We denote by  $(\mathcal{F}_t, t \geq 0)$  the canonical filtration associated with this Poisson point measure and  $(T_k, k \geq 0)$  the sequence of jumps times, that is the random sequence of times of arrivals given by the Poisson point measure  $Q$ .

Let  $\bar{z} \in \bar{\mathbb{V}}$ . For every function  $g \in \mathcal{C}^1(\mathcal{U} \times \mathcal{X}, \mathbb{R})$  and  $t \geq 0$ , the process  $(\bar{v}_t, t \geq 0)$  starting from  $\bar{z}$  verifies

$$\begin{aligned} \langle \bar{v}_t, g \rangle &:= \langle \bar{z}, g \rangle + \int_0^t \int_{\mathcal{U} \times \mathcal{X}} \frac{\partial g}{\partial x}(u, x) \cdot \mu(x, \nu_s) \bar{v}_s(du, dx) ds \\ &+ \int_E \mathbb{1}_{\{u \in \mathbb{G}(\bar{v}_{s^-})\}} \mathbb{1}_{\{r \leq B_n(X_{s^-}^u, \nu_{s^-}) K_n(X_{s^-}^u, \nu_{s^-}, \mathbf{y})\}} \left( \sum_{i=1}^n g(ui, \mathbf{y}^i) - g(u, X_{s^-}^u) \right) \\ &\quad \times Q(ds, du, dr, dn, d\mathbf{y}). \end{aligned} \quad (5.1)$$

The first line in this stochastic differential equation corresponds to the deterministic evolution of the trait between the branching events. The second line describes the branching events where  $u$  is the label of the branching individual,  $n$  is the number of children and  $\mathbf{y}$  their traits at birth.

Note that the dynamical construction of the marginal measure-valued process only does not ensure uniqueness: the individual for the next branching event is chosen according to its trait, and thus two individuals with the same trait can be indistinctly chosen to be the branching one. The labeling of individuals allows us to overcome this problem. However, any other labeling method could work, see *e.g.* [25]. Here we consider non-bounded rates and we follow with small adaptations the proof of Lemma 2.5 in [46]. Let  $T > 0$ . We prove the non-accumulation of branching events on  $[0, T]$ . First we use equation (5.1) applied to the constant function equal to 1, that gives the number of individuals in the population, denoted  $(N_t, t \geq 0)$ . Using Assumption 1.3, we have for all  $t < T_k \wedge T$

$$\mathbb{E}_{\bar{z}}[N_t] \leq N_0 + \int_0^t b_0 \left( \mathbb{E}_{\bar{z}} \left[ \sum_{u \in \mathbb{G}(t)} |X_s^u| \right] + \mathbb{E}_{\bar{z}}[N_s] \right) ds.$$

Next, we take a sequence of functions  $(g_n, n \in \mathbb{N})$  where for all  $n$ ,  $g_n \in \mathcal{C}^1(\mathcal{U} \times \mathcal{X}, \mathbb{R}_+)$  and  $\lim_{n \rightarrow \infty} g_n(u, x) = |x|$  and  $\lim_{n \rightarrow \infty} \frac{\partial g_n}{\partial x}(u, x) \cdot \mu(x, \nu) \leq |\mu(x, \nu)|$  for all  $\nu$ . Applying equation (5.1) to these functions and using Assumptions 1.1 and 1.4, we have when  $n \rightarrow +\infty$ ,

$$\mathbb{E}_{\bar{z}} \left[ \sum_{u \in \mathbb{G}(t)} |X_t^u| \right] \leq \langle \bar{z}, |\cdot| \rangle + \int_0^t \mu_0 \left( \mathbb{E}_{\bar{z}} \left[ \sum_{u \in \mathbb{G}(s)} |X_s^u| \right] + \mathbb{E}_{\bar{z}} \left[ \sum_{u \in \mathbb{G}(s)} X_s^u \right] + \mathbb{E}_{\bar{z}}[N_s] \right) ds.$$

According to Grönwall's lemma, for all  $t < T_k \wedge T$ , we have

$$\mathbb{E}_{\bar{z}}[N_t] + \mathbb{E}_{\bar{z}} \left[ \sum_{u \in \mathbb{G}(t)} |X_t^u| \right] \leq (N_0 + \langle \bar{z}, |\cdot| \rangle) e^{(b_0 + \mu_0)t} < \infty.$$

The number of individuals is thus almost surely finite at finite time as well as the trait of every individual. Assumption 1.2 ensures that in finite time, there is no accumulation of branching events that does not change the size of the population.  $\square$

**5.2. Proof of Section 2.** Theorem 2.1 is proved following the steps of the proof of Theorem 1 in [3]. Let  $\psi \in \mathcal{D}$  satisfying Assumption 2, and  $((E_t, \bar{x}_t), t \geq 0)$  be the  $\bar{\mathbb{W}}$ -valued branching process of generator  $\hat{\mathcal{L}}_\psi$ , defined in (2.4). We will first need to introduce some notations. We will denote  $(U_k, k \geq 0)$  the sequence of  $\mathcal{U}$ -valued random variables giving the label of the branching individuals at the jumps times  $(T_k, k \geq 0)$ . Let  $(N_k, k \geq 0)$  be the sequence of  $\mathbb{N}$ -valued random variables giving the number of children at each branching event and we denote for brevity  $A_k := (U_k, N_k)$  for all  $k \geq 0$ . At the  $k$ -th branching time  $T_k$ , we denote  $\mathcal{Y}_k$  the  $\mathcal{X}^{N_k}$ -valued random variable giving the vector of offspring traits. Finally we introduce for all  $k \geq 0$ ,  $\mathcal{V}_k := (T_k, A_k, \mathcal{Y}_k)$ . We similarly define  $((\hat{T}_k, \hat{U}_k, \hat{N}_k, \hat{\mathcal{Y}}_k), k \geq 0)$ , the sequence of jumps times, labels of the branching individual, number of children and trait of these children at birth in the spinal construction. Notice that the distribution of number of children and traits depend on whether the branching individual is the spinal one or not. We will also use for all  $k \geq 0$ ,  $\hat{\mathcal{V}}_k := (\hat{T}_k, \hat{A}_k, \hat{\mathcal{Y}}_k)$  where  $\hat{A}_k := (\hat{U}_k, \hat{N}_k)$ . At time  $s \in [\hat{T}_{k-1}, \hat{T}_k)$ , the label of the spinal individual is denoted  $E_k$  and its trait  $Y_k$ . For a given initial population  $\bar{z} = \sum_{i=1}^n \delta_{(i, x^i)} \in \bar{\mathbb{V}}$ , we use by convention  $U_0 = \emptyset$ ,  $N_0 = n$ ,  $\mathcal{Y}_0 = (x^i, 1 \leq i \leq n)$  almost surely. The same convention holds for the spine process. For all  $0 \leq k$ , we introduce the associated filtrations

$$\mathcal{F}_k = \sigma(\mathcal{V}_i, 0 \leq i \leq k), \text{ and } \hat{\mathcal{F}}_k = \sigma(E_k, (\hat{\mathcal{V}}_i, 0 \leq i \leq k)).$$

Notice that these notations, summarized in Figure 4, are well-defined until the explosion time of the branching and spine processes and that the vectors  $(\mathcal{V}_k, k \geq 0)$  where for all  $k \geq 0$ ,  $\mathcal{V}_k := (T_k, A_k, \mathcal{Y}_k)$  characterize the trajectories of  $(\bar{\nu}_t, t \geq 0)$  until the explosion time.

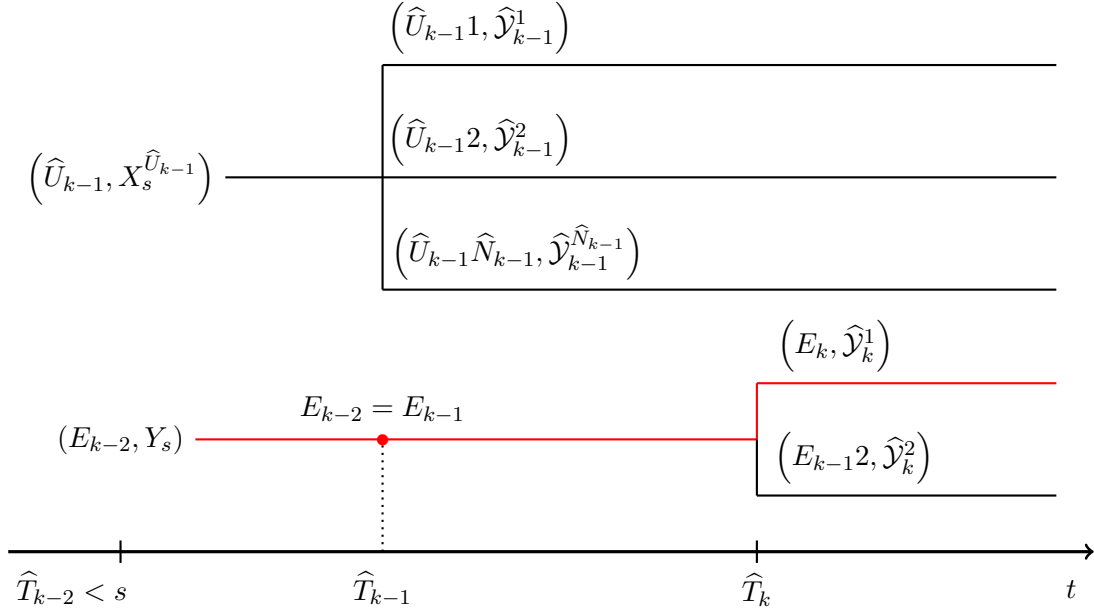


FIGURE 4. Sequential notations for the spine process.

For every initial population  $\bar{z} \in \bar{\mathbb{V}}$  and every  $k \geq 0$ , we introduce the set of sequences of  $k$  branching events that lead to non-extinguished trajectories  $\mathfrak{U}_k(\bar{z}) \subset (\mathcal{U} \times \mathbb{N})^k$ , starting from  $\bar{z}$ . We also introduce for all  $a \in \mathfrak{U}_k(\bar{z})$  and all  $0 \leq i \leq k$ ,  $\mathbb{G}_i(a)$  the set of labels of individuals living between the  $i$ -th and  $(i+1)$ -th branching events in the population where all the branching events were given by  $a$ . By decomposing the branching process  $(\bar{\nu}_t \geq 0)$  based on the sequences  $a \in \mathfrak{U}_k(\bar{z})$  and sampling at time  $t$  an individual  $e \in \mathbb{G}_k(a)$ , we get that for every non-negative

measurable function  $H$  on  $\mathcal{U} \times \mathbb{D}([0, t], \mathbb{V})$

$$\begin{aligned} \mathbb{E}_{\bar{z}} \left[ \mathbb{1}_{\{T_{\text{Exp}} > t, \mathbb{G}(t) \neq \emptyset\}} H(U_t, (\bar{\nu}_s, s \leq t)) \right] = \\ \sum_{k \geq 0} \sum_{a \in \mathfrak{A}_k(\bar{z})} \sum_{e \in \mathbb{G}_k(a)} \mathbb{E}_{\bar{z}} \left[ H(e, (\bar{\nu}_s, s \leq t)) p_e(\bar{\nu}_t) \mathbb{1}_{\{T_k \leq t < T_{k+1}\}} \prod_{i=0}^k \mathbb{1}_{\{A_i = a_i\}} \right]. \end{aligned} \quad (5.2)$$

The expectation on the right-hand side of (5.2) is linked by a Girsanov-type result to the spinal construction, as shown in Lemma 5.1. The difference with our proof and the proof of Theorem 1 in [3] lies among other things in the demonstration of this lemma.

**Lemma 5.1.**

For any  $k > 0$ ,  $\bar{z} \in \bar{\mathbb{V}}$  and  $a = ((u_i, n_i), 0 \leq i \leq k) \in \mathfrak{A}_k(\bar{z})$ , let  $F$  be a non-negative measurable function on  $\Pi_{i=1}^k (\mathbb{R}_+ \times \mathcal{U} \times \mathbb{N} \times \mathcal{X}^{n_i})$ . For any  $e \in \mathbb{G}_k(a)$ ,

$$\begin{aligned} \mathbb{E}_{\bar{z}} \left[ F(\mathcal{V}_i, 0 \leq i \leq k) \prod_{i=0}^k \mathbb{1}_{\{A_i = a_i\}} \right] = \langle z, \psi(\cdot, z, 0) \rangle \\ \times \mathbb{E}_{\bar{z}} \left[ \xi_k \left( E_k, (\widehat{\mathcal{V}}_i, 0 \leq i \leq k) \right) F \left( \widehat{\mathcal{V}}_i, 0 \leq i \leq k \right) \mathbb{1}_{\{E_k = e\}} \prod_{i=0}^k \mathbb{1}_{\{\widehat{A}_i = a_i\}} \right], \end{aligned}$$

where

$$\xi_k \left( E_k, (\widehat{\mathcal{V}}_i, 0 \leq i \leq k) \right) := \frac{1}{\psi \left( Y_{\widehat{T}_k}, \chi_{\widehat{T}_k} \right)} \exp \left( \int_0^{\widehat{T}_k} \frac{\mathcal{G}\psi(Y_s, \chi_s)}{\psi(Y_s, \chi_s)} ds \right).$$

We prove this lemma by induction on the number of branching events  $k$ . We will first state a technical lemma to avoid too long computations.

**Lemma 5.2.** For all  $k > 0$ ,  $\bar{z} \in \bar{\mathbb{V}}$  and  $a = ((u_i, n_i), 0 \leq i \leq k) \in \mathfrak{A}_k(\bar{z})$ ,

$$\frac{\xi_k \left( E_k, (\widehat{\mathcal{V}}_i, 0 \leq i \leq k) \right)}{\xi_{k-1} \left( E_{k-1}, (\widehat{\mathcal{V}}_i, 0 \leq i \leq k-1) \right)} = \frac{\psi \left( Y_{\widehat{T}_k}, \chi_{\widehat{T}_k} \right)}{\psi \left( Y_{\widehat{T}_{k-1}}, \chi_{\widehat{T}_{k-1}} \right)} \exp \left( \int_{\widehat{T}_{k-1}}^{\widehat{T}_k} \lambda(Y_s, \chi_s) ds \right) \quad a.s.$$

where, for all  $s \in [\widehat{T}_{k-1}, \widehat{T}_k)$ ,

$$\lambda(Y_s, \chi_s) := \widehat{\tau}_{\text{tot}}(Y_s, \chi_s) - \tau(\chi_s), \quad \text{with } \tau(\chi_s) := \int_{\mathcal{X}} B(x, \chi_s) \chi_s(dx). \quad (5.3)$$

We recall that  $\widehat{\tau}_{\text{tot}}$  is the total branching rate of the spine process defined in (2.10).

*Proof of Lemma 5.2.* The recursive equality for  $\xi_k$  derives from the fundamental theorem of calculus. Let  $k \geq 1$ ,

$$\frac{\xi_k \left( E_k, (\widehat{\mathcal{V}}_i, 0 \leq i \leq k) \right)}{\xi_{k-1} \left( E_{k-1}, (\widehat{\mathcal{V}}_i, 0 \leq i \leq k-1) \right)} = \frac{\psi \left( Y_{\widehat{T}_{k-1}}, \chi_{\widehat{T}_{k-1}} \right)}{\psi \left( Y_{\widehat{T}_k}, \chi_{\widehat{T}_k} \right)} \exp \left( \int_{\widehat{T}_{k-1}}^{\widehat{T}_k} \frac{\mathcal{G}\psi(Y_s, \chi_s)}{\psi(Y_s, \chi_s)} ds \right).$$

We notice that  $\mathcal{G}$  introduced in (2.3), verify for all  $(x_e, \nu) \in \mathbb{W}$

$$\frac{[\mathcal{G} - G] \psi(x_e, \nu)}{\psi(x_e, \nu)} = \widehat{\tau}_{\text{tot}}(x_e, \nu) - \int_{\mathcal{X}} B(x, \nu) \nu(dx),$$

where  $\widehat{\tau}_{\text{tot}}$  is defined in (2.10). Then, using derivative chain rule we get that

$$G(\ln \circ \psi)(x_e, \nu) = \frac{G\psi(x_e, \nu)}{\psi(x_e, \nu)}.$$

Between successive branching events, the evolution of the process is purely deterministic and we can apply the fundamental theorem of calculus:

$$\int_{\widehat{T}_{k-1}}^{\widehat{T}_k} G(\ln \circ \psi)(Y_s, \chi_s) ds = \ln \left( \psi \left( Y_{\widehat{T}_k^-}, \chi_{\widehat{T}_k^-} \right) \right) - \ln \left( \psi \left( Y_{\widehat{T}_{k-1}}, \chi_{\widehat{T}_{k-1}} \right) \right).$$

We thus have

$$\exp \left( \int_{\widehat{T}_{k-1}}^{\widehat{T}_k} \frac{\mathcal{G}\psi(Y_s, \chi_s)}{\psi(Y_s, \chi_s)} ds \right) = \frac{\psi \left( Y_{\widehat{T}_k^-}, \chi_{\widehat{T}_k^-} \right)}{\psi \left( Y_{\widehat{T}_{k-1}}, \chi_{\widehat{T}_{k-1}} \right)} \exp \left( \int_{\widehat{T}_{k-1}}^{\widehat{T}_k} \lambda(Y_s, \chi_s) ds \right),$$

that concludes the proof.  $\square$

This lemma establishes that the exponential path-integral term is the suitable change of measure to compensate for the deterministic dynamics between jumps.

*Proof of Lemma 5.1.* Following the proof of Theorem 1 in [3], the result is established by induction on the number  $k$  of branching events. The original branching process and its associated spine process may stop branching in finite time, in this case the total numbers of branching events, respectively  $N_{\text{tot}}$  and  $\widehat{N}_{\text{tot}}$ , are finite. For all  $k$  such that  $k > N_{\text{tot}}$ , the  $k$ -th branching event of the original construction arrives at  $T_k = +\infty$ . In this case we set  $A_k = (\emptyset, 0)$  by convention. The same convention is used for the spinal construction. Let  $\bar{z} := \sum_{i=1}^n \delta_{(i, x^i)} \in \overline{\mathbb{V}}$  be the initial population and let  $F$  be a non-negative measurable function on  $\mathbb{R}_+ \times \mathcal{U} \times \mathbb{N} \times \mathcal{X}^n$ . Then, by definition,  $\mathcal{V}_0 = \widehat{\mathcal{V}}_0 = (0, (\emptyset, n), (x^i, 1 \leq i \leq n))$  almost surely.

Therefore, for all  $e \in \mathbb{G}(0)$

$$\mathbb{E}_{\bar{z}} \left[ \xi_0 \left( E_0, \widehat{\mathcal{V}}_0 \right) F \left( \widehat{\mathcal{V}}_0 \right) \mathbb{1}_{\{E_0=e\}} \mathbb{1}_{\{\widehat{A}_0=(\emptyset, n)\}} \right] = F(\mathcal{V}_0) \mathbb{1}_{\{A_0=(\emptyset, n)\}} \mathbb{E}_{\bar{z}} \left[ \frac{\mathbb{1}_{\{E_0=e\}}}{\psi(Y_0, z)} \right].$$

We recall that the individual  $1 \leq i \leq n$  of trait  $x^i$  in the population  $\bar{z}$  is chosen to be the spinal individual with probability  $\psi(x^i, z) \langle z, \psi(\cdot, z) \rangle^{-1}$ . Then we have

$$\mathbb{E}_{\bar{z}} \left[ \xi_0 \left( E_0, \widehat{\mathcal{V}}_0 \right) F \left( \widehat{\mathcal{V}}_0 \right) \mathbb{1}_{\{E_0=e\}} \mathbb{1}_{\{\widehat{A}_0=(\emptyset, n)\}} \right] = \frac{1}{\langle z, \psi(\cdot, z) \rangle} \mathbb{E}_{\bar{z}} \left[ \mathbb{1}_{\{A_0=(\emptyset, n)\}} F(0, A_0, \mathcal{Y}_0) \right].$$

Thus the result holds for  $k = 0$ . Now let  $k \geq 1$  and assume that the following induction hypothesis holds at rank  $k - 1$ .

**Induction Hypothesis.** For every  $a = (a_i, 0 \leq i \leq k - 1) \in (\mathcal{U} \times \mathbb{N})^{k-1}$  with  $a_i = (u_i, n_i)$ , every non-negative measurable function  $H$  on  $\bigotimes_{i=1}^{k-1} (\mathbb{R}_+ \times \mathcal{U} \times \mathbb{N} \times \mathcal{X}^{n_i})$  and every label  $e \in \mathbb{G}_{k-1}(a)$ :

$$\begin{aligned} & \mathbb{E}_{\bar{z}} \left[ H(\mathcal{V}_i, 0 \leq i \leq k - 1) \prod_{i=0}^{k-1} \mathbb{1}_{\{A_i=a_i\}} \right] = \langle z, \psi(\cdot, z) \rangle \\ & \times \mathbb{E}_{\bar{z}} \left[ \xi_{k-1} \left( E_{k-1}, \left( \widehat{\mathcal{V}}_i, 0 \leq i \leq k - 1 \right) \right) H \left( \widehat{\mathcal{V}}_i, 0 \leq i \leq k - 1 \right) \mathbb{1}_{\{E_{k-1}=e\}} \prod_{i=0}^{k-1} \mathbb{1}_{\{A_i=a_i\}} \right]. \end{aligned} \quad (5.4)$$

Let  $a = (a_i, 0 \leq i \leq k) \in (\mathcal{U} \times \mathbb{N})^n$  with  $a_i = (u_i, n_i)$  and  $e \in \mathbb{G}_k(a)$ . We denote  $a' = (a_i, 0 \leq i \leq k - 1)$  and take  $F_k^a$  a non-negative measurable function on  $\bigotimes_{i=1}^k (\mathbb{R}_+ \times \mathcal{X}^{n_i})$  such that, for all  $((t_i, y_i), 0 \leq i \leq k) \in \bigotimes_{i=1}^n (\mathbb{R}_+ \times \mathcal{X}^{n_i})$ :

$$F_k^a((t_i, y_i), 0 \leq i \leq n) := F_{k-1}^{a'}((t_i, y_i), 0 \leq i \leq k - 1) I(t_k - t_{k-1}) F(y_k), \quad (5.5)$$

where  $F_{k-1}^{a'}$ ,  $I$  and  $F$  are non-negative measurable functions, respectively on  $\bigotimes_{i=1}^{k-1} (\mathbb{R}_+ \times \mathcal{X}^{n_i})$ ,  $\mathbb{R}_+$  and  $\mathcal{X}^{n_k}$ . We denote by  $e' \in \mathbb{G}_{k-1}(a')$  the deterministic label of the ancestor of  $e$  just before the  $k$ -th jump. Thus there exists a unique  $j \in \mathbb{N}^* \cup \{\emptyset\}$  satisfying  $e'j = e$ . If  $j \neq \emptyset$ , it means that  $e'$  is the branching individual in the  $k$ -th jump, otherwise it is any other individual in the population. We introduce the  $\mathbb{N}^* \cup \{\emptyset\}$ -valued random variable  $J_k$  choosing the label of the spinal individual at the  $k$ -th branching event, so that  $E_k = E_{k-1}J_k$  almost surely.

Following the proof of Lemma 1 in [3], we express both sides of the equality (5.4) for  $k \geq 1$  conditionally on the filtrations at the previous step  $k-1$ . We recall that, for all  $1 \leq j \leq k$

$$\mathcal{F}_j = \sigma(\mathcal{V}_i, 0 \leq i \leq j), \text{ and } \widehat{\mathcal{F}}_j = \sigma\left(E_j, \left(\widehat{\mathcal{V}}_i, 0 \leq i \leq j\right)\right).$$

The following proposition is crucial for applying the induction hypothesis and concluding the proof. We present the proof of this proposition separately, as it forms the core of the change of measure between both processes.

**Proposition 5.3.** *With the previously introduced notations, we define the deterministic functions of  $\widehat{\mathcal{F}}_{k-1}$ -measurable random variables*

$$C\left(\widehat{\mathcal{V}}_i, 0 \leq i \leq k-1\right) := \mathbb{E}\left[I\left(T_k - \widehat{T}_{k-1}\right) F\left(\mathcal{Y}_k\right) \mathbb{1}_{\{A_k = a_k\}} \middle| \widehat{\mathcal{F}}_{k-1}\right] \quad (5.6)$$

and

$$\begin{aligned} \widehat{C}\left(e', \left(\widehat{\mathcal{V}}_i, 0 \leq i \leq k-1\right)\right) &:= \mathbb{1}_{\{E_{k-1} = e'\}} \mathbb{E}\left[\frac{\xi_k\left(E_k, \left(\widehat{\mathcal{V}}_i, 0 \leq i \leq k-1\right)\right)}{\xi_{k-1}\left(E_{k-1}, \left(\widehat{\mathcal{V}}_i, 0 \leq i \leq k-1\right)\right)} \mathbb{1}_{\{J_k = j\}}\right. \\ &\quad \left. \times I\left(\widehat{T}_k - \widehat{T}_{k-1}\right) F\left(\widehat{\mathcal{Y}}_k\right) \mathbb{1}_{\{\widehat{A}_k = a_k\}} \middle| \widehat{\mathcal{F}}_{k-1}\right]. \end{aligned}$$

We have that, almost surely on the event that the last spine is chosen

$$\widehat{C}\left(e', \left(\widehat{\mathcal{V}}_i, 0 \leq i \leq k-1\right)\right) = \mathbb{1}_{\{E_{k-1} = e'\}} C\left(\widehat{\mathcal{V}}_i, 0 \leq i \leq k-1\right).$$

Notice that the function  $C$  in (5.6) takes the expectation of the random variable  $\mathcal{V}_k$ , that follows the dynamics of the original process, conditionally on the previous state of the spinal process.

*Proof of Proposition 5.3.* We start by using Lemma 5.2 to get that

$$\begin{aligned} \widehat{C}\left(e', \left(\widehat{\mathcal{V}}_i, 0 \leq i \leq k-1\right)\right) &:= \mathbb{1}_{\{E_{k-1} = e'\}} \mathbb{E}\left[\frac{\psi\left(Y_{\widehat{T}_{k-1}}, \chi_{\widehat{T}_{k-1}}\right)}{\psi\left(Y_{\widehat{T}_k}, \chi_{\widehat{T}_k}\right)} \exp\left(\int_{\widehat{T}_{k-1}}^{\widehat{T}_k} \lambda(Y_s, \chi_s) ds\right)\right. \\ &\quad \left. \times \mathbb{1}_{\{J_k = j\}} I\left(\widehat{T}_k - \widehat{T}_{k-1}\right) F\left(\widehat{\mathcal{Y}}_k\right) \mathbb{1}_{\{\widehat{A}_k = a_k\}} \middle| \widehat{\mathcal{F}}_{k-1}\right]. \end{aligned}$$

Next we condition the right-hand side expression on the next jump time  $\widehat{T}_k$

$$\begin{aligned} \widehat{C}\left(e', \left(\widehat{\mathcal{V}}_i, 0 \leq i \leq k-1\right)\right) &= \mathbb{1}_{\{E_{k-1} = e'\}} \mathbb{E}\left[\exp\left(\int_{\widehat{T}_{k-1}}^{\widehat{T}_k} \lambda(Y_s, \chi_s) ds\right) I\left(\widehat{T}_k - \widehat{T}_{k-1}\right)\right. \\ &\quad \left. \times \psi\left(Y_{\widehat{T}_k}, \chi_{\widehat{T}_k}\right) \mathbb{E}\left[\frac{F\left(\widehat{\mathcal{Y}}_k\right) \mathbb{1}_{\{J_k = j\}} \mathbb{1}_{\{\widehat{A}_k = a_k\}}}{\psi\left(Y_{\widehat{T}_k}, \chi_{\widehat{T}_k}\right)} \middle| \widehat{\mathcal{F}}_{k-1}, \widehat{T}_k\right] \middle| \widehat{\mathcal{F}}_{k-1}\right]. \quad (5.7) \end{aligned}$$



We handle this last equation by complete induction, on the event  $\{j = \emptyset\} \cup \{j \neq \emptyset\}$ , to show that

$$\begin{aligned} \psi \left( Y_{\widehat{T}_k^-}, \chi_{\widehat{T}_k^-} \right) \mathbb{E} \left[ \frac{F \left( \widehat{\mathcal{Y}}_k \right) \mathbb{1}_{\{J_k=j\}} \mathbb{1}_{\{\widehat{A}_k=a_k\}}}{\psi \left( Y_{\widehat{T}_k}, \chi_{\widehat{T}_k} \right)} \middle| \widehat{\mathcal{F}}_{k-1}, \widehat{T}_k \right] = \\ \frac{B_{n_k} \left( X_{\widehat{T}_k^-}^{u_k}, \chi_{\widehat{T}_k^-} \right)}{\widehat{\tau}_{\text{tot}} \left( \chi_{\widehat{T}_k^-} \right)} \int_{\mathcal{X}^{\widehat{N}_k}} F \left( \mathbf{y} \right) K_{\widehat{N}_k} \left( X_{\widehat{T}_k^-}^{\widehat{U}_k}, \chi_{\widehat{T}_k^-}, \mathbf{y} \right) \mathcal{M}_n(d\mathbf{y}). \quad (5.8) \end{aligned}$$

**Branching outside the spine.** If  $j = \emptyset$ , then the branching individual is not the spinal one, and  $e = e'j = e'$ . We follow the same conditioning than for the branching process, and use the fact that on the event  $\{j = \emptyset\}$  the trait of the spinal individual is  $\widehat{\mathcal{F}}_{k-1}$ -measurable. Using the expression of the  $\psi$ -biased distribution  $\widehat{K}$  of  $\widehat{\mathcal{Y}}_k$ , defined in (2.5), we get

$$\begin{aligned} \mathbb{E} \left[ \frac{F \left( \widehat{\mathcal{Y}}_k \right) \mathbb{1}_{\{J_k=\emptyset\}} \mathbb{1}_{\{\widehat{A}_k=a_k\}}}{\psi \left( Y_{\widehat{T}_k}, \chi_{\widehat{T}_k} \right)} \middle| \widehat{\mathcal{F}}_{k-1}, \widehat{T}_k \right] = \mathbb{E} \left[ \mathbb{1}_{\{\widehat{A}_k=a_k\}} \mathbb{1}_{\{J_k=\emptyset\}} \right. \\ \left. \times \int_{\mathcal{X}^{\widehat{N}_k}} \frac{F \left( \mathbf{y} \right)}{\widehat{\Gamma}_{n_k} \left( Y_{\widehat{T}_k^-}, X_{\widehat{T}_k^-}^{u_k}, \chi_{\widehat{T}_k^-} \right)} K_{n_k} \left( X_{\widehat{T}_k^-}^{\widehat{U}_k}, \chi_{\widehat{T}_k^-}, \mathbf{y} \right) \mathcal{M}_n(d\mathbf{y}) \middle| \widehat{\mathcal{F}}_{k-1}, \widehat{T}_k \right]. \end{aligned}$$

We then recall the distribution of  $\widehat{A}_k$  outside the spine, established in Dynamics 2.2 :

$$\mathbb{E} \left[ \mathbb{1}_{\{\widehat{A}_k=a_k\}} \mathbb{1}_{\{J_k=\emptyset\}} \middle| \widehat{\mathcal{F}}_{k-1}, \widehat{T}_k \right] = \frac{\widehat{\Gamma}_{n_k} \left( Y_{\widehat{T}_k^-}, X_{\widehat{T}_k^-}^{u_k}, \chi_{\widehat{T}_k^-} \right) B_{n_k} \left( X_{\widehat{T}_k^-}^{u_k}, \chi_{\widehat{T}_k^-} \right)}{\psi \left( Y_{\widehat{T}_k^-}, \chi_{\widehat{T}_k^-} \right) \widehat{\tau}_{\text{tot}} \left( \chi_{\widehat{T}_k^-} \right)}.$$

This gives (5.8) on the event  $\{j = \emptyset\}$ .

**Spine branching.** We follow the same computations when  $j \neq \emptyset$ , that corresponds to the case when the branching individual is the spinal one, i.e  $u_k = e'$ . In this case, the distribution of  $Y_{\widehat{T}_k}$  now depends on the traits  $\widehat{\mathcal{Y}}_k$ . Thus conditioning on  $\widehat{\mathcal{Y}}_k$  and using the distribution of the next spinal individual, defined in (2.9), we have

$$\begin{aligned} \mathbb{1}_{\{j \neq \emptyset\}} \mathbb{E} \left[ \frac{F \left( \widehat{\mathcal{Y}}_k \right) \mathbb{1}_{\{J_k=j\}} \mathbb{1}_{\{\widehat{A}_k=a_k\}}}{\psi \left( Y_{\widehat{T}_k}, \chi_{\widehat{T}_k} \right)} \middle| \widehat{\mathcal{F}}_{k-1}, \widehat{T}_k \right] = \mathbb{E} \left[ \mathbb{1}_{\{\widehat{A}_k=(e', n_k)\}} \right. \\ \left. \times \mathbb{E} \left[ \frac{F \left( \widehat{\mathcal{Y}}_k \right)}{\sum_{i=1}^{\widehat{N}_k} \psi \left( \widehat{\mathcal{Y}}_k^i, \chi_{\widehat{T}_k^-} - \delta_{Y_{\widehat{T}_k^-}} + \sum_{l=1}^{\widehat{N}_k} \delta_{\widehat{\mathcal{Y}}_k^l} \right)} \middle| \widehat{\mathcal{F}}_{k-1}, \widehat{T}_k, \widehat{A}_k \right] \middle| \widehat{\mathcal{F}}_{k-1}, \widehat{T}_k \right]. \end{aligned}$$

We then use the distribution  $\widehat{K}^*$  of  $\widehat{\mathcal{Y}}_k$ , defined in (2.6) when the branching individual is the spinal one, conditioning on  $\widehat{A}_k$ , defined in Dynamics 2.3.

$$\begin{aligned} \mathbb{1}_{\{\widehat{A}_k=(e', n_k)\}} \mathbb{E} \left[ \frac{F \left( \widehat{\mathcal{Y}}_k \right)}{\sum_{i=1}^{\widehat{N}_k} \psi \left( \widehat{\mathcal{Y}}_k^i, \chi_{\widehat{T}_k^-} - \delta_{Y_{\widehat{T}_k^-}} + \sum_{l=1}^{\widehat{N}_k} \delta_{\widehat{\mathcal{Y}}_k^l} \right)} \middle| \widehat{\mathcal{F}}_{k-1}, \widehat{T}_k, \widehat{A}_k \right] = \\ \mathbb{1}_{\{\widehat{A}_k=(e', n_k)\}} \int_{\mathcal{X}^{n_k}} \frac{F \left( \mathbf{y} \right)}{\widehat{\Gamma}_{n_k}^* \left( Y_{\widehat{T}_k^-}, \chi_{\widehat{T}_k^-} \right)} K_{n_k} \left( Y_{\widehat{T}_k^-}, \chi_{\widehat{T}_k^-}, \mathbf{y} \right) \mathcal{M}_{n_k} (d\mathbf{y}). \end{aligned}$$

Finally we use the distribution of  $\widehat{A}_k$  when the branching individual is the spinal one, defined in Dynamics 2.3.

$$\mathbb{E} \left[ \mathbb{1}_{\{\widehat{A}_k=(e',n_k)\}} \middle| \widehat{\mathcal{F}}_{k-1}, \widehat{T}_k \right] = \frac{\widehat{\Gamma}_{n_k}^* \left( Y_{\widehat{T}_k^-}, \chi_{\widehat{T}_k^-} \right) B_{n_k} \left( Y_{\widehat{T}_k^-}, \chi_{\widehat{T}_k^-} \right)}{\psi \left( Y_{\widehat{T}_k^-}, \chi_{\widehat{T}_k^-} \right) \widehat{\tau}_{\text{tot}} \left( \chi_{\widehat{T}_k^-} \right)}.$$

This gives (5.8) on the event  $\{j \neq \emptyset\}$ .

Now, combining (5.7) and (5.8), we get

$$\begin{aligned} \widehat{C} \left( e', \left( \widehat{\mathcal{V}}_i, 0 \leq i \leq k-1 \right) \right) &= \mathbb{1}_{\{E_{k-1}=e'\}} \mathbb{E} \left[ \exp \left( \int_{\widehat{T}_{k-1}}^{\widehat{T}_k} \lambda(Y_s, \chi_s) ds \right) I \left( \widehat{T}_k - \widehat{T}_{k-1} \right) \right. \\ &\quad \left. \times \frac{B_{n_k} \left( X_{\widehat{T}_k^-}^{u_k}, \chi_{\widehat{T}_k^-} \right)}{\widehat{\tau}_{\text{tot}} \left( \chi_{\widehat{T}_k^-} \right)} \int_{\mathcal{X}^{n_k}} F(\mathbf{y}) K_{n_k} \left( X_{\widehat{T}_k^-}^{u_k}, \chi_{\widehat{T}_k^-}, \mathbf{y} \right) \mathcal{M}_{n_k}(d\mathbf{y}) \middle| \widehat{\mathcal{F}}_{k-1} \right]. \end{aligned}$$

Using that the time between two jump follows an inhomogeneous exponential law of instantaneous rate  $\widehat{\tau}_{\text{tot}}$ , we have

$$\begin{aligned} \widehat{C} \left( e', \left( \widehat{\mathcal{V}}_i, 0 \leq i \leq k-1 \right) \right) &= \mathbb{1}_{E_{k-1}=e'} \int_{\widehat{T}_{k-1}}^{+\infty} \exp \left( \int_{\widehat{T}_{k-1}}^t \lambda(Y_s, \chi_s) - \widehat{\tau}_{\text{tot}}(Y_s, \chi_s) ds \right) \\ &\quad \times I(t - \widehat{T}_{k-1}) B_{n_k} \left( X_t^{u_k}, \chi_t \right) \int_{\mathcal{X}^{n_k}} F(\mathbf{y}) K_{n_k} \left( X_t^{u_k}, \chi_t, \mathbf{y} \right) \mathcal{M}_{n_k}(d\mathbf{y}) dt. \quad (5.9) \end{aligned}$$

Finally, using in (5.9) the fact that  $\lambda$ , defined in (5.3), is the difference of branching rates between the spine process and the original process, we get that

$$\begin{aligned} \widehat{C} \left( e', \left( \widehat{\mathcal{V}}_i, 0 \leq i \leq k-1 \right) \right) &= \mathbb{1}_{E_{k-1}=e'} \int_{\widehat{T}_{k-1}}^{+\infty} \tau(\chi_t) \exp \left( - \int_{\widehat{T}_{k-1}}^t \tau(\chi_s) ds \right) \\ &\quad \times I(t - \widehat{T}_{k-1}) \frac{B_{n_k} \left( X_t^{u_k}, \chi_t \right)}{\tau(\chi_t)} \int_{\mathcal{X}^{n_k}} F(\mathbf{y}) K_{n_k} \left( X_t^{u_k}, \chi_t, \mathbf{y} \right) \mathcal{M}_{n_k}(d\mathbf{y}) dt. \end{aligned}$$

Now, conditioning in chain the expression of  $C$  defined in (5.6), we get

$$C \left( \widehat{\mathcal{V}}_i, 0 \leq i \leq k-1 \right) = \mathbb{E} \left[ I(T_k - \widehat{T}_{k-1}) \mathbb{E} \left[ \mathbb{1}_{\{A_k=a_k\}} F(\mathcal{Y}_k) \middle| \widehat{\mathcal{F}}_{k-1}, T_k \right] \middle| \widehat{\mathcal{F}}_{k-1} \right].$$

Using the conditional distribution of  $\mathcal{Y}_k$  and  $A_k$ , we have

$$\mathbb{E} \left[ \mathbb{1}_{\{A_k=a_k\}} F(\mathcal{Y}_k) \middle| \widehat{\mathcal{F}}_{k-1}, T_k \right] = \frac{B_{n_k} \left( X_{T_k^-}^{u_k}, \chi_{T_k^-} \right)}{\tau \left( \chi_{T_k^-} \right)} \int_{\mathcal{X}^{n_k}} F(\mathbf{y}) K_{n_k} \left( X_{T_k^-}^{u_k}, \chi_{T_k^-}, \mathbf{y} \right) \mathcal{M}_{n_k}(d\mathbf{y}).$$

We conclude the proof using that the time between two jumps follows an inhomogeneous exponential law of instantaneous rate  $\tau(\cdot)$ .  $\square$

Now, using the result of Proposition 5.3, we can proceed with the inductive proof of Lemma 5.1. Using (5.5) and conditioning the  $\mathcal{F}_k$ -measurable sequence of random variables  $(\mathcal{V}_i, 0 \leq i \leq k)$  on the filtration  $\mathcal{F}_{k-1}$ , that contains the information on the original process until just after the

$(k - 1)$ -th branching event, we have

$$\mathbb{E}_{\bar{z}} \left[ F_k^a((T_i, \mathcal{Y}_i), 0 \leq i \leq k) \prod_{i=0}^k \mathbb{1}_{\{A_i=a_i\}} \right] = \mathbb{E}_{\bar{z}} \left[ F_{k-1}^{a'}((T_i, \mathcal{Y}_i), 0 \leq i \leq k-1) C(\mathcal{V}_i, 0 \leq i \leq k-1) \prod_{i=0}^{k-1} \mathbb{1}_{\{A_i=a_i\}} \right], \quad (5.10)$$

where we used the deterministic function  $C$  of  $\mathcal{F}_{k-1}$ -measurable random variables

$$C(\mathcal{V}_i, 0 \leq i \leq k-1) := \mathbb{E} [I(T_k - T_{k-1}) F(\mathcal{Y}_k) \mathbb{1}_{\{A_k=a_k\}} | \mathcal{F}_{k-1}].$$

We introduce the function  $H$  defined by

$$H(\mathcal{V}_i, 0 \leq i \leq k-1) := F_{k-1}^{a'}((T_i, \mathcal{Y}_i), 0 \leq i \leq k-1) C(\mathcal{V}_i, 0 \leq i \leq k-1).$$

Then we can apply the induction hypothesis (5.4) with this function  $H$  in (5.10)

$$\begin{aligned} \mathbb{E}_{\bar{z}} \left[ F_k^a((T_i, \mathcal{Y}_i), 0 \leq i \leq k) \prod_{i=0}^k \mathbb{1}_{\{A_i=a_i\}} \right] &= \langle z, \psi(\cdot, z) \rangle \\ &\times \mathbb{E}_{\bar{z}} \left[ \xi_{k-1} \left( E_{k-1}, (\widehat{\mathcal{V}}_i, 0 \leq i \leq k-1) \right) F_{k-1}^{a'} \left( (\widehat{T}_i, \widehat{\mathcal{Y}}_i), 0 \leq i \leq k-1 \right) \right. \\ &\quad \left. \times C \left( \widehat{\mathcal{V}}_i, 0 \leq i \leq k-1 \right) \mathbb{1}_{\{E_{k-1}=e'\}} \prod_{i=0}^{k-1} \mathbb{1}_{\{A_i=a_i\}} \right]. \end{aligned}$$

Then, using Proposition (5.3), and the fact that  $\{E_k = e\} = \{E_{k-1} = e'\} \cap \{J_k = j\}$ , we finally get that

$$\begin{aligned} \mathbb{E}_{\bar{z}} \left[ F_k^a((T_i, \mathcal{Y}_i), 0 \leq i \leq k) \prod_{i=0}^k \mathbb{1}_{\{A_i=a_i\}} \right] &= \langle z, \psi(\cdot, z) \rangle \\ &\times \mathbb{E}_{\bar{z}} \left[ \xi_k \left( E_k, (\widehat{\mathcal{V}}_i, 0 \leq i \leq k) \right) F_k^a \left( (\widehat{T}_i, \widehat{\mathcal{Y}}_i), 0 \leq i \leq k \right) \mathbb{1}_{\{E_k=e\}} \prod_{i=0}^k \mathbb{1}_{\{\widehat{A}_i=a_i\}} \right]. \end{aligned}$$

From this equality and using a monotone class argument for the functions  $F_k^a$  defined in (5.5), we obtain (5.4) at rank  $k$ , that concludes the proof.  $\square$

*Proof of Theorem 2.1.* Let  $\psi \in \mathcal{D}$ ,  $t \geq 0$ , and  $\bar{z} \in \overline{\mathbb{V}}$ . Let  $((E_t, \bar{\chi}_t), t \geq 0)$  be the time-inhomogeneous  $\overline{\mathbb{W}}$ -valued branching process of generator  $\widehat{\mathcal{L}}_\psi$  defined in (2.4). Let  $\widehat{T}_{\text{Exp}}$  denote its explosion time and  $((Y_t, \chi_t), t \geq 0)$  its projection on  $\mathbb{W}$ . For every  $t < T_{\text{Exp}}$ , there exists  $k \in \mathbb{N}$  such that  $\{T_k \leq t < T_{k+1}\}$  where  $T_{k+1} = +\infty$  if there is no more jumps after  $T_k$ . Thus we can write

$$\mathbb{E}_{\bar{z}} \left[ \mathbb{1}_{\{T_{\text{Exp}} > t, \mathbb{G}(t) \neq \emptyset\}} H(U_t, (\bar{\nu}_s, s \leq t)) \right] = \sum_{k \geq 0} \mathbb{E}_{\bar{z}} \left[ \mathbb{1}_{\{T_k \leq t < T_{k+1}, \mathbb{G}(t) \neq \emptyset\}} H(U_t, (\bar{\nu}_s, s \leq t)) \right].$$

For every  $t \geq 0$ ,  $k \in \mathbb{N}$ , every non-extinguished sequence  $a = (a_i, 1 \leq i \leq k) \in \mathfrak{A}_k(\bar{z})$  with  $a_i = (u_i, n_i)$  and every  $e \in \mathbb{G}_k(a)$ , there exists a non-negative measurable function  $F_{k,a}^{t,e}$  on

$\bigotimes_{i=1}^k (\mathbb{R}_+ \times \mathcal{U} \times \mathbb{N} \times \mathcal{X}^{n_i})$ , such that

$$\mathbb{E}_{\bar{z}} \left[ \mathbb{1}_{\{T_{\text{Exp}} > t, \mathbb{G}(t) \neq \emptyset\}} H(U_t, (\bar{v}_s, s \leq t)) \right] = \sum_{k \geq 0} \sum_{a \in \mathfrak{A}_k(\bar{z})} \sum_{e \in \mathbb{G}_k(a)} \mathbb{E}_{\bar{z}} \left[ \mathbb{E} \left[ \mathbb{1}_{\{T_{k+1} > t\}} | \mathcal{F}_k \right] \mathbb{1}_{\{T_k \leq t\}} F_{k,a}^{t,e}(\mathcal{V}_i, 0 \leq i \leq k) \prod_{i=0}^k \mathbb{1}_{\{A_i = a_i\}} \right].$$

We apply Lemma 5.1 with the non-negative measurable function  $F$  defined for all sequence  $(v_i, 0 \leq i \leq k) \in \bigotimes_{i=1}^k (\mathbb{R}_+ \times \mathcal{U} \times \mathbb{N} \times \mathcal{X}^{n_i})$ , where for all  $0 \leq i \leq k$ ,  $v_i := (t_i, a_i, y_i)$ , by

$$F(v_i, 0 \leq i \leq k) := \mathbb{P} \left( S_{k+1} > t - t_k \mid \bigcap_{0 \leq i \leq k} \{\mathcal{V}_i = v_i\} \right) \mathbb{1}_{t_k \leq t} F_{k,a}^{t,e}(v_i, 0 \leq i \leq k),$$

where  $S_{k+1}$  is the random variable giving the  $(k+1)$ -th inter-arrival time of jumps in the original process. Note that  $S_{k+1}$  follows the same probability law as  $T_{k+1} - T_k$ . We thus get, for all  $e \in \mathbb{G}_k(a)$

$$\begin{aligned} \mathbb{E}_{\bar{z}} \left[ \mathbb{1}_{\{\mathbb{G}(t) \neq \emptyset\} \cap \{T_{\text{Exp}} > t\}} H(U_t, (\bar{v}_s, s \leq t)) \right] &= \langle z, \psi(\cdot, z) \rangle \sum_{k \geq 0} \sum_{a \in \mathfrak{A}_k(\bar{z})} \sum_{e \in \mathbb{G}_k(a)} \mathbb{E}_{\bar{z}} \left[ \mathbb{1}_{\{E_k = e\}} \right. \\ &\quad \left. \times \xi_k(E_k, (\widehat{\mathcal{V}}_i, 0 \leq i \leq k)) F(\widehat{\mathcal{V}}_i, 0 \leq i \leq k) \prod_{i=0}^k \mathbb{1}_{\{\widehat{A}_i = a_i\}} \right]. \end{aligned} \quad (5.11)$$

Then, using that  $\lambda$  introduced in (5.3) verifies  $\lambda := \widehat{\tau}_{\text{tot}} - \tau$ ,

$$\begin{aligned} \mathbb{P} \left( S_{k+1} > t - \widehat{T}_k \mid \bigcap_{0 \leq i \leq k} \{\widehat{\mathcal{V}}_i = v_i\} \right) \mathbb{1}_{\{\widehat{T}_k \leq t\}} &= e^{\int_{\widehat{T}_k}^t \lambda(Y_s, \chi_s) ds} e^{-\int_{\widehat{T}_k}^t \widehat{\tau}_{\text{tot}}(Y_s, \chi_s) ds} \mathbb{1}_{\{\widehat{T}_k \leq t\}} \\ &= e^{\int_{\widehat{T}_k}^t \lambda(Y_s, \chi_s) ds} \mathbb{E} \left[ \mathbb{1}_{\{\widehat{T}_{k+1} > t\}} | \widehat{\mathcal{F}}_k \right] \mathbb{1}_{\{\widehat{T}_k \leq t\}}. \end{aligned}$$

We recall that on the event  $\{\widehat{T}_{k+1} > t\}$ , there is no jump in the interval  $(\widehat{T}_k, t]$ , thus, applying Lemma 5.2 on this interval we get

$$\begin{aligned} \xi_k(E_k, (\widehat{\mathcal{V}}_i, 0 \leq i \leq k)) \mathbb{P} \left( S_k > t - \widehat{T}_k \mid \bigcap_{0 \leq i \leq k} \{\widehat{\mathcal{V}}_i = v_i\} \right) \mathbb{1}_{\{\widehat{T}_k \leq t\}} &= \\ \frac{1}{\psi(Y_t, \chi_t)} \exp \left( \int_0^t \frac{\mathcal{G}\psi(Y_s, \chi_s)}{\psi(Y_s, \chi_s)} ds \right) \mathbb{E} \left[ \mathbb{1}_{\{\widehat{T}_k \leq t < \widehat{T}_{k+1}\}} | \widehat{\mathcal{F}}_k \right]. \end{aligned} \quad (5.12)$$

On the event  $\{\widehat{T}_k \leq t < \widehat{T}_{k+1}\}$ ,  $E_k = E(t)$  almost surely, then using (5.11) and (5.12) gives

$$\begin{aligned} \mathbb{E}_{\bar{z}} \left[ \mathbb{1}_{\{T_{\text{Exp}} > t, \mathbb{G}(t) \neq \emptyset\}} H(U_t, (\bar{v}_s, s \leq t)) \right] &= \langle z, \psi(\cdot, z) \rangle \sum_{k \geq 0} \sum_{a \in \mathfrak{A}_k(\bar{z})} \sum_{e \in \mathbb{G}_k(a)} \mathbb{E}_{\bar{z}} \left[ \mathbb{1}_{\{E(t) = e\}} \right. \\ &\quad \left. \times \frac{\mathbb{1}_{\{\widehat{T}_k \leq t < \widehat{T}_{k+1}\}}}{\psi(Y_t, \chi_t)} \exp \left( \int_0^t \frac{\mathcal{G}\psi(Y_s, \chi_s)}{\psi(Y_s, \chi_s)} ds \right) F_{k,a}^{t,e}(\widehat{\mathcal{V}}_i, 0 \leq i \leq k) \prod_{i=0}^k \mathbb{1}_{\{\widehat{A}_i = a_i\}} \right]. \end{aligned}$$

Reconstructing the right-hand side and using the fact that the spine process does not extinct, we have

$$\mathbb{E}_{\bar{z}} \left[ \mathbb{1}_{\{T_{\text{Exp}} > t, \mathbb{G}(t) \neq \emptyset\}} H(U_t, (\bar{\nu}_s, s \leq t)) \right] = \langle z, \psi(\cdot, z) \rangle \mathbb{E}_{\bar{z}} \left[ \mathbb{1}_{\{\hat{T}_{\text{Exp}} > t\}} \frac{p_{E_t}(\bar{\chi}_t)}{\psi(Y_t, \chi_t)} \exp \left( \int_0^t \frac{\mathcal{G}\psi(Y_s, \chi_s)}{\psi(Y_s, \chi_s)} ds \right) H(E_t, (\bar{\chi}_s, s \leq t)) \right],$$

that concludes the proof.  $\square$

**5.3. Proofs of Section 3.** In this section we derive the results on the limiting martingale and a  $L \log L$  criterion.

*Proof of Proposition 3.1.* Let  $\psi \in \mathcal{D}$  satisfying Assumption 2, we first show for all  $t \geq 0$ , the integrability of the random variable  $W_t(\psi)$  introduced in (3.1). As Assumption 1 is satisfied, we can apply Corollary 2.5 for any  $t \geq 0$  to the positive function  $f$  on  $\mathbb{D}([0, t], \mathcal{X} \times \bar{\mathbb{V}})$ , such that for all  $u \in \mathbb{G}(t)$ ,

$$f((X_s^u, \bar{\nu}_s), 0 \leq s \leq t) := \exp \left( - \int_0^t \frac{\mathcal{G}\psi(X_s^u, \nu_s)}{\psi(X_s^u, \nu_s)} ds \right).$$

Corollary 2.5 applied to the function  $f$ , ensures that for any initial condition  $\bar{z}$ ,

$$\mathbb{E}_{\bar{z}} \left[ \sum_{u \in \mathbb{G}(t)} \exp \left( - \int_0^t \frac{\mathcal{G}\psi(X_s^u, \nu_s)}{\psi(X_s^u, \nu_s)} ds \right) \psi(X_t^u, \nu_t) \right] = \langle z, \psi(\cdot, z) \rangle. \quad (5.13)$$

This last identity guarantees the integrability of  $W_t(\psi)$ .

Now, for all  $r, t \in \mathbb{R}_+^*$ , such that  $t \geq r$ , we decompose the individuals alive in the population at time  $t$  according to their ancestors at time  $r$ . We then have

$$\begin{aligned} \mathbb{E}_{\bar{z}} [W_t(\psi) | \mathcal{F}_r] &= \sum_{v \in \mathbb{G}(r)} \exp \left( - \int_0^r \frac{\mathcal{G}\psi(X_s^v, \nu_s)}{\psi(X_s^v, \nu_s)} ds \right) \\ &\quad \times \mathbb{E}_{\bar{z}} \left[ \sum_{u \in \mathbb{G}(t), s.t. v \preceq u} \exp \left( - \int_r^t \frac{\mathcal{G}\psi(X_s^u, \nu_s)}{\psi(X_s^u, \nu_s)} ds \right) \psi(X_t^u, \nu_t) \middle| \mathcal{F}_r \right]. \end{aligned}$$

Now, in order to establish that the conditional expectation is equal to  $\psi(X_r^v, \nu_r)$  for all  $v \in \mathbb{G}(r)$ , we apply Corollary 2.5 to the positive functions  $f_v$  on  $\mathcal{U} \times \mathbb{D}([0, t], \bar{\mathbb{V}})$ , such that

$$f_v(u, (\bar{\nu}_s, 0 \leq s \leq t)) := \mathbb{1}_{v \preceq u} \exp \left( - \int_r^t \frac{\mathcal{G}\psi(X_s^u, \nu_s)}{\psi(X_s^u, \nu_s)} ds \right)$$

and use the Markov property. We get that for all  $v \in \mathbb{G}(r)$

$$\mathbb{E} \left[ \sum_{\substack{u \in \mathbb{G}(t), \\ s.t. v \preceq u}} \exp \left( - \int_r^t \frac{\mathcal{G}\psi(X_s^u, \nu_s)}{\psi(X_s^u, \nu_s)} ds \right) \psi(X_t^u, \nu_t) \middle| \mathcal{F}_r \right] = \langle \nu_r, \psi(\cdot, \nu_r) \rangle \mathbb{E}_{\nu_r} [\mathbb{1}_{\{v \preceq E_t\}}].$$

Finally using the new spinal individual distribution (2.9), we get

$$\mathbb{E} [W_t(\psi) | \mathcal{F}_r] = \sum_{v \in \mathbb{G}(r)} \exp \left( - \int_0^r \frac{\mathcal{G}\psi(X_s^v, \nu_s)}{\psi(X_s^v, \nu_s)} ds \right) \psi(X_r^v, \nu_r).$$

That concludes the proof.  $\square$

To establish Theorem 3.4, we need the following lemma from measure theory.

**Lemma 5.4.** *Let  $(\Omega, \mathcal{F}, \mu)$  be a probability space and let  $\hat{\mu}$  be a finite non negative measure on  $\Omega$ . Let  $(\mathcal{F}_t, 0 \leq t)$  be increasing  $\sigma$ -fields such that  $\sigma\left(\bigcup_{0 \leq t} \mathcal{F}_t\right) = \mathcal{F}$ , and  $\hat{\mu}_t, \mu_t$  be the restrictions of  $\hat{\mu}$  and  $\mu$  to  $\mathcal{F}_t$ . Suppose that there exists a non-negative  $\mathcal{F}_t$ -martingale  $(W_t, 0 \leq t)$  such that for all  $t \geq 0$*

$$\frac{d\hat{\mu}_t}{d\mu_t} = W_t.$$

Then, denoting  $W := \limsup_{t \rightarrow +\infty} W_t$ , we have the following dichotomy:

- (1)  $\int W d\mu = \int W_0 d\mu$  if and only if  $W < +\infty$   $\hat{\mu}$ -a.s.
- (2)  $W = 0$   $\mu$ -a.s. if and only if  $W = +\infty$   $\hat{\mu}$ -a.s.

*Proof.* We refer to [2, 45] for the proof of this result in discrete time. The extension to continuous time changes of measure uses Kolmogorov's extension theorem, see Durrett Appendix A [21] for further details.  $\square$

We now state the following lemma, that is the dual proposition of Theorem 3.4.

**Lemma 5.5.** *Let  $\psi \in \mathcal{D}$  satisfying Assumption 7, let  $(\widehat{W}_t(\psi), t \leq \widehat{T}_{Exp})$  be the  $\widehat{\mathcal{F}}_t$ -adapted process such that, for all  $t \leq \widehat{T}_{Exp}$ ,*

$$\widehat{W}_t(\psi) := \sum_{u \in \widehat{\mathbb{G}}(t)} \psi(X_t^u, \chi_t) \exp\left(-\int_0^t \frac{\mathcal{G}\psi(X_s^u, \chi_s)}{\psi(X_s^u, \chi_s)} ds\right), \quad (5.14)$$

where  $\widehat{\mathbb{G}}(t)$  is the set of labels of individuals living in the spinal population at time  $t$ .

Under Assumption 5, (3.2) implies that,  $\widehat{T}_{Exp} = \infty$  and  $\limsup_{t \rightarrow +\infty} \widehat{W}_t(\psi) < +\infty$  almost surely.

Under Assumption 6, (3.3) implies that  $\limsup_{t \rightarrow +\infty} \widehat{W}_t(\psi) = +\infty$  almost surely.

This lemma is the innovative part of this proof. It involves the decomposition of the spinal process as a process with immigration where the spinal individual provides new individuals at a  $\psi$ -biased rate, and the new spine is the new source of immigration. However as the function  $\psi$  is not constant, the spinal construction also changes the dynamics of individuals outside the spine, that do not behave as those in the original process  $(\nu_t, t \geq 0)$ . Assumption 5 is used to control the behavior of the individuals outside the spine to prove the non degeneracy of the limiting martingale. For the second part of this theorem, we use Assumption 6 to establish that the contribution of the offspring without the new spine is enough to ensure the degeneracy.

*Proof of Theorem 3.4.* Theorem 2.1 applied with any function  $\psi \in \mathcal{D}$  satisfying Assumption 7 exhibits the martingale  $(W_t(\psi), t \geq 0)$  as a Radon-Nikodym derivative. We can thus apply Lemma 5.4 to the spinal change of measure. The martingale  $W_t(\psi)$  and its limit  $W(\psi)$ , introduced in (3.1), verify the following dichotomy:

- (1)  $\mathbb{E}_{\bar{z}}[W(\psi)] = \mathbb{E}_{\bar{z}}[W_0(\psi)]$  if and only if  $\limsup_{t \rightarrow +\infty} \widehat{W}_t(\psi) < +\infty$  a.s.
- (2)  $W(\psi) = 0$  a.s. if and only if  $\limsup_{t \rightarrow +\infty} \widehat{W}_t(\psi) = +\infty$  a.s.

In this case, we use (5.13) to obtain that

$$\mathbb{E}_{\bar{z}}[W_0(\psi)] = \langle z, \psi(\cdot, z) \rangle,$$

and a direct application of Lemma 5.5 concludes the proof.  $\square$

*Proof of Lemma 5.5.* Let  $\psi \in \mathcal{D}$  satisfying Assumption 7, and introduce for all  $k \in \mathbb{N}$ ,

$$\widehat{p}_n^* := \sup_{(x, \nu) \in \mathbb{W}} \widehat{p}_n^*(x, \nu), \quad \underline{p}_n^* := \inf_{(x, \nu) \in \mathbb{W}} \widehat{p}_n^*(x, \nu),$$

$$\widehat{K}_n^*(\cdot) := \sup_{(x,\nu) \in \mathbb{W}} \widehat{K}_n^*(x, \nu, \cdot), \quad \text{and} \quad \underline{\widehat{K}}_n^*(\cdot) := \inf_{(x,\nu) \in \mathbb{W}} \widehat{K}_n^*((x, \nu, \cdot)),$$

where  $(\widehat{p}_n^*, n \geq 0)$  and  $\widehat{K}_n^*$  are respectively the law of the number of children and the measure giving the offspring traits for the  $\psi$ -spine, defined respectively in (2.7) and (2.6).

We first state two technical lemmas, proven in Appendix B.

**Lemma 5.6.** *For all  $\psi \in \mathcal{D}$  satisfying Assumption 7, the criteria*

$$\sum_{n \geq 1} \widehat{p}_n^* \int_{\mathcal{X}^n} \sup_{(x,\nu) \in \mathbb{W}} \left[ \log \left( \sum_{i=1}^n \psi(y^i, \nu^+(x, \mathbf{y})) \right) \right] \widehat{K}_n^*(\mathbf{y}) \mathcal{M}_n(d\mathbf{y}) < +\infty$$

implies that

$$\limsup_{t \rightarrow +\infty} \psi(Y_t, \chi_t) \exp \left( - \int_0^t \frac{\mathcal{G}\psi(Y_s, \chi_s)}{\psi(Y_s, \chi_s)} ds \right) = 0 \quad a.s.$$

We recall the notation  $\nu^+(x, \mathbf{y}) := \nu + \sum_{i=1}^n \delta_{y^i} - \delta_x$ , introduced in (2.2).

**Lemma 5.7.** *For all  $\psi \in \mathcal{D}$  satisfying Assumptions 6 and 7, the criteria*

$$\sum_{n \geq 1} \widehat{p}_n^* \int_{\mathcal{X}^n} \inf_{(x,\nu) \in \mathbb{W}} \left[ \log \left( \sum_{i=1}^n \psi(y^i, \nu^+(x, \mathbf{y})) \right) \right] \widehat{K}_n^*(\mathbf{y}) \mathcal{M}_n(d\mathbf{y}) = +\infty$$

implies that for any sequence of integers  $(j_n, n \geq 2)$  such that  $1 \leq j_n \leq n$ ,

$$\sum_{n \geq 2} \widehat{p}_n^* \int_{\mathcal{X}^n} \inf_{\substack{(x,\nu) \in \mathbb{W} \\ i \neq j_n}} \left[ \log \left( \sum_{i=1}^n \psi(y^i, \nu^+(x, \mathbf{y})) \right) \right] \widehat{K}_n^*(\mathbf{y}) \mathcal{M}_n(d\mathbf{y}) = +\infty.$$

To establish Lemma 5.5, we follow the conceptual decomposition of the spine process, first introduced by Lyons, Pemantle and Peres [45], that is, viewing the spine as an immigration source into a process without a spine. We consider the  $\psi$ -spinal process  $(\dot{\chi}_t, t \geq 0)$  describing all the individuals outside the spine and its associated marginal process  $\dot{\chi}_t$ , defined for all  $t \geq 0$  by:

$$\dot{\chi}_t := \bar{\chi}_t - \delta_{(E_t, Y_t)}, \quad \text{and} \quad \dot{\chi}_t := \chi_t - \delta_{Y_t}.$$

We denote by  $\dot{\mathbb{G}}(t)$  the random set of labels of non-spinal individuals living at time  $t$

$$\dot{\mathbb{G}}(t) = \left\{ u \in \mathcal{U} \setminus \{E_t\} : \int_{\mathcal{U} \times \mathcal{X}} \mathbb{1}_{\{v=u\}} \bar{\chi}_t(dv, dx) > 0 \right\}.$$

We also introduce for all  $t \geq 0$ ,

$$\dot{W}_t(\psi) := \sum_{u \in \dot{\mathbb{G}}(t)} \psi(X_t^u, \chi_t) \exp \left( - \int_0^t \Lambda(X_s^u, \chi_s) ds \right),$$

where for all  $x, \nu \in \mathbb{W}$ ,

$$\Lambda(x, \nu) := \frac{\mathcal{G}\psi(x, \nu)}{\psi(x, \nu)}.$$

• We handle first the **degenerated case**, and suppose that Assumption 6 holds and that

$$\sum_{n \geq 1} \widehat{p}_n^* \int_{\mathcal{X}^n} \inf_{(x,\nu) \in \mathbb{W}} \left[ \log \left( \sum_{i=1}^n \psi(y^i, \nu^+(x, \mathbf{y})) \right) \right] \widehat{K}_n^*(\mathbf{y}) \mathcal{M}_n(d\mathbf{y}) = +\infty. \quad (5.15)$$

We notice that, almost surely for all  $t \geq 0$ , the martingale  $\widehat{W}_t(\psi)$  introduced in (5.14) verifies

$$\dot{W}_t(\psi) \leq \widehat{W}_t(\psi). \quad (5.16)$$

We denote  $(\widehat{T}_k^*, k \geq 0)$  the sequence of random jumps times of the spine, and  $(\widehat{N}_k^*, k \geq 0)$  the sequence of random variables giving the number of children at each branching event of the spine. For every  $k \geq 0$ ,  $(X_k^i, 1 \leq i \leq \widehat{N}_k^*)$  is the random vector giving the types of the children of the spine, among them the trait of the new spine is denoted  $Y_{\widehat{T}_k^*}$ . We also introduce the sequence of filtrations  $(\mathcal{F}_k, k \geq 1)$  such that for all  $k \geq 0$ :

$$\mathcal{F}_k := \widehat{\mathcal{F}}_{\widehat{T}_{k+1}^*}. \quad (5.17)$$

where  $\widehat{\mathcal{F}}_t$  is the canonical filtration of the spinal process up to time  $t$ . Thus,  $\mathcal{F}_k$  corresponds to the information on the process until the time of the  $(k+1)$ -th jump. We notice that

$$\begin{aligned} \dot{\mathcal{W}}_{\widehat{T}_k^*} &= \dot{\mathcal{W}}_{\widehat{T}_k^* -} + \sum_{i=1}^{\widehat{N}_k^*} \psi \left( \widehat{X}_k^i, \chi_{\widehat{T}_k^*} \right) \exp \left( - \int_0^{\widehat{T}_k^*} \Lambda \left( X_s^i, \chi_s \right) ds \right) \\ &\quad - \psi \left( Y_{\widehat{T}_k^*}, \chi_{\widehat{T}_k^*} \right) \exp \left( - \int_0^{\widehat{T}_k^*} \Lambda \left( Y_s, \chi_s \right) ds \right). \end{aligned}$$

Using that for all  $t \geq 0$ ,  $\dot{\mathcal{W}}_t$  is almost surely non-negative, and using the upper bound of  $\Lambda$  in Assumption 7, we obtain

$$\begin{aligned} \dot{\mathcal{W}}_{\widehat{T}_k^*} &\geq \left( \sum_{i=1}^{\widehat{N}_k^*} \psi \left( \widehat{X}_k^i, \chi_{\widehat{T}_k^*} \right) - \psi \left( Y_{\widehat{T}_k^*}, \chi_{\widehat{T}_k^*} \right) \right) \exp \left( -C\widehat{T}_k^* \right) \\ &\geq \left( \sum_{i=1}^{\widehat{N}_k^*} \psi \left( \widehat{X}_k^i, \chi_{\widehat{T}_k^*} \right) - \max_{1 \leq j \leq n} \psi \left( X_k^j, \chi_{\widehat{T}_k^*} \right) \right) \exp \left( -C\widehat{T}_k^* \right). \end{aligned} \quad (5.18)$$

Let us introduce for all  $K > 0$  and all  $k \geq 1$ , the event

$$B_k^K := \left\{ \log \left( \sum_{i=1}^{\widehat{N}_k^*} \psi \left( \widehat{X}_k^i, \chi_{\widehat{T}_k^*} \right) - \max_{1 \leq j \leq k} \psi \left( X_k^j, \chi_{\widehat{T}_k^*} \right) \right) \geq kK \right\} \in \mathcal{F}_n.$$

We will now establish that

$$\sum_{k \geq 1} \mathbb{P} \left( B_k^K \mid \mathcal{F}_{k-1} \right) = +\infty.$$

First notice that

$$\begin{aligned} \sum_{k \geq 1} \mathbb{P} \left( B_k^K \mid \mathcal{F}_{k-1} \right) &= \sum_{k \geq 1} \sum_{n \geq 1} \widehat{p}_n^* \left( Y_{\widehat{T}_k^* -}, \chi_{\widehat{T}_k^* -} \right) \int_{\mathcal{X}^n} \widehat{K}_n^* \left( Y_{\widehat{T}_k^* -}, \chi_{\widehat{T}_k^* -}, \mathbf{y} \right) \\ &\quad \times \mathbb{1} \left\{ \log \left( \sum_{i=1}^n \psi \left( y^i, \chi_{\widehat{T}_k^*} \right) - \max_{1 \leq j \leq n} \psi \left( y^j, \chi_{\widehat{T}_k^*} \right) \right) \geq kK \right\} \mathcal{M}_n(d\mathbf{y}). \end{aligned}$$

Notice that the indicator function is always zero for  $n = 1$ . Taking the infimum over  $\mathbb{W}$ , we get that

$$\begin{aligned} \sum_{n \geq 1} \mathbb{P} \left( B_k^K \mid \mathcal{F}_{k-1} \right) &\geq \sum_{n \geq 2} \widehat{p}_n^* \int_{\mathcal{X}^n} \widehat{K}_n^* (\mathbf{y}) \\ &\quad \times \sum_{k \geq 1} \mathbb{1} \left\{ \inf_{(x, \nu) \in \mathbb{W}} \left( \log \left( \sum_{i=1}^n \psi \left( y^i, \nu^+(x, \mathbf{y}) \right) - \max_{1 \leq j \leq n} \psi \left( y^j, \nu^+(x, \mathbf{y}) \right) \right) \geq kK \right\} \mathcal{M}_n(d\mathbf{y}). \end{aligned}$$



Finally, using that  $\sum_{n \geq 2} \widehat{p}_n^* \int_{\mathcal{X}^n} \widehat{K}_n^*(\mathbf{y}) \mathcal{M}_n(d\mathbf{y}) \leq 1$  and that, for all  $A \in \mathbb{R}$ ,

$$\sum_{k \geq 0} \mathbb{1}_{\{A \geq kK\}} \geq -1 + \frac{A}{K},$$

we get

$$\begin{aligned} \sum_{k \geq 1} \mathbb{P} \left( B_k^K \middle| \mathcal{F}_{k-1} \right) &\geq -1 + \frac{1}{K} \sum_{n \geq 2} \widehat{p}_n^* \int_{\mathcal{X}^n} \widehat{K}_n^*(\mathbf{y}) \\ &\quad \times \inf_{(x, \nu) \in \mathbb{W}} \left( \log \left( \sum_{i=1}^n \psi(y^i, \nu^+(x, \mathbf{y})) - \max_{1 \leq j \leq n} \psi(y^j, \nu^+(x, \mathbf{y})) \right) \right) \mathcal{M}_n(d\mathbf{y}). \end{aligned}$$

Lemma 5.7 with (5.15) ensures that the lower bound is infinite and therefore

$$\sum_{k \geq 1} \mathbb{P} \left( B_k^K \middle| \mathcal{F}_{k-1} \right) = +\infty.$$

We can thus apply the conditional second Borel-Cantelli lemma, see Theorem 4.3.4 in [21]. For all  $K > 0$ ,

$$\limsup_{k \rightarrow \infty} \left( \sum_{i=1}^{\widehat{N}_k^*} \psi \left( \widehat{X}_k^i, \chi_{\widehat{T}_k^*} \right) - \psi \left( Y_{\widehat{T}_k^*}, \chi_{\widehat{T}_k^*} \right) \right) e^{-Kk} = +\infty \quad \text{a.s.}$$

Furthermore, the bounds on the branching rate in Assumption 7 ensure that  $\widehat{T}_k^*$  grows linearly almost surely to infinity as  $k$  tends to infinity. Thus, using (5.18) we get that  $\limsup_{t \rightarrow \infty} \mathring{W}_t(\psi) = +\infty$  and relation (5.16) concludes the proof.

• Now we treat the **non-degenerated** case, and suppose that Assumption 5 holds and that

$$\sum_{n \geq 1} \widehat{p}_n^* < +\infty, \quad \sup_{n \geq 1} \int_{\mathcal{X}^n} \widehat{K}_n^*(\mathbf{y}) \mathcal{M}_n(d\mathbf{y}) < +\infty, \quad (5.19)$$

and

$$\sum_{n \geq 1} \widehat{p}_n^* \int_{\mathcal{X}^n} \sup_{(x, \nu) \in \mathbb{W}} \left[ \log \left( \sum_{i=1}^n \psi(y^i, \nu^+(x, \mathbf{y}), t) \right) \right] \widehat{K}_n^*(\mathbf{y}) \mathcal{M}_n(d\mathbf{y}) < +\infty. \quad (5.20)$$

We notice that

$$\mathring{W}_t(\psi) = \widehat{W}_t(\psi) - \psi(Y_t, \chi_t) e^{-\int_0^t \Lambda(Y_s, \chi_s) ds}. \quad (5.21)$$

In the following we prove that  $\limsup_{t \rightarrow +\infty} \mathring{W}_t(\psi) < +\infty$  almost surely, and use Lemma 5.6 to conclude the proof.

Let  $\widehat{E} = \mathcal{U} \times \mathbb{R}_+ \times \mathbb{N} \times \mathcal{X}^{\mathbb{N}}$  and  $\widehat{E}^* = \mathbb{R}_+ \times \mathbb{N} \times \mathcal{X}^{\mathbb{N}}$ . We introduce  $\widehat{Q}(ds, du, dr, dn, d\mathbf{y})$  and  $\widehat{Q}^*(ds, dr, dn, d\mathbf{y})$  two independent Poisson point measures on  $\mathbb{R}_+ \times \widehat{E}$  and  $\mathbb{R}_+ \times \widehat{E}^*$  with respective intensity  $ds(\sum_{v \in \mathcal{U}} \delta_v(du)) dr \sum_{i \geq 1} \delta_i(dn) \mathcal{M}_i(d\mathbf{y})$  outside the spine and  $ds dr \sum_{i \geq 1} \delta_i(dn) \mathcal{M}_i(d\mathbf{y})$  for the spine. We denote by  $(\widehat{\mathcal{F}}_t^*, t \geq 0)$  the canonical filtration associated with  $\widehat{Q}^*$ . Recall that the set of labels of the living individuals in the population outside the spine is denoted by  $\mathring{\mathbb{G}}(t)$  and the random index of the new spine, introduced in (2.9), is denoted by  $J(x, \nu, \mathbf{y})$ .

We localize the process  $(\mathring{W}_t(\psi), t \geq 0)$  to avoid explosion and use Itô's formula to explicit  $\mathring{W}_t(\psi)$ . Let  $\widehat{T}^m$  be the  $m$ -th branching event of the spinal process for  $m \geq 1$ .

For all  $T \leq \widehat{T}^m$ ,

$$\begin{aligned}
 \dot{W}_T(\psi) &= \dot{W}_0 - \int_0^T I_t dt + \int_0^T \int_{\widehat{E}} \widehat{I}_t \widehat{Q}(dt, du, dr, dn, d\mathbf{y}) \\
 &+ \int_0^T \int_{\widehat{E}^*} \left\{ \sum_{i=1}^n \mathbb{1}_{\{i \neq J(Y_{t-}, \chi_{t-}, \mathbf{y})\}} \psi(y^i, \chi_t^+(Y_t, \mathbf{y})) \exp\left(-\int_0^t \Lambda(Y_s, \chi_s) ds\right) \right. \\
 &\quad \left. + \sum_{v \in \mathring{\mathbb{G}}_{t-}} [\psi(X_t^v, \chi_t^+(Y_t, \mathbf{y})) - \psi(X_t^v, \chi_{t-})] \exp\left(-\int_0^t \Lambda(X_s^v, \chi_s) ds\right) \right\} \\
 &\quad \times \mathbb{1}_{\{r \leq \widehat{B}_n^*(Y_{t-}, \chi_{t-}) \widehat{K}_n^*(Y_{t-}, \chi_{t-}, \mathbf{y})\}} \widehat{Q}^*(dt, dr, dn, d\mathbf{y}) \quad (5.22)
 \end{aligned}$$

where

$$I_t := \sum_{u \in \mathring{\mathbb{G}}_t} [\mathcal{G}\psi - G\psi](X_t^u, \chi_t) e^{-\int_0^t \Lambda(X_s^u, \chi_s) ds},$$

and

$$\begin{aligned}
 \widehat{I}_t &:= \left\{ \left[ \sum_{i=1}^n \psi(y^i, \chi_t^+(X_t^u, \mathbf{y})) - \psi(X_t^u, \chi_t^+(X_t^u, \mathbf{y})) \right] \exp\left(-\int_0^t \Lambda(X_s^u, \chi_s) ds\right) \right. \\
 &\quad \left. + \sum_{v \in \mathring{\mathbb{G}}_{t-}} [\psi(X_t^v, \chi_t^+(X_t^u, \mathbf{y})) - \psi(X_t^v, \chi_{t-})] \exp\left(-\int_0^t \Lambda(X_s^v, \chi_s) ds\right) \right\} \\
 &\quad \times \mathbb{1}_{\{u \in \mathring{\mathbb{G}}_t\}} \mathbb{1}_{\{r \leq \widehat{B}_n(Y_t, X_t^u, \chi_{t-}) \widehat{K}_n(Y_t, X_t^u, \chi_{t-}, \mathbf{y})\}}. \quad (5.23)
 \end{aligned}$$

$I_t$  describes the deterministic evolution of types between the time of jump,  $\widehat{I}_t$  corresponds to the jumps outside the spine, and the last integral in (5.22) is the contribution from the spinal individual.

Using the expression of the operator  $\mathcal{G}$  and Assumption 5 we have for all  $t > 0$ , all  $u \in \mathring{\mathbb{G}}_t$

$$\begin{aligned}
 I_t &= \sum_{u \in \mathring{\mathbb{G}}_t} \exp\left(-\int_0^t \Lambda(X_s^u, \chi_s) ds\right) \sum_{n \geq 0} B_n(X_t^u, \chi_t) \\
 &\quad \times \int_{\mathcal{X}^n} \left[ \sum_{i=1}^n \psi(y^i, \chi_t^+(X_t^u, \mathbf{y})) - \psi(X_t^u, \chi_t^+(X_t^u, \mathbf{y})) \right] K_n(X_t^u, \chi_t, \mathbf{y}) \mathcal{M}_n(d\mathbf{y}).
 \end{aligned}$$

Notice that Assumption 5 implies that, for all  $(x, (x_e, \nu), t) \in \mathcal{X} \times \mathbb{W}$ ,

$$\widehat{B}_n(x_e, x, \nu) \widehat{K}_n(x_e, x, \nu, \mathbf{y}) = B_n(x, \nu) K_n(x, \nu, \mathbf{y}).$$

Thus using Assumption 5 in (5.23) we get

$$\int_0^T \int_{\widehat{E}} \widehat{I}_t dt dr \sum_{i \geq 1} \delta_i(dn) \mathcal{M}_i(d\mathbf{y}) = \int_0^T I_t dt.$$

Thus, using Assumption 5 we get, for all  $T \leq \widehat{T}^m$

$$\begin{aligned}
 \mathbb{E} \left[ \dot{W}_T(\psi) \middle| \widehat{\mathcal{F}}_T^* \right] &= \mathbb{E} \left[ \dot{W}_0(\psi) \right] \\
 &+ \int_0^T \int_{\widehat{E}^*} \sum_{i=1}^n \mathbb{1}_{\{i \neq J(Y_{t-}, \chi_{t-}, t, \mathbf{y})\}} \psi(y^i, \chi_t^+(Y_t, \mathbf{y})) \exp\left(-\int_0^t \Lambda(Y_s, \chi_s) ds\right) \\
 &\quad \times \mathbb{1}_{\{r \leq \widehat{B}_n^*(Y_{t-}, \chi_{t-}) \widehat{K}_n^*(Y_{t-}, \chi_{t-}, \mathbf{y})\}} \widehat{Q}^*(dt, dr, dn, d\mathbf{y}).
 \end{aligned}$$

Using the positivity of  $\psi$  we get a majorant by adding the new spine in the sum. We then take the supremum over  $\mathbb{W}$  in the indicator function and notice that for all  $n \geq 0$ , all  $(x, \nu) \in \mathbb{W}$ ,

$$\widehat{B}_n^*(x, \nu) = \widehat{p}_n^*(x, \nu) \sum_{k \geq 0} \widehat{B}_k^*(x, \nu) \leq \widehat{p}_n^* \bar{\tau},$$

where we used Assumption 7 for the last inequality. Putting these steps together and using the lower bound for  $\mathcal{G}\psi/\psi$  in Assumption 7 we get:

$$\begin{aligned} \mathbb{E} \left[ \widehat{W}_T(\psi) \middle| \widehat{\mathcal{F}}_T^* \right] &\leq \mathbb{E} \left[ \widehat{W}_0(\psi) \right] + \int_0^T \int_{\widehat{E}^*} \mathbb{1}_{\{r \leq \bar{\tau} \widehat{p}_n^* \widehat{K}_n^*(\mathbf{y})\}} \\ &\quad \times \exp \left\{ \sup_{(x, \nu) \in \mathbb{W}} \left( \log \left( \sum_{i=1}^k \psi(y^i, \nu_+(x, \mathbf{y})) \right) \right) - ct \right\} \widehat{Q}^*(dt, dr, dn, d\mathbf{y}). \end{aligned} \quad (5.24)$$

Finally we use (5.19) to get an upper bound of the term in the indicator

$$\bar{\tau} \widehat{p}_n^* \widehat{K}_n^*(\mathbf{y}) \leq \bar{\tau} \left( \sum_{i \geq 1} \bar{p}_i^* \right) \left( \sup_{i \geq 1} \int_{\mathcal{X}^i} \widehat{K}_i^*(\mathbf{y}) \mathcal{M}_k(d\mathbf{y}) \right) \left( \frac{\bar{p}_n^*}{\sum_{i \geq 1} \bar{p}_i^*} \right) \left( \frac{\widehat{K}_n^*(\mathbf{y})}{\int_{\mathcal{X}^n} \widehat{K}_n^*(\mathbf{w}) \mathcal{M}_n(d\mathbf{w})} \right),$$

and introduce the sequences of independent random variables  $(S_k, k \geq 1)$  of exponential law of parameter  $\bar{\tau} \left( \sum_{i \geq 1} \bar{p}_i^* \right) \left( \sup_{i \geq 1} \int_{\mathcal{X}^i} \widehat{K}_i^*(\mathbf{y}) \mathcal{M}_i(d\mathbf{y}) \right)$ . We also introduce the independent sequence of couples of random variables  $((N_k, \mathbf{Y}_k), k \geq 1)$  such that for all  $k$ ,  $N_k$  follows the law  $(\bar{p}_n^* / \sum_{i \geq 1} \bar{p}_i^*)$  and  $\mathbf{Y}_k | N_k$  follows the law  $\widehat{K}_{N_k}^*(\mathbf{y}) / \int_{\mathcal{X}^N} \widehat{K}_{N_k}^*(\mathbf{w}) \mathcal{M}_{N_k}(d\mathbf{w})$ . Thus we can write (5.24) with a compound Poisson process

$$\begin{aligned} \mathbb{E} \left[ \widehat{W}_T(\psi) \middle| \widehat{\mathcal{F}}_T^* \right] &\leq \mathbb{E} \left[ \widehat{W}_0(\psi) \right] \\ &\quad + \sum_{k \geq 1} \exp \left\{ \sup_{(x, \nu) \in \mathbb{W}} \left( \log \left( \sum_{j=1}^{N_k} \psi(Y_k^j, \nu_+(x, \mathbf{Y}_k)) \right) \right) - c \sum_{j=1}^k S_j \right\}. \end{aligned}$$

In order to show that the series is almost surely finite, we introduce the following sequence of events: for every  $K > 0$  and every  $k \geq 0$

$$B_k^K := \left\{ \sup_{(x, \nu) \in \mathbb{W}} \left( \log \left( \sum_{j=1}^{N_k} \psi(Y_k^j, \nu_+(x, \mathbf{Y}_k)) \right) \right) \geq kK \right\}.$$

Using the law of the couple of random variables  $((N_k, \mathbf{Y}_k), k \geq 1)$  we get that

$$\begin{aligned} \sum_{k \geq 1} \mathbb{P}(B_k^K) &= \sum_{k \geq 1} \sum_{n \geq 1} \frac{\bar{p}_n^*}{\sum_{i \geq 1} \bar{p}_i^*} \\ &\quad \times \int_{\mathcal{X}^n} \mathbb{1}_{\left\{ \sup_{(x, \nu) \in \mathbb{W}} (\log(\sum_{j=1}^n \psi(y^j, \nu_+(x, \mathbf{y}))) \geq kK \right\}} \frac{\widehat{K}_n^*(\mathbf{y}) \mathcal{M}_n(d\mathbf{y})}{\int_{\mathcal{X}^n} \widehat{K}_n^*(\mathbf{w}) \mathcal{M}_n(d\mathbf{w})}. \end{aligned}$$

Using that  $\int_{\mathcal{X}^n} \widehat{K}_n^*(\mathbf{w}) \mathcal{M}_n(d\mathbf{w}) \geq 1$ ,  $\sum_{i \geq 1} \bar{p}_i^* \geq 1$  and  $\sum_{k \geq 1} \mathbb{1}_{A \geq kK} \leq 0 \vee (A/K)$ , we get

$$\begin{aligned} \sum_{k \geq 1} \mathbb{P}(B_k^K) &\leq \\ &\quad 0 \vee \left( \frac{1}{K} \sum_{k \geq 1} \bar{p}_k^* \int_{\mathcal{X}^n} \sup_{(x, \nu) \in \mathbb{W}} \left( \log \left( \sum_{j=1}^k \psi(y^j, \nu_+(x, \mathbf{y})) \right) \right) \widehat{K}_k^*(\mathbf{y}) \mathcal{M}_n(d\mathbf{y}) \right). \end{aligned}$$

We now use the Borel Cantelli lemma -see Theorem 2.3.1 in [21]- with (5.20) to ensure that

$$\limsup_{n \rightarrow \infty} \frac{\sup_{(x, \nu) \in \mathbb{W}} \left( \log \left( \sum_{j=1}^{N_n} \psi \left( Y_n^j, \nu_+(x, \mathbf{Y}_n) \right) \right) \right)}{n} = 0 \quad \text{a.s.}$$

Notice that and (5.19) ensures that, asymptotically,  $\sum_{j=1}^k S_j$  grows linearly with  $k$ . Thus

$$\sum_{k \geq 1} \exp \left\{ \sup_{(x, \nu) \in \mathbb{W}} \left( \log \left( \sum_{j=1}^{N_k} \psi \left( Y_k^j, \nu_+(x, \mathbf{Y}_k) \right) \right) \right) - c \sum_{j=1}^k S_j \right\} < \infty \quad \text{a.s.}$$

The upper bound for the branching rate in Assumption 7 ensures that  $\widehat{T}^m \rightarrow \infty$  almost surely as  $m$  tends to infinity, and Fatou's lemma gives that  $\sup_{t \geq 0} \mathbb{E}_{\bar{z}} \left[ \widehat{W}_t(\psi) | \widehat{\mathcal{F}}_\infty^* \right] < \infty$ . Thus, the quenched submartingale  $(\widehat{W}_{t \wedge \widehat{T}^m}(\psi), t \geq 0)$  converges almost surely to a finite random variable and

$$\limsup_{t \rightarrow +\infty} \widehat{W}_t(\psi) < +\infty \quad \text{a.s.}$$

We conclude the proof using (5.21) and Lemma 5.6.  $\square$

#### APPENDIX A. PROOF OF PROPOSITION 2.4

The spine process  $((E_t, \bar{\chi}_t), t \geq 0)$  defined by Dynamics 2.2 and 2.3 can be rigorously expressed as the solution of a SDE driven by a multivariate point measure.

We recall that the offspring traits at birth  $\mathbf{y} = (y^1, \dots, y^n)$  of an individual of trait  $x$  in a population  $\nu$  at time  $t$  is given by the law  $K_n(x, \nu, \mathbf{y}) \mathcal{M}_n(d\mathbf{y})$ . Let  $\widehat{E} = \mathbb{R}_+ \times \mathcal{U} \times \mathbb{R}_+ \times \mathbb{N} \times \mathcal{X}^{\mathbb{N}}$  and  $\widehat{E}^* = \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{N} \times \mathcal{X}^{\mathbb{N}}$ . Let  $\widehat{Q}(ds, du, dr, dn, d\mathbf{y})$  and  $\widehat{Q}^*(ds, dr, dn, d\mathbf{y})$  be two independent Poisson point measures on  $\widehat{E}$  and  $\widehat{E}^*$  with respective intensity  $ds(\sum_{v \in \mathcal{U}} \delta_v(du)) dr \sum_{i \geq 1} \delta_i(dn) \mathcal{M}_i(d\mathbf{y})$  outside the spine and  $ds dr \sum_{i \geq 1} \delta_i(dn) \mathcal{M}_i(d\mathbf{y})$  for the spine. We denote by  $(\widehat{\mathcal{F}}_t, t \geq 0)$  the canonical filtration associated with these Poisson point measures. The  $\widehat{\mathcal{F}}_t$ -adapted set of labels of the living individuals in the population outside the spine is denoted  $\widehat{\mathbb{G}}(t)$ . Finally we recall the notation  $J(x, \nu, \mathbf{y})$ , introduced in (2.9), for the random index of the new spine after the branching of a spine of trait  $x$  in a population  $\nu$  at time  $t$  to an offspring  $\mathbf{y}$ .

Let  $(e, \bar{z}) \in \overline{\mathbb{W}}$ . Under Assumptions 1 and 3, the process  $((E_t, \bar{\chi}_t), t \geq 0)$  is the unique  $\widehat{\mathcal{F}}_t$ -adapted solution, for every function  $g \in \mathcal{C}^1(\mathcal{U} \times \mathcal{U} \times \mathcal{X}, \mathbb{R})$  and  $t \geq 0$ , of the following equation

$$\begin{aligned} \langle \bar{\chi}_t, g(E_t, \cdot) \rangle &:= \langle \bar{z}, g(e, \cdot) \rangle + \int_0^t \int_{\mathcal{U} \times \mathcal{X}} \frac{\partial g}{\partial x}(E_s, u, x) \cdot \mu(x, \chi_s) \bar{\chi}_s(du, dx) ds \\ &+ \int_{\widehat{E}} \left[ \sum_{i=1}^n g(E_{s-}, ui, \mathbf{y}) - g(E_{s-}, u, X_{s-}^u) \right] \\ &\quad \times \mathbb{1}_{\{u \in \widehat{\mathbb{G}}_{s-}\}} \mathbb{1}_{\{r \leq \widehat{B}_n(Y_{s-}, X_{s-}^u, \chi_{s-}) \widehat{K}_n(Y_{s-}, X_{s-}^u, \chi_{s-}, \mathbf{y})\}} \widehat{Q}(ds, du, dr, dn, d\mathbf{y}) \\ &+ \int_{\widehat{E}^*} \left[ \sum_{i=1}^n g(E_{s-} J(Y_{s-}, \chi_{s-}, \mathbf{y}), E_{s-} J(Y_{s-}, \chi_{s-}, \mathbf{y}), \mathbf{y}) - g(E_{s-}, E_{s-}, Y_{s-}) \right] \\ &\quad \times \mathbb{1}_{\{r \leq \widehat{B}_k^*(Y_{s-}, \chi_{s-}) \widehat{K}_k^*(Y_{s-}, \chi_{s-}, \mathbf{y})\}} \widehat{Q}^*(ds, dr, dn, d\mathbf{y}). \end{aligned} \quad (\text{A.1})$$

This assertion is shown following the same computations than for proof of Proposition 1.1 and [46], using the positivity of the  $\psi$  function. Assumption 3 leads to the same bound as Assumption 1.3 in the case of a spine process.

Now we establish the expression of the generator of the marginal spine process. We recall the notation  $\nu^+(x, \mathbf{y}) := \nu + \sum_{i=1}^n \delta_{y^i} - \delta_x$ , introduced in (2.2). Taking expectations of (A.1) for the marginal process on  $\mathbb{W}$ , we derive the non-homogeneous infinitesimal operator of  $((Y_t, \chi_t), t \geq 0)$ , following steps in [25]. It is given by the operator  $\widehat{L}_\psi$ , defined for every  $F \in \mathcal{D}$ , introduced in (2.1), and  $(x_e, \nu) \in \mathbb{W}$ , by

$$\begin{aligned} \widehat{L}_\psi F(x_e, \nu) &:= GF(x_e, \nu) \\ &+ \sum_{n \geq 0} \left\{ \widehat{B}_n^*(x_e, \nu) \int_{\mathcal{X}^n} \left[ \frac{\sum_{i=1}^n F \psi(y^i, \nu^+(x_e, \mathbf{y}))}{\sum_{j=1}^n \psi(y^j, \nu^+(x_e, \mathbf{y}))} - F(x_e, \nu) \right] \widehat{K}_n^*(x_e, \nu, \mathbf{y}) \mathcal{M}_n(d\mathbf{y}) \right. \\ &\quad + \int_{\mathcal{X}} \widehat{B}_n(x, \nu) \int_{\mathcal{X}^n} \left[ F(x_e, \nu^+(x, \mathbf{y})) - F(x_e, \nu) \right] \widehat{K}_n(x, \nu, \mathbf{y}) \mathcal{M}_n(d\mathbf{y}) \nu(dx) \\ &\quad \left. - \widehat{B}_n(x_e, \nu) \int_{\mathcal{X}^n} \left[ F(x_e, \nu^+(x_e, \mathbf{y})) - F(x_e, \nu) \right] \widehat{K}_n(x_e, \nu, \mathbf{y}) \mathcal{M}_n(d\mathbf{y}) \right\}. \end{aligned}$$

The first line gives the dynamical evolution between branching events. The second is related to spinal branching events and the choice of a new spinal individual among the offspring population. The last two lines describe the branching events outside the spine, for all individuals but the spinal one. Using that  $G[\psi F] = FG\psi + \psi GF$ , we get

$$\frac{G[\psi F]}{\psi}(\cdot) - \frac{G\psi}{\psi}F(\cdot) = GF(\cdot) + \sum_{n \geq 0} \frac{B(\cdot)}{\psi(\cdot)} \widehat{T}_n(\cdot),$$

where the jump part  $\widehat{T}_n$  is defined for all  $(x_e, \nu) \in \mathbb{W}$  by

$$\begin{aligned} \widehat{T}_n(\rho_e) &:= p_n(\rho_e) \int_{\mathcal{X}^n} \sum_{i=1}^n \psi F(y_i, \nu^+(x_e, \mathbf{y})) - \psi F(x_e, \nu^+(x_e, \mathbf{y})) K_n(\rho_e, \mathbf{y}) \mathcal{M}_n(d\mathbf{y}) \\ &\quad + \int_{\mathcal{X}} p_n(x, \nu) \int_{\mathcal{X}^n} [\psi F(x_e, \nu^+(x, \mathbf{y})) - \psi F(\rho_e)] K_n(x, \nu, \mathbf{y}) \mathcal{M}_n(d\mathbf{y}) \nu(dx) \\ &\quad - p_n(\rho_e) \int_{\mathcal{X}^n} \left[ \sum_{i=1}^n \psi(y_i, \nu^+(x_e, \mathbf{y})) - \psi(x_e, \nu^+(x_e, \mathbf{y})) \right] F(\rho_e) K_n(\rho_e, \mathbf{y}) \mathcal{M}_n(d\mathbf{y}) \\ &\quad - \int_{\mathcal{X}} p_n(x, \nu) \int_{\mathcal{X}^n} [\psi(x_e, \nu^+(x, \mathbf{y})) - \psi(\rho_e)] F(\rho_e) K_n(x, \nu, \mathbf{y}) \mathcal{M}_n(d\mathbf{y}) \nu(dx), \end{aligned} \quad (\text{A.2})$$

where we used the notation  $\rho_e := (x_e, \nu)$ . Rearranging the terms in (A.2) we get

$$\begin{aligned} \widehat{T}_n(\rho_e) &= p_n(\rho_e) \int_{\mathcal{X}^n} \sum_{i=1}^n \psi(y_i, \nu^+(x_e, \mathbf{y})) [F(y_i, \nu^+(x_e, \mathbf{y})) - F(\rho_e)] K_n(\rho_e, \mathbf{y}) \mathcal{M}_n(d\mathbf{y}) \\ &\quad + \int_{\mathcal{X}} p_n(x, \nu) \\ &\quad \quad \times \int_{\mathcal{X}^n} [F(x_e, \nu^+(x, \mathbf{y})) - F(\rho_e)] \psi(x_e, \nu^+(x, \mathbf{y})) K_n(x, \nu, \mathbf{y}) \mathcal{M}_n(d\mathbf{y}) \nu(dx) \\ &\quad - p_n(\rho_e) \int_{\mathcal{X}^n} [F(x_e, \nu^+(x_e, \mathbf{y})) - F(\rho_e)] \psi(x_e, \nu^+(x_e, \mathbf{y})) K_n(\rho_e, \mathbf{y}) \mathcal{M}_n(d\mathbf{y}). \end{aligned}$$

We conclude the proof using the branching rates introduced in Dynamics 2.2 and 2.3.

#### APPENDIX B. PROOF OF LEMMAS 5.6 AND 5.7

*Proof of Lemma 5.6.* We recall the notations  $((\widehat{T}_k^*, \widehat{N}_k^*, (\widehat{X}_k^i, 1 \leq i \leq \widehat{N}_k^*)), k \geq 0)$  for the sequence of jumps times, number of children and type corresponding to branching events of the spine. The

trait of the new spine is denoted  $Y_{\widehat{T}_k^*}$ . We also recall the sequence of filtrations  $(\mathcal{F}_k, k \geq 1)$ , introduced in (5.17). Using these notations, with the lower bound for  $\mathcal{G}\psi/\psi$  in Assumption 7 and the positivity of the function  $\psi$ , we have for all  $k \geq 0$

$$\log\left(\psi\left(Y_{\widehat{T}_k^*}, \chi_{\widehat{T}_k^*}\right)\right) - \int_0^{\widehat{T}_k^*} \Lambda(Y_s, \chi_s) ds \leq \log\left(\sum_{i=1}^{\widehat{N}_k^*} \psi\left(\widehat{X}_k^i, \chi_{\widehat{T}_k^*}\right)\right) - c\widehat{T}_k^* \quad (\text{B.1})$$

as  $Y_{T_k} \in \{X_k^i, 1 \leq i \leq N_k\}$ . We introduce for all  $K > 0$  and all  $k \geq 1$ , the event

$$B_k^K := \left\{ \log\left(\sum_{i=1}^{\widehat{N}_k^*} \psi\left(\widehat{X}_k^i, \chi_{\widehat{T}_k^*}\right)\right) \geq kK \right\} \in \mathcal{F}_k.$$

First notice that

$$\begin{aligned} \sum_{k \geq 1} \mathbb{P}\left(B_k^K \mid \mathcal{F}_{k-1}\right) &= \sum_{k \geq 1} \sum_{n \geq 1} \widehat{p}_n^*\left(Y_{\widehat{T}_k^*}, \chi_{\widehat{T}_k^*}\right) \\ &\quad \times \int_{\mathcal{X}^n} \mathbb{1}_{\left\{\log\left(\sum_{i=1}^n \psi\left(y^i, \chi_{\widehat{T}_k^*}\right)\right) \geq kK\right\}} \widehat{K}_n^*\left(Y_{\widehat{T}_k^*}, \chi_{\widehat{T}_k^*}, \mathbf{y}\right) \mathcal{M}_n(d\mathbf{y}). \end{aligned}$$

Taking the supremum over  $\mathbb{W}$ , we get that

$$\begin{aligned} \sum_{k \geq 1} \mathbb{P}\left(B_k^K \mid \mathcal{F}_{k-1}\right) &\leq \\ &\sum_{n \geq 1} \widehat{p}_n^* \int_{\mathcal{X}^n} \sum_{k \geq 1} \mathbb{1}_{\left\{\sup_{(x, \nu) \in \mathbb{W}} \left(\log\left(\sum_{i=1}^n \psi\left(y^i, \nu^+(x, \mathbf{y})\right)\right)\right) \geq kK\right\}} \widehat{K}_n^*(\mathbf{y}) \mathcal{M}_n(d\mathbf{y}). \end{aligned}$$

We recall the notation  $\nu^+(x, \mathbf{y}) := \nu + \sum_{i=1}^n \delta_{y^i} - \delta_x$ , introduced in (2.2). Finally, using that for all  $A \in \mathbb{R}$ ,  $\sum_{k \geq 1} \mathbb{1}_{A \geq kK} \leq 0 \vee (A/K)$  we get

$$\begin{aligned} \sum_{k \geq 1} \mathbb{P}\left(B_k^K \mid \mathcal{F}_{k-1}\right) &\leq \\ &0 \vee \left( \frac{1}{K} \sum_{n \geq 1} \widehat{p}_n^* \int_{\mathcal{X}^n} \sup_{(x, \nu) \in \mathbb{W}} \left[ \log\left(\sum_{i=1}^n \psi\left(y^i, \nu^+(x, \mathbf{y})\right)\right) \right] \widehat{K}_n^*(\mathbf{y}) \mathcal{M}_n(d\mathbf{y}) \right). \end{aligned}$$

Thus (5.20) ensures that

$$\sum_{k \geq 1} \mathbb{P}\left(B_k^K \mid \mathcal{F}_{k-1}\right) < +\infty,$$

and the second Borel-Cantelli Lemma- see Theorem 4.3.4 in [21]- gives that for all  $K > 0$ , eventually

$$\log\left(\psi\left(Y_{\widehat{T}_k^*}, \chi_{\widehat{T}_k^*}\right)\right) \leq kK \quad \text{a.s.}$$

Furthermore, the bounds on the jump rate in Assumption 7 ensure that  $\widehat{T}_k^*$  grows linearly almost surely to infinity as  $k$  tends to infinity. Thus, using (B.1) we get that almost surely

$$\limsup_{k \rightarrow \infty} \left[ \log\left(\psi\left(Y_{\widehat{T}_k^*}, \chi_{\widehat{T}_k^*}\right)\right) - \int_0^{\widehat{T}_k^*} \Lambda(Y_s, \chi_s) ds \right] = -\infty.$$

The fact that  $\lim_{k \rightarrow \infty} \widehat{T}_k^* = +\infty$  almost surely concludes the proof.  $\square$

*Proof of Lemma 5.7.* First, notice that, using a log-sum inequality along with the positivity of the function  $\psi$  we get, for all  $n \geq 2$ ,  $(x, \nu) \in \mathbb{W}$ ,  $\mathbf{y} \in \mathcal{X}^n$  and  $j_n \in \{1, \dots, n\}$

$$\begin{aligned} \log \left( \sum_{i=1}^n \psi(y^i, \nu^+(x, \mathbf{y})) \right) &= \log \left( \sum_{i=1}^n \mathbb{1}_{i \neq j_n} \psi(y^i, \nu^+(x, \mathbf{y})) + \psi(y^{j_n}, \nu^+(x, \mathbf{y})) \right) \\ &\leq \log(2) + \frac{\psi(y^{j_n}, \nu^+(x, \mathbf{y}))}{\sum_{i=1}^n \psi(y^i, \nu^+(x, \mathbf{y}))} \log(\psi(y^{j_n}, \nu^+(x, \mathbf{y}))) \\ &\quad + \frac{\sum_{i=1}^n \mathbb{1}_{i \neq j_n} \psi(y^i, \nu^+(x, \mathbf{y}))}{\sum_i \psi(y^i, \nu^+(x, \mathbf{y}))} \log \left( \sum_{i=1}^n \mathbb{1}_{i \neq j_n} \psi(y^i, \nu^+(x, \mathbf{y})) \right). \end{aligned}$$

Thus, taking the infimum in the previous inequality and using (3.3), we have for any arbitrary state  $(x_0, \nu_0) \in \mathbb{W}$

$$\begin{aligned} \sum_{n \geq 2} \widehat{p}_n^*(\rho_0) \int_{\mathcal{X}^n} \frac{\psi(y^{j_n}, \nu_0^+(x_0, \mathbf{y}))}{\sum_{i=1}^n \psi(y^i, \nu_0^+(x_0, \mathbf{y}))} \log(\psi(y^{j_n}, \nu_0^+(x_0, \mathbf{y}))) \widehat{K}_n^*(\rho_0, \mathbf{y}) \mathcal{M}_n(d\mathbf{y}) \\ + \sum_{n \geq 2} \widehat{p}_n^* \int_{\mathcal{X}^n} \inf_{(x, \nu) \in \mathbb{W}} \left[ \log \left( \sum_{i \neq j_n} \psi(y^i, \nu^+(x, \mathbf{y})) \right) \right] \widehat{K}_n^*(\mathbf{y}) \mathcal{M}_n(d\mathbf{y}) = +\infty \quad (\text{B.2}) \end{aligned}$$

where we used the notation  $\rho_0 := x_0, \nu_0$ . We conclude the proof using Assumption 6 to ensure that the first term in (B.2) is finite.  $\square$

### APPENDIX C. ALGORITHMIC CONSTRUCTION

We propose an algorithm that generates trajectories of the  $\psi$ -spine process introduced in (4.3). We denote by  $F_{\text{div}}^{-1}$ ,  $F_{\text{loss}}^{-1}$  and  $F_{\text{loss}}^{*-1}$  the generalized inverse of the cumulative distribution functions of the random variables  $\widehat{\Lambda}$ ,  $\widehat{\Theta}$  and  $\Theta$ , defined in (4.4) and (4.5). Using the deterministic evolution of the traits and knowing the branching events, it is easy to recover the traits at all time of the individuals.

We first generate a realization of an homogeneous Poisson point process of intensity 1 on the interval  $[t_1, t_2]$ , starting from  $n$  roots following classical algorithms [49]. It returns a list  $T_{\text{div}} = [t_1, T_1, T_2, \dots]$  of increasing times, a list  $I_{\text{div}} = [i_1, i_2, \dots]$  containing the numbering of the individuals that branched at these times, a list  $L_{\text{div}} = [\lambda_1, \lambda_2, \dots]$  of fractions of mass at birth, and the numbering  $E$  of the spinal individual. We call numbering a labeling method of the individuals in the population that does not encodes the whole lineage of the individual. The numbering choice is arbitrary and we chose to add every new individuals at the end of the hidden list, thus at each division event a parent is chosen uniformly in the population and its child is added at the end of the list. This computation is handled by the function `Tree` explained in Figure 5, where  $i_0$  is the initial numbering of the spinal individual.

From the list  $I_{\text{div}}$ , it is possible to retrieve the lineage of all individuals using any chosen labeling method. In our case we used the Ulam-Harris-Neveu notations and the algorithm generating the list  $U$  of labels from the list  $I$  previously obtained and the initial size of the population  $n$  is detailed Figure 6. The last line of this algorithm lists the labels of individuals in increasing order, grouping siblings together. The label of the spinal individual is thus  $U[E]$ .

To compute a trajectory of the spinal process on the interval  $[0, T]$ , we need to establish the time of jumps and their outcomes. Using the deterministic evolution of the traits and knowing the branching events, it is easy to recover the traits at all time for the individuals. Furthermore, depending on the statistics that one want to evaluate on the system, it might not be necessary to compute the traits of all individuals. For that reason we propose an algorithmic construction of a trajectory of the spinal process, returning the list of jumps times, events and labels of the individuals living at time  $T$ . This algorithm, presented in Figure 7 also distinguishes the spinal

```

function  $(T_{\text{div}}, I_{\text{div}}, E, L_{\text{div}}) = \text{Tree}(t_1, t_2, n, i_0, F_{\text{div}}^{-1})$ 
 $(T_{\text{div}}, N) = ([t_1], n)$ 
% Step 1; generating the division times
while  $T_{\text{div}}[\text{end}] < t_2$  do
  Generate  $u \sim \text{uniform}(0, 1)$ 
  append  $T_{\text{div}}[\text{end}] - \ln(u)/N$  to  $T_{\text{div}}$  ▷ Branching time
   $N \leftarrow N + 1$  ▷ New population size
end while
if  $T_{\text{div}}[\text{end}] > t_2$  then pop  $T_{\text{div}}$ 
% Step 2; generating the fractions at birth and choosing the spinal individual
 $(I_{\text{div}}, E) = ([0, \dots, 0], i_0)$  ▷  $I_{\text{div}}$  of size  $\text{length}(T_{\text{div}})$ 
for  $i \in \{1, \dots, \text{length}(T_{\text{div}}) - 1\}$  do
  Generate  $I \sim \text{uniform}(\{1, \dots, n + i - 1\})$  ▷ Branching individual
   $I_{\text{div}}[i] \leftarrow I$ 
  Generate  $v, q \sim \text{uniform}(0, 1)$ 
  append  $F_{\text{div}}^{-1}(u)$  to  $L_{\text{div}}$  ▷ Fraction  $\lambda$  at birth
  if  $E = i$  and  $p > F_{\text{div}}^{-1}(u)$  then  $E \leftarrow n + i$  ▷ New spinal individual position
end for
end function

```

FIGURE 5. Simulation algorithm of the jumps times until time  $T$ .

```

function  $U = \text{Labels}(I_{\text{div}}, n)$ 
 $N_{\text{div}} = \text{length}(I_{\text{div}})$ 
 $U = [1, \dots, n, 0, \dots, 0]$  ▷  $U$  is of size  $n + N_{\text{div}}$ 
for  $i \in \{1, \dots, N_{\text{div}}\}$  do
   $(U[I_{\text{div}}[i]], U[n + i]) \leftarrow (U[I_{\text{div}}[i]]1, U[I_{\text{div}}[i]]2)$  ▷  $U[I_{\text{div}}[i]] \in \mathcal{U}$ 
end for
 $U \leftarrow \text{QuickSort}(U)$  ▷ sorting algorithm in  $O((N_{\text{div}} + n) \log(N_{\text{div}} + n))$ 
end function

```

FIGURE 6. Labeling function.

individual in the population by returning the label of the spinal individual in the living ones at time  $T$ . To simplify we will denote  $\text{PPP}_{[t_1, t_2]}(c(\cdot))$  the list of times given by a Poisson point process of intensity  $c(\cdot)$  on the interval  $[t_1, t_2]$ , computed using Lewis' thinning algorithm [40].

The output tuple  $(T_{\text{div}}, I_{\text{div}}, E, L_{\text{div}}, T_{\text{loss}}, T_{\text{loss}}^*, I_{\text{loss}}, L_{\text{loss}}, L_{\text{loss}}^*)$  is the minimal information needed to construct the spinal process introduced in Section 4. The list  $U$  of labels of the individuals can be computed with the function  $\text{Labels}(I_{\text{div}}, n)$ . Notice that a lot of operations can be parallelized in this algorithm, unlike in the classical Lewis' algorithm. Moreover, depending on the statistic that one wants to compute on this process, the algorithm can be further simplified. For example, if one takes interest in the total biomass of the population, the indexes of the individuals that branched are not necessary.

**Proposition C.1.** *Let  $T > 0$  and  $T_{\text{div}}^i = (T_{\text{div}}^i, 0 \leq i \leq N_{\text{div}} + 1)$  be a sequence of increasing branching times with  $T_{\text{div}}^0 := 0$  and  $T_{\text{div}}^{N_{\text{div}}+1} := T$ . The size  $\widehat{N}$  of the spinal population starting from  $N_0$  individuals is such that*

$$\int_0^T \widehat{N}_s ds = (N_0 + N_{\text{div}})T - \sum_{i=0}^{N_{\text{div}}} T_{\text{div}}^i,$$



**Input:** model parameters:  $d, F_{\text{div}}^{-1}, F_{\text{loss}}^{-1}, F_{\text{loss}}^{*-1}, T$ , initial condition:  $z = [x^1, \dots, x^n]$ .

**Output:**  $(T_{\text{div}}, I_{\text{div}}, E, L_{\text{div}}, T_{\text{loss}}, T_{\text{loss}}^*, I_{\text{loss}}, L_{\text{loss}}, L_{\text{loss}}^*)$

% Initializing the spine

Generate  $u \sim \text{uniform}(0, 1)$

$i_0 = 1$

**while**  $\sum_{i=1}^{i_0} x^i / (\sum_{i=1}^n x^i) < u$  **do**

$i_0 \leftarrow i_0 + 1$

**end while**

% Generating binary spinal tree on division events

$(T_{\text{div}}, I_{\text{div}}, E, L_{\text{div}}) = \text{Tree}(0, T, n, i_0, F_{\text{div}}^{-1})$

% Generating loss events for the spine and individuals outside the spine

$(T_{\text{loss}}, T_{\text{loss}}^*, I_{\text{loss}}) = ([\cdot], [\cdot], [\cdot])$

**for**  $i \in \{1, \dots, \text{length}(T_{\text{div}})\}$  **do**

% For the individuals outside the spine

Generate  $P \sim \text{PPP}_{[T_{\text{div}}[i-1], T_{\text{div}}[i]]}(K_{\text{loss}}(n+i-1)^2 d(\cdot))$

▷ Times of loss events

append  $P$  to  $T_{\text{loss}}$

Generate  $I_u \sim (\text{uniform}(\{1, \dots, n+i-1\})^{\text{length}(T_{\text{loss}}[i])})$

▷ Distributing loss events

append  $I_u$  to  $I_{\text{loss}}$

% For the spine

Generate  $P^* \sim \text{PPP}_{[T_{\text{div}}[i-1], T_{\text{div}}[i]]}((n+i-1)d(\cdot))$

▷ Times of loss events

append  $P^*$  to  $T_{\text{loss}}^*$

**end for**

Generate  $L_u \sim (\text{uniform}(0, 1))^{\text{length}(T_{\text{loss}})}$

▷ Generating fractions lost

$L_{\text{loss}} = F_{\text{loss}}^{-1}(L_u)$

Generate  $L_u^* \sim (\text{uniform}(0, 1))^{\text{length}(T_{\text{loss}}^*)}$

▷ Generating fractions lost

$L_{\text{loss}}^* = F_{\text{loss}}^{*-1}(L_u^*)$

FIGURE 7. Simulation algorithm of the jumps times until time  $T$

and

$$\int_0^T \widehat{N}_s^2 ds = (N_0 + N_{\text{div}})^2 T - (2N_0 - 1) \sum_{i=1}^{N_{\text{div}}} T_{\text{div}}^i - 2 \sum_{i=1}^{N_{\text{div}}} iT_{\text{div}}^i.$$

*Proof.* The first integral term is directly computed using that

$$\sum_{i=0}^{N_{\text{div}}} i(T_{\text{div}}^{i+1} - T_{\text{div}}^i) = N_{\text{div}}T - \sum_{i=0}^{N_{\text{div}}} T_{\text{div}}^i. \quad (\text{C.1})$$

The second integral term is computed by noticing that

$$\int_0^T \widehat{N}_s^2 ds = N_0^2 T + 2N_0 \sum_{i=0}^{N_{\text{div}}} i(T_{\text{div}}^{i+1} - T_{\text{div}}^i) + \sum_{i=0}^{N_{\text{div}}} i^2 (T_{\text{div}}^{i+1} - T_{\text{div}}^i),$$

and that

$$\sum_{i=0}^{N_{\text{div}}} i^2 (T_{\text{div}}^{i+1} - T_{\text{div}}^i) = N_{\text{div}}^2 T - \sum_{i=1}^{N_{\text{div}}} (2i-1)T_{\text{div}}^i. \quad (\text{C.2})$$

Using (C.1) and (C.2) and rearranging the terms concludes the proof.  $\square$

**Proposition C.2.** For all  $0 \leq i \leq N_{\text{div}}$ , let  $(T_{\text{loss}}^{i,j}, 1 \leq j \leq N_{\text{loss}}^i)$  be the sequence of loss events times in  $[T_{\text{div}}^i, T_{\text{div}}^{i+1})$ . If there is no loss event in this period of time,  $N_{\text{loss}}^i = 0$  and we use the

formalism  $T_{\text{loss}}^{i, N^i_{\text{loss}}} = T_{\text{loss}}^{i, 1} = T_{\text{div}}^{i-1}$ . The fraction of lost masses are denoted by  $\theta_{i,j}$  and the label of the individual that suffered the loss is denoted  $(i, j)$ . Then

$$\int_0^T B_s ds = \sum_{i=0}^{N_{\text{div}}} \left[ \frac{B_{T_{\text{div}}^i}}{\mu} \left( e^{\mu(T_{\text{div}}^{i+1} - T_{\text{div}}^i)} - 1 \right) - \sum_{k=1}^{N^i_{\text{loss}}} \frac{1 - \theta_{i,k}}{\mu} X_{T_{\text{div}}^i}^{(i,k)} \left( e^{\mu(T_{\text{div}}^{i+1} - T_{\text{div}}^i)} - e^{\mu(T_{\text{loss}}^{i,k} - T_{\text{div}}^i)} \right) \right].$$

*Proof.* We recall that  $T_{\text{div}}^{N_{\text{div}}+1} = T$ , and then

$$\int_0^T B_s ds = \sum_{i=0}^{N_{\text{div}}} \left[ \int_{T_{\text{div}}^i}^{T_{\text{loss}}^{i,1}} B_s ds + \sum_{j=2}^{N^i_{\text{loss}}} \int_{T_{\text{loss}}^{i,j-1}}^{T_{\text{loss}}^{i,j}} B_s ds + \int_{T_{\text{loss}}^{i, N^i_{\text{loss}}}}^{T_{\text{div}}^{i+1}} B_s ds \right].$$

Furthermore, for all  $0 \leq i \leq N_{\text{div}}$ , all  $1 \leq j \leq N^i_{\text{loss}} - 1$  and all time  $t \in (T_{\text{loss}}^{i,j}, T_{\text{loss}}^{i,j+1})$ , the total biomass verifies

$$B_t = \left( B_{T_{\text{div}}^i} - \sum_{k=1}^j (1 - \theta_{i,k}) X_{T_{\text{div}}^i}^{(i,k)} \right) \exp(\mu(t - T_{\text{div}}^i)).$$

Thus we integrate on those intervals to get

$$\begin{aligned} \int_0^T B_s ds &= \sum_{i=0}^{N_{\text{div}}} \left[ \frac{B_{T_{\text{div}}^i}}{\mu} e^{-\mu T_{\text{div}}^i} \left( e^{\mu T_{\text{loss}}^{i,1}} - e^{-\mu T_{\text{div}}^i} \right) \right. \\ &\quad + \sum_{j=2}^{N^i_{\text{loss}}} \left( \frac{B_{T_{\text{div}}^i}}{\mu} - \sum_{k=1}^{j-1} \frac{1 - \theta_{i,k}}{\mu} X_{T_{\text{div}}^i}^{(i,k)} \right) e^{-\mu T_{\text{div}}^i} \left( e^{\mu T_{\text{loss}}^{i,j}} - e^{\mu T_{\text{loss}}^{i,j-1}} \right) \\ &\quad \left. + \left( \frac{B_{T_{\text{div}}^i}}{\mu} - \sum_{k=1}^{N^i_{\text{loss}}} \frac{1 - \theta_{i,k}}{\mu} X_{T_{\text{div}}^i}^{(i,k)} \right) e^{-\mu T_{\text{div}}^i} \left( e^{\mu T_{\text{div}}^{i+1}} - e^{\mu T_{\text{loss}}^{i, N^i_{\text{loss}}}} \right) \right]. \end{aligned}$$

Notice that for all  $0 \leq i \leq N_{\text{div}}$  the sum over  $j$  is telescoping for the term in  $B_{T_{\text{div}}^i}$  and thus

$$\begin{aligned} \int_0^T B_s ds &= \sum_{i=0}^{N_{\text{div}}} \left[ \frac{B_{T_{\text{div}}^i}}{\mu} e^{-\mu T_{\text{div}}^i} \left( e^{\mu T_{\text{div}}^{i+1}} - e^{\mu T_{\text{div}}^i} \right) \right. \\ &\quad - \sum_{j=2}^{N^i_{\text{loss}}} \sum_{k=1}^{j-1} \frac{1 - \theta_{i,k}}{\mu} X_{T_{\text{div}}^i}^{(i,k)} e^{-\mu T_{\text{div}}^i} \left( e^{\mu T_{\text{loss}}^{i,j}} - e^{\mu T_{\text{loss}}^{i,j-1}} \right) \\ &\quad \left. - \sum_{k=1}^{N^i_{\text{loss}}} \frac{1 - \theta_{i,k}}{\mu} X_{T_{\text{div}}^i}^{(i,k)} e^{-\mu T_{\text{div}}^i} \left( e^{\mu T_{\text{div}}^{i+1}} - e^{-\mu T_{\text{loss}}^{i, N^i_{\text{loss}}}} \right) \right]. \end{aligned}$$

Inverting the sums over  $j$  and  $k$  in the second line and using the fact that the sum over  $j$  becomes telescoping we get

$$\int_0^T B_s ds = \sum_{i=0}^{N_{\text{div}}} \left[ \frac{B_{T_{\text{div}}^i}}{\mu} e^{-\mu T_{\text{div}}^i} \left( e^{\mu T_{\text{div}}^{i+1}} - e^{\mu T_{\text{div}}^i} \right) - \sum_{k=1}^{N_{\text{loss}}^i - 1} \frac{1 - \theta_{i,k}}{\mu} X_{T_{\text{div}}^i}^{(i,k)} e^{-\mu T_{\text{div}}^i} \left( e^{\mu T_{\text{loss}}^{i, N_{\text{loss}}^i}} - e^{\mu T_{\text{loss}}^i} \right) - \sum_{k=1}^{N_{\text{loss}}^i} \frac{1 - \theta_{i,k}}{\mu} X_{T_{\text{div}}^i}^{(i,k)} e^{-\mu T_{\text{div}}^i} \left( e^{\mu T_{\text{div}}^{i+1}} - e^{\mu T_{\text{loss}}^{i, N_{\text{loss}}^i}} \right) \right].$$

We conclude the proof using the distributive property of the sum and rearranging the terms.  $\square$

Notice that in order to compute  $B_{T_{\text{div}}^i}$ , we need to know the past loss events of each individuals. Thus two approaches are possible: forward or backward. The backward method avoid the storage of the traits of each individuals but is computationally heavy as many calculations are performed multiple times. In fact at every branching event, the past events of the common ancestor are integrated twice. The forward method, on the other hand, avoid these multiplicities but need to store at every branching event the traits of every individual. According to our available RAM, we chose the forward method. This method is also more convenient as the computation of  $\Pi_T$  is then straightforward.

## REFERENCES

- [1] David J. Aldous. Stochastic models and descriptive statistics for phylogenetic trees, from Yule to today. *Statistical Science*, 16(1):23–34, 2001.
- [2] Krishna B. Athreya. Change of measures for Markov chains and the LlogL theorem for branching processes. *Bernoulli*, 6(2):323–338, 2000.
- [3] Vincent Bansaye. Spine for interacting populations and sampling. *Bernoulli*, 30(2):1555 – 1585, 2024.
- [4] Vincent Bansaye and Sylvie Méléard. *Stochastic models for structured populations*, volume 16. Springer, 2015.
- [5] Vincent Bansaye, Jean-François Delmas, Laurence Marsalle, and Viet Chi Tran. Limit theorems for Markov processes indexed by continuous time Galton-Watson trees. *The Annals of Applied Probability*, 21(6):2263–2314, 2011.
- [6] Vincent Bansaye, and Viet Chi Tran. Branching Feller diffusion for cell division with parasite infection. *ALEA: Latin American Journal of Probability and Mathematical Statistics*, 8:95–127, 2011.
- [7] Jean Bertoin and Bastien Mallein. Biggins’ martingale convergence for branching Lévy processes. *Electronic Communications in Probability*, 23:1 – 12, 2018.
- [8] John D. Biggins and Andreas E. Kyprianou. Measure change in multitype branching. *Advances in Applied Probability*, 36(2):544–581, 2004.
- [9] John D Biggins. Martingale convergence in the branching random walk. *Journal of Applied Probability*, 14(1):25–37, 1977.
- [10] Kenneth J. Brown and Song-Sun Lin. On the existence of positive eigenfunctions for an eigenvalue problem with indefinite weight function. *Journal of Mathematical Analysis and Applications*, 75(1):112–120, 1980.
- [11] Nicolas Champagnat and Sylvie Méléard. Invasion and adaptive evolution for individual-based spatially structured populations. *Journal of Mathematical Biology*, 55:147–188, 2007.
- [12] Brigitte Chauvin and Alain Rouault. K-P-P 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.
- [13] Brigitte Chauvin, Alain Rouault, and Anton Wakolbinger. Growing conditioned trees. *Stochastic Processes and their Applications*, 39(1):117–130, 1991.
- [14] David Cheek and Samuel GG Johnston. Ancestral reproductive bias in branching processes. *Journal of Mathematical Biology*, 86(5):70, 2023.
- [15] Raphaël Chetrite and Hugo Touchette. Nonequilibrium Markov processes conditioned on large deviations. *Annales Henri Poincaré*, 16:2005–2057, 2015.
- [16] Bertrand Cloez. Limit theorems for some branching measure-valued processes. *Advances in Applied Probability*, 49(2):549–580, 2017.
- [17] Joel E Cohen. Population growth and earth’s human carrying capacity. *Science*, 269(5222):341–346, 1995.

- [18] Jérôme Coville. On a simple criterion for the existence of a principal eigenfunction of some nonlocal operators. Journal of Differential Equations, 249(11):2921–2953, 2010.
- [19] Pierre Del Moral. Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications. Springer New York, New York, NY, 2004.
- [20] Luc Devroye. General Principles in Random Variate Generation, pages 27–82. Springer New York, New York, NY, 1986.
- [21] Rick Durrett. Probability: theory and examples, volume 49. Cambridge university press, 2019.
- [22] Maren Eckhoff, Andreas E. Kyprianou, and Matthias Winkel. Spines, skeletons and the strong law of large numbers for superdiffusions. The Annals of Probability, 43(5):2545–2610, 2015.
- [23] János Engländer. Branching diffusions, superdiffusions and random media. Probability Surveys, 4:303 – 364, 2007.
- [24] János Engländer, Simon C. Harris, and Andreas E. Kyprianou. Strong Law of Large Numbers for branching diffusions. Annales de l’Institut Henri Poincaré, Probabilités et Statistiques, 46(1):279 – 298, 2010.
- [25] Nicolas Fournier and Sylvie Méléard. A microscopic probabilistic description of a locally regulated population and macroscopic approximations. The Annals of Applied Probability, 14(4):1880 – 1919, 2004.
- [26] Coralie Fritsch, Jérôme Harmand, and Fabien Campillo. A modeling approach of the chemostat. Ecological Modelling, 299:1–13, 2015.
- [27] Hans-Otto Georgii and Ellen Baake. Supercritical multitype branching processes: The ancestral types of typical individuals. Advances in Applied Probability, 35(4):1090–1110, 2003.
- [28] Yoav Git, John W. Harris, and Simon C. Harris. Exponential growth rates in a typed branching diffusion. The Annals of Applied Probability, 17(2):609–653, 2007.
- [29] Robert Hardy and Simon C. Harris. A spine approach to branching diffusions with applications to  $L_p$ -convergence of martingales. Séminaire de Probabilités XLII, 1979:281–330, 2009.
- [30] Simon C. Harris, Samuel G. G. Johnston, and Matthew I. Roberts. The coalescent structure of continuous-time Galton-Watson trees. The Annals of Applied Probability, 30(3): 1368–1414, 2020.
- [31] Simon C. Harris and Matthew I. Roberts. The many-to-few lemma and multiple spines. Annales de l’Institut Henri Poincaré, Probabilités et Statistiques, 53(1):226 – 242, 2017.
- [32] Simon C. Harris and David Williams. Large-deviations and martingales for a typed branching diffusion, 1. Asterisque, 236:133–154, 1996.
- [33] Theodore E. Harris et al. The theory of branching processes, volume 6. Springer Berlin, 1963.
- [34] Benoît Henry, Sylvie Méléard, and Viet Chi Tran. Time reversal of spinal processes for linear and non-linear branching processes near stationarity. Electronic Journal of Probability, 28:1 – 27, 2023.
- [35] Samuel G. G. Johnston and Amaury Lambert. The coalescent structure of uniform and Poisson samples from multitype branching processes. The Annals of Applied Probability, 33(6A): 4820 – 4857, 2023.
- [36] Olav Kallenberg. Stability of critical cluster fields. Mathematische Nachrichten, 77(1):7–43, 1977.
- [37] Harry Kesten and Bernt P. Stigum. A limit theorem for multidimensional Galton-Watson processes. The Annals of Mathematical Statistics, 37(5):1211–1223, 1966.
- [38] Thomas Kurtz, Russell Lyons, Robin Pemantle, and Yuval Peres. A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes. Classical and Modern Branching Processes, 84:181–185, 1997.
- [39] Andreas E. Kyprianou. Travelling wave solutions to the K-P-P equation: alternatives to Simon Harris’ probabilistic analysis. Annales de l’Institut Henri Poincaré (B) Probability and Statistics, 40(1):53–72, 2004.
- [40] Peter A. W Lewis and Gerald S. Shedler. Simulation of nonhomogeneous Poisson processes by thinning. Naval Research Logistics Quarterly, 26(3):403–413, 1979.
- [41] Rong-Li Liu, Yan-Xia Ren, and Renming Song. LlogL criterion for a class of superdiffusions. Journal of Applied Probability, 46(2):479–496, 2009.
- [42] Rong-Li Liu, Yan-Xia Ren, and Renming Song. LlogL condition for supercritical branching Hunt processes. Journal of Theoretical Probability, 24(1):170–193, 2011.
- [43] Rong-Li Liu, Yan-Xia Ren, and Renming Song. Strong Law of Large Numbers for a class of superdiffusions. Acta applicandae mathematicae, 123(1):73–97, 2013.
- [44] Russell Lyons. A simple path to Biggins’ martingale convergence for branching random walk. Classical and Modern Branching Processes, 84:217–221, 1997.
- [45] Russell Lyons, Robin Pemantle, and Yuval Peres. Conceptual proofs of LLogL criteria for mean behavior of branching processes. The Annals of Probability, 23(3):1125–1138, 1995.
- [46] Aline Marguet. Uniform sampling in a structured branching population. Bernoulli, 25(4A):2649 – 2695, 2019.
- [47] Charles Medous. <https://github.com/charles-medous/Spinal-method-for-Yule-model>, 2024.
- [48] Charles Medous. Spinal constructions for continuous type-space branching processes with interactions. arXiv, eprint: 2309.15449, 2024. <https://arxiv.org/abs/2309.15449>
- [49] Roger E. Miles. On the homogeneous planar Poisson point process. Mathematical Biosciences, 6:85–127, 1970.
- [50] Jacques Neveu. Arbres et processus de Galton-Watson. Annales de l’I.H.P. Probabilités et statistiques, 22(2):199–207, 1986.

- [51] Yoshihiko Ogata. On Lewis' simulation method for point processes. *IEEE Transactions on Information Theory*, 27(1):23–31, 1981.
- [52] Pázsit, Imre and Pál, Lénard. *Neutron fluctuations: A treatise on the physics of branching processes*, Elsevier, 2007
- [53] Paul A. Rees. *Chapter 5 - Population Growth*, pages 121–137. Academic Press, 2018.
- [54] Yan-Xia Ren, Renming Song, and Ting Yang. Spine decomposition and LlogL criterion for superprocesses with non-local branching mechanisms. *ALEA Lat. Am. J. Probab. Math. Stat*, 19(1):163–208, 2022.

### **Acknowledgments.**

I would like to thank L. Coquille, A. Marguet and C. Smadi for their useful advises and comments during this work. I thank V. Bansaye for stimulating discussions on the subject of this paper and S. Billiard for fruitful discussions on the Yule model.

This work is supported by the French National Research Agency in the framework of the "France 2030" program (ANR-15-IDEX-0002) and by the LabEx PERSYVAL-Lab (ANR-11-LABX-0025-01). I acknowledge partial support by the Chair "Modélisation Mathématique et Biodiversité" of VEOLIA-Ecole Polytechnique-MNHN-F.X.

CHARLES MEDOUS, UNIV. GRENOBLE ALPES, CNRS, IF, 38000 GRENOBLE, FRANCE;

UNIV. GRENOBLE ALPES, INRIA, 38000 GRENOBLE, FRANCE

*Email address:* `charles.medous@univ-grenoble-alpes.fr`