ECOLE POLYTECHNIQUE

CENTRE DE MATHÉMATIQUES APPLIQUÉES
UMR CNRS 7641

91128 PALAISEAU CEDEX (FRANCE). Tél: 01 69 33 41 50. Fax: 01 69 33 30 11

http://www.cmap.polytechnique.fr/

Slow and fast scales for superprocess limits of age-structured populations

Sylvie Méléard, Chi Viet Tran

R.I. N^0 689 July 2010

Slow and fast scales for superprocess limits of age-structured populations

Sylvie Méléard*and Viet Chi Tran†

July 20, 2010

Abstract

A superprocess limit for an interacting birth-death particle system modelling a population with trait and age-structures is established. Traits of newborn offspring are inherited from the parents except when mutations occur, while ages are set to zero. Because of interactions between individuals, standard approaches based on the Laplace transform do not hold. We use a martingale problem approach and a separation of the slow (trait) and fast (age) scales. While the trait marginals converge in a pathwise sense to a superprocess, the age dynamics, on another time scale, averages to an equilibrium that depends on traits. The convergence of the whole process depending on trait and age, only holds for finite-dimensional time-marginals. We apply our results to the study of examples illustrating different cases of trade-off between competition and senescence.

Keywords: Interacting particle system, age-structure, superprocess, slow and fast scales, trait-structured density-dependent population.
AMS Subject Classification: 60J80, 60K35, 60G57.

1 Introduction

We consider an asexual population in which each individual's ability to survive and reproduce is characterized by a quantitative trait, such as for example the body size, or the rate of food intake. As emphasized by Charlesworth [7], most of these abilities also depend on age. We are interested in this paper to study the joint effects of age and trait structures in the interplay between ecology and evolution. Evolution, acting on the trait distribution of the population, is the consequence of three basic mechanisms: heredity, which transmits traits to new offsprings, mutation, driving a variation in the trait values in the population, and selection between these different trait values, which is due to ecological interactions. Some questions on evolution are strongly related to the age structure. For example, we would like to understand how the age influences the persistence of the population or the trait's evolution or which age structure will appear in long time scales for a given trait. More generally, different life story strategies can be expressed in term of age.

We are interested in the dynamics of large populations composed of small individuals with fast births and deaths, and characterized by quantitative traits and by their physical ages. These models describe allometric demographies with resource constraints: lifetimes and gestation lengths are proportional to individuals' biomasses. The latter are all the more small as the population is large. A main point in our work, biologically and mathematically, is the fact that

*CMAP, Ecole Polytechnique
†Equipe Probabilité Statistique, Laboratoire Paul Painlevé, USTL

at each birth, the age is reset to 0, inducing a large asymmetry between mother and daughters, and a difference of time scales between age and trait. When a mutation occurs, the new mutant trait is close to its ancestor's one, inducing a slow variation of the trait. Hence the dynamics is governed by two different time-scaled phenomena. The other main point is that in our model, interactions between individuals are taken into account, yielding nonlinear mathematical objects.

We start with an individual-centered description of the population dynamics, as point measure-valued process, taking into account all reproductive and death events, and age. Such models have been introduced in Méléard and Tran [23] (see also Ferrière and Tran [13]), and generalize the trait-structured case developed in Champagnat, Ferrière, Méléard [6, 5], and the case with only age in Tran [31].

Our aim is to study an approximation of this model under the setting described above. There are qualitatively different asymptotic behaviors arising from the separation between the fast age and slow trait time scales. While the trait marginals converge in a pathwise sense to a superprocess, the age dynamics, on another time scale, stabilizes in an equilibrium that depends on traits. The convergence of the whole process depending on trait and age, only holds for finite-dimensional time-marginals.

The literature concerning superprocess limits is huge but at our knowledge, nothing has been done in the setting we are interested in, with slow and fast variables, and nonlinearity. The techniques we use are based on martingale properties and generalize to this infinite dimensional setting the treatment of the slow-fast scales phenomena for diffusion processes, developed by Kurtz [20], and by Ball *et al.* [2]. In Champagnat *et al.* [5, 6], superprocess approximations are obtained for large density-dependent populations composed of small individuals, only structured by heritable traits. Evans and Steinsaltz study in [12] the damage segregation at cell fissioning as an explanation of senescence. Nevertheless in their model there is no interaction between cells, and at each birth, the age (as damage) is distributed asymmetrically in each daughter cell following a distribution centered on the mother's age. Thus this age behaves more as a trait than the physical (reseting to 0) age, that we consider here. Athreya *et al.* [1], Bose and Kaj [3, 4], Dawson *et al.* [9], Dynkin [10], Fleischmann *et al.* [14], Kaj and Sagitov [19], Wang [32] study cases without interaction, with techniques based on Laplace characterizations that do not hold anymore when interactions between particles are allowed. Our results generalize Athreya *et al.* [1], Bose and Kaj [3], where an averaging phenomenon is proved in the case where birth and death rates do not depend on age. Starting with particles with exponential lifetimes, they show that in the limit, the age structure stabilizes into an exponential distribution. In our case with dependence, the lifelength of an individual cannot be governed by an age distribution function independent of trait, as in *e.g.* [14, 32], or by a positive continuous additive functional, as in *e.g.* [9, 10, 19].

In Section 2, we start by introducing the microscopic trait and age-structured population model that has been studied in Méléard and Tran [23] and Ferrière and Tran [13]. In Section 3, we establish a limit theorem for large populations with fast births and deaths.

Our main result is stated in Theorem 3.1 where the averaging phenomenon is obtained. In Section 4, we consider as an illustration two models where the population is structured by size and physical or biological age. The physical age measures the time since birth, while the biological age is intrinsic to the individual and may depend on its traits. We see on simulations that larger senescence, which leads to shorter lives for individuals, may result in longer persistence for the population.

Notation: For a given metric space E, we denote by $\mathbb{D}([0,T],E)$ the space of right continuous and left limited (càd-làg) functions from $[0,T]$ to E. This space is embedded with the Skorohod topology (*e.g.* [26, 17]).

If $\mathcal{X}$ is a subset of $\mathbb{R}^d$, we denote by $\mathcal{M}_F(\mathcal{X})$ the set of finite measures on $\mathcal{X}$, which will be usually embedded with the topology of weak convergence. Nevertheless, if $\mathcal{X}$ is unbounded, we will also consider the topology of vague convergence. If we need to differentiate both topological spaces, we will denote by $(\mathcal{M}_F(\mathcal{X}), w)$, respectively $(\mathcal{M}_F(\mathcal{X}), v)$, the space of measures endowed by the weak (resp. vague) topology. For a measurable real bounded function f, and a measure $\mu \in \mathcal{M}_F(\mathcal{X})$, we will denote

$$\langle \mu, f \rangle = \int_{\mathcal{X}} f(x)\mu(dx).$$

For $\ell \in \mathbb{N}$, we denote by $\mathcal{C}_b^\ell(\mathcal{X}, \mathbb{R})$ the space of real bounded functions f of class $\mathcal{C}^\ell$ with bounded derivatives. In the sequel, the space $\mathcal{C}_b^{0,1}(\mathcal{X} \times \mathbb{R}_+, \mathbb{R})$ denotes the space of continuous bounded real functions $\varphi(x,a)$ on $\mathcal{X} \times \mathbb{R}_+$ of class $\mathcal{C}^1$ with respect to a with bounded derivatives.

2 Microscopic age and trait structured particle system

We consider a discrete population in continuous time where the individuals reproduce, age and die with rates depending on a hereditary trait and on their age. An individual is characterized by a quantitative trait $x \in \mathcal{X}$ where $\mathcal{X}$ is a closed subset of $\mathbb{R}^d$ and by its physical age $a \in \mathbb{R}_+$, *i.e.* the time since its birth. The individuals reproduce asexually during their lives, and the trait from the parent is transmitted to its offspring except when a mutation occurs. The individuals compete for resources. This interaction between individuals involves selection pressure and senescence. It is described by a kernel comparing the competitors' traits and ages.

We are interested in approximating the dynamics of a large population whose size is parametrized by some integer n. This parameter can be seen as the order of the carrying capacity, when the total amount of resources is assumed to be fixed. If the parameter n is large, there will be many individuals with little per capita resource and we renormalize the individual biomass by the weight $1/n$.
We consider here allometric demographies where the lifetime and gestation length of each individual are proportional to its biomass. Thus the birth and death rates are of order n, while preserving the demographic balance. As a consequence the right scale to observe a nontrivial limit in the age structure, as n increases, is of order $1/n$.

The population at time t is represented by a point measure as follows:

$$X_t^n = \frac{1}{n} \sum_{i=1}^{N_t^n} \delta_{(X_i(t), A_i(t))}, \tag{2.1}$$

where $N_t^n = \langle nX_t^n, 1 \rangle$ is the number of individuals alive at time t, and $X_i(t)$ and $A_i(t)$ denote respectively the trait and age of individual i at time t (individuals are ranked in lexicographical order for instance).

The dynamics of X^n is given as follows:

- The birth of an individual with trait $x \in \mathcal{X}$ and age $a \in \mathbb{R}_+$ is given by $n\, r(x,a) + b(x,a)$. The new offspring is of age 0 at birth. Moreover, it inherits of the trait x of its ancestor with probability $1 - p(x,a) \in [0,1]$ and is a mutant with probability $p(x,a) \in [0,1]$. The mutant trait is then $x + h$, where the variation h is randomly chosen following the distribution $\pi^n(x, dh)$.

- Individuals age with velocity n, so that the physical age at time t of an individual born at time c is $a = n(t - c)$.

- The intrinsic death rate of an individual with trait $x \in \mathcal{X}$ and age $a \in \mathbb{R}_+$ is given by $n\,r(x,a) + d(x,a)$. The competition between individuals (x,a) and (y,α) is described by the value $U((x,a),(y,\alpha))$ of a kernel U. In a population described by the measure $X \in \mathcal{M}_F(\mathcal{X} \times \mathbb{R}_+)$, the total interaction on an individual (x,a) is thus:

$$XU(x,a) = \int_{\mathcal{X}\times\mathbb{R}_+} U((x,a),(y,\alpha))X(dy,d\alpha), \tag{2.2}$$

and its total death rate is $n\,r(x,a) + d(x,a) + XU(x,a)$.

Assumption 2.1. 1. The birth and death rates b and d are continuous on $\mathcal{X} \times \mathbb{R}_+$ and bounded respectively by $\bar{b}$ and $\bar{d}$.

2. The function r is continuous on $\mathcal{X} \times \mathbb{R}_+$ and there exist two positive constants $\underline{r}$ and $\bar{r}$ such that $\forall (x,a) \in \mathcal{X} \times \mathbb{R}_+,\ \underline{r} \leq |r(x,a)| \leq \bar{r}$.

3. The competition kernel U is continuous on $(\mathcal{X} \times \mathbb{R}_+)^2$ and is bounded by $\bar{U}$. □

Assumption 2.1-(2) implies that every individual has a finite lifetime that is stochastically upper-bounded by an exponential of parameter $n\underline{r}$.

If the competition kernel U is positive on $(\mathcal{X} \times \mathbb{R}_+)^2$, it can model a competition of the logistic type: the more important the size of the population is and the higher the death rate by competition is. For examples of such kernels we refer to [23].

Assumption 2.2. For any $x \in \mathcal{X}$, the mutation kernel $\pi^n(x,dh)$ has its support in $\mathcal{X} - \{x\} = \{h \in \mathbb{R}^d \,|\, x+h \in \mathcal{X}\}$. We consider two cases:

1. The trait space $\mathcal{X}$ is a compact subset of $\mathbb{R}^d$ and there exists a generator A of a Feller semi-group on $\mathcal{C}_b(\mathcal{X},\mathbb{R})$ with domain $\mathcal{D}(A)$ dense in $\mathcal{C}_b(\mathcal{X},\mathbb{R})$ such that:

$$\forall f \in \mathcal{D}(A), \qquad \lim_{n\to+\infty} \sup_{x\in\mathcal{X}} \left| n \int_{\mathcal{X}-\{x\}} \big(f(x+h) - f(x)\big)\pi^n(x,dh) - Af(x) \right| = 0. \tag{2.3}$$

2. The trait space $\mathcal{X}$ is a closed subset of $\mathbb{R}^d$ and we assume in addition that there exists $\ell_1 \geq \ell_0 \geq 2$ with $\mathcal{C}_b^{\ell_1}(\mathcal{X},\mathbb{R}) \subset \mathcal{D}(A)$ and $\forall f \in \mathcal{C}_b^{\ell_1}(\mathcal{X},\mathbb{R})$, $\forall x \in \mathcal{X}$.

$$|Af(x)| \leq C \sum_{\substack{|k|\leq \ell_0 \\ k=(k_1,\dots,k_d)}} |D^k f(x)| \tag{2.4}$$

and

$$\sup_{x\in\mathcal{X}} \left| n \int_{\mathcal{X}-\{x\}} \big(f(x+h) - f(x)\big)\pi^n(x,dh) - Af(x) \right| \leq \varepsilon_n \sum_{\substack{|k|\leq \ell_1 \\ k=(k_1,\dots,k_d)}} \|D^k f\|_\infty, \tag{2.5}$$

where $D^k f(x) = \partial_{x_1}^{k_1} \dots \partial_{x_d}^{k_d} f(x)$, ε_n is a sequence tending to 0 as n tends to infinity and C is a constant.

Remark 2.3. Both Assumptions (2.3) and (2.5) describe small mutation steps. The stronger hypothesis (2.5) is required when $\mathcal{X}$ is not compact, to obtain the tightness in the proof of Theorem 3.1.

Example 2.4. Let us give some examples of mutation kernels satisfying (2.3) or (2.4) and (2.5).

1. In the case where $\mathcal{X} = [x_1, x_2]$, the mutation kernel $\pi^n(x, dh)$ is a Gaussian distribution with mean 0 and variance σ^2/n, conditioned to $[x_1 - x, x_2 - x]$. In this case, elementary computation shows that for $f \in \mathcal{C}_b^2([x_1, x_2], \mathbb{R})$ such that $f'(x_1) = f'(x_2) = 0$, $Af(x) = \frac{\sigma^2}{2} f''(x)$, which satisfies (2.3).

2. In the case where $\mathcal{X} = \mathbb{R}^d$, a possible choice of mutation kernel $\pi^n(x, dh)$ is a Gaussian distribution with mean 0 and covariance matrix $\Sigma(x)/n$, with $\Sigma(x) = (\Sigma_{ij}(x), 1 \leq i, j \leq d)$. The generator A is given for $f \in \mathcal{C}_b^2(\mathbb{R}^d, \mathbb{R})$ by $Af(x) = \frac{1}{2} \sum_{i,j=1}^d \Sigma_{ij}(x) \partial^2_{ij} f(x)$. If the function Σ is bounded, then Assumption (2.4) is fulfilled. If moreover, the third moments of $\pi^n(x, dh)$ are bounded (in x), then (2.5) is satisfied. □

Let us now describe the generator L^n of the $\mathcal{M}_F(\mathcal{X} \times \mathbb{R}_+)$-valued Markov process X^n, which sums the aging phenomenon and the ecological dynamics of the population. As developed in Dawson [8] Theorem 3.2.6, the set of cylindrical functions defined for each $\mu \in \mathcal{M}_F(\mathcal{X} \times \mathbb{R}_+)$ by $F_\varphi(\mu) = F(\langle \mu, \varphi \rangle)$, with $F \in \mathcal{C}_b^1(\mathbb{R}, \mathbb{R})$ and $\varphi \in \mathcal{C}_b^{0,1}(\mathcal{X} \times \mathbb{R}_+, \mathbb{R})$, generates the set of bounded measurable functions on $\mathcal{M}_F(\mathcal{X} \times \mathbb{R}_+)$. For such function, for $\mu \in \mathcal{M}_F(\mathcal{X} \times \mathbb{R}_+)$,

$$\begin{aligned} L^n F_\varphi(\mu) = {} & n \langle \mu, \partial_a \varphi(.) \rangle F'_\varphi(\mu) \\ & + n \int_{\mathcal{X} \times \mathbb{R}_+} \Big[\big(nr(x,a) + d(x,a) + \mu U(x,a) \big) \Big(F_\varphi\big(\mu - \frac{1}{n} \delta_{(x,a)}\big) - F_\varphi(\mu) \Big) \\ & + \big(nr(x,a) + b(x,a) \big) \int_{\mathbb{R}^d} \Big(F_\varphi\big(\mu + \frac{1}{n} \delta_{(x+h,0)}\big) - F_\varphi(\mu) \Big) K^n(x,a,dh) \Big] \, \mu(dx, da), \end{aligned} \tag{2.6}$$

where

$$K^n(x, a, dh) = p(x,a) \; \pi^n(x, dh) + (1 - p(x,a)) \; \delta_0(dh). \tag{2.7}$$

Under the condition $\sup_{n \in \mathbb{N}^*} \mathbb{E}\left(\langle X_0^n, 1 \rangle\right) < +\infty$, it has been proved in Méléard-Tran [23] (see also the case without age in Fournier-Méléard [16] and the case without trait in Tran [31]), that there exists for any n, a càd-làg Markov process with generator L^n, which can be obtained as solution of a stochastic differential equation driven by a Point Poisson measure. Trajectorial uniqueness also holds for this equation. The construction provides an exact individual-based simulation algorithm (see [13]).

A slight adaptation of the proofs in [16] allows us to get the

Proposition 2.5. *(i) Under Assumptions 2.1, and if*

$$\sup_{n \in \mathbb{N}^*} \mathbb{E}(\langle X_0^n, 1 \rangle^3) < +\infty, \tag{2.8}$$

then for all $T > 0$,

$$\sup_{n \in \mathbb{N}^*} \mathbb{E}\Big(\sup_{t \in [0,T]} \langle X_t^n, 1 \rangle^3 \Big) < +\infty. \tag{2.9}$$

(ii) Moreover, for $n \in \mathbb{N}^*$ and a test function $\varphi \in \mathcal{C}_b^{0,1}(\mathcal{X} \times \mathbb{R}_+, \mathbb{R})$, the process $M^{n,\varphi}$ defined by

$$\begin{aligned} M_t^{n,\varphi} = {} & \langle X_t^n, \varphi \rangle - \langle X_0^n, \varphi \rangle - n \int_0^t \langle X_s^n, \partial_a \varphi(x,a) \rangle \, ds \\ & - \int_0^t \int_{\mathcal{X} \times \mathbb{R}_+} \Big(\big(nr(x,a) + b(x,a) \big) \int_{\mathbb{R}^d} \varphi(x+h, 0) K^n(x,a,dh) \\ & - \big(nr(x,a) + d(x,a) + X_s^n U(x,a) \big) \varphi(x,a) \Big) X_s^n(dx, da) \, ds \end{aligned} \tag{2.10}$$

is a square integrable martingale started at 0 with quadratic variation:

$$\langle M^{n,\varphi}\rangle_t = \frac{1}{n}\int_0^t\int_{\mathcal{X}\times\mathbb{R}_+}\Big(\big(nr(x,a)+b(x,a)\big)\int_{\mathbb{R}^d}\varphi^2(x+h,0)K^n(x,a,dh) + \big(nr(x,a)+d(x,a)+X^n_sU(x,a)\big)\varphi^2(x,a)\Big)X^n_s(dx,da)\,ds. \quad (2.11)$$

Notice that in (2.10), the mutation rate is hidden in the kernel K^n:

$$\big(nr(x,a)+b(x,a)\big)\int_{\mathbb{R}^d}\varphi(x+h,0)K^n(x,a,dh) = \big(nr(x,a)+b(x,a)\big)(1-p(x,a))\varphi(x,0) + \big(nr(x,a)+b(x,a)\big)p(x,a)\int_{\mathbb{R}^d}\varphi(x+h,0)\pi^n(x,dh). \quad (2.12)$$

3 Superprocess limit

We now investigate the limit when n increases to $+\infty$. In the limit, we obtain a continuum of individuals in which the individualities are lost. It is in particular difficult to keep track of the age-distribution when n tends to infinity. Because of the non-local branching (a mother of age $a > 0$ gives birth to a daughter of age 0) and because the aging velocity tends to infinity, it is impossible to obtain the uniform tightness of the sequence of measure-valued process $(X^n_{\cdot}(dx,da))_{n\in\mathbb{N}^*}$, as it can be observed considering (2.10). Assuming that the function φ only depends on a, the term of the form

$$\int_0^t\int_{\mathcal{X}\times\mathbb{R}_+} nr(x,a)\left(\varphi(0)-\varphi(a)\right)X^n_s(dx,da)ds$$

cannot be tight if $r(x,a)$ is bounded and X^n is tight. Therefore, we will be led to firstly show the uniform tightness of the trait marginal of the process X^n and then to prove that in the limit, an averaging phenomenon appears for the age dynamics. Indeed, this "fast" evolving component stabilizes in an equilibrium that depends on the dynamics of the "slow" trait component.

We generalize to measure-valued processes, averaging techniques of Ball *et al.* [2], Kurtz [20] for diffusion processes. A specificity in our case is that the fast-scaling is related to time, since age is involved. In addition, notice that the competition between individuals creates a large dependence between the age and trait distributions. At our knowledge this dependence has never been investigated before in the literature.

Let us introduce the marginal $\bar{X}^n_t(dx)$ of $X^n_t(dx,da)$ defined for any bounded and measurable function f on $\mathcal{X}$ and for any $t\in\mathbb{R}_+$ by

$$\int_{\mathcal{X}} f(x)\bar{X}^n_t(dx) = \int_{\mathcal{X}\times\mathbb{R}_+} f(x)X^n_t(dx,da). \quad (3.1)$$

Our main result states the convergence of the sequence $(\bar{X}^n)_{n\in\mathbb{N}^*}$ to a nonlinear super-process. The nonlinearity remains at the slow time scale in the growth rate, which is preserved in this asymptotics. Moreover, fast mutations are compensated by small mutation steps. Fast births and deaths provide stochastic fluctuations in the limit.

Theorem 3.1. *Assume Hypotheses 2.1 and 2.2, (2.8) and assume that there exists $X_0\in\mathcal{M}_F(\mathcal{X}\times\mathbb{R}_+)$ such that $\lim_{n\to+\infty}X^n_0 = X_0$ in $(\mathcal{M}_F(\mathcal{X}\times\mathbb{R}_+),w)$.*

For any $x\in\mathcal{X}$, let us introduce the age probability density

$$\widehat{m}(x,a) = \frac{\exp\big(-\int_0^a r(x,\alpha)d\alpha\big)}{\int_0^{+\infty}\exp\big(-\int_0^a r(x,\alpha)d\alpha\big)da}. \quad (3.2)$$

Then, for each $T>0$, the sequence $(\bar{X}^n)_{n\in\mathbb{N}^*}$ converges in law in $\mathbb{D}([0,T],(\mathcal{M}_F(\mathcal{X}),w))$ to the unique superprocess $\bar{X}\in\mathcal{C}([0,T],\mathcal{M}_F(\mathcal{X}))$ such that for any function $f\in\mathcal{D}(A)$,

$$\begin{aligned}M_t^f=\langle\bar{X}_t,f\rangle-\langle\bar{X}_0,f\rangle-&\int_0^t\int_{\mathcal{X}}\Big(\widehat{(p\,r)}(x)Af(x)\\&+\big[\widehat{b}(x)-\big(\widehat{d}(x)+\bar{X}_s\widehat{U}(x)\big)\big]f(x)\Big)\bar{X}_s(dx)\,ds\end{aligned}\tag{3.3}$$

is a square integrable martingale with quadratic variation:

$$\langle M^f\rangle_t=\int_0^t\int_{\mathcal{X}}2\widehat{r}(x)f^2(x)\bar{X}_s(dx)\,ds.\tag{3.4}$$

Here, any $\widehat{\psi}(x)$ is defined for a bounded function $\psi(x,a)$ by

$$\widehat{\psi}(x)=\int_{\mathbb{R}_+}\psi(x,a)\widehat{m}(x,a)da,$$

and $\bar{X}_t\widehat{U}(x)$ is given by

$$\bar{X}_t\widehat{U}(x)=\int_{\mathcal{X}}\left(\int_{\mathbb{R}_+}\int_{\mathbb{R}_+}U((x,a),(y,\alpha))\widehat{m}(y,\alpha)d\alpha\ \widehat{m}(x,a)da\right)\bar{X}_t(dy).$$

□

Theorem 3.1 states that in the limit, an averaging phenomenon happens and the "fast" age component finally submits to the dynamics of the "slow" trait component. Since the fast-scaling (involving age) is related to the time, the stable age distribution $\widehat{m}(x,a)da$ given in (3.2) is obtained for each trait x as the long time limit in the age-structured population where all coefficients except $r(x,a)$ are zero.

Before proving this slow-fast limit theorem, let us insist on the main difficulty created by the interacting competition mechanism. Indeed, the branching property fails and it impedes the use of Laplace-transform techniques, as it had been almost systematically done in the papers in the past studying particle pictures with age-structure. Our model generalizes the age-structure population process studied in Athreya *et al.* [1], Bose and Kaj [3], in which birth and death rates are equal to a constant λ. In that case, the limiting behaviour of the renormalized critical birth and death process appears as a particular case of Theorem 3.1 with $\widehat{m}(x,a)da=\lambda e^{-\lambda a}da$. In Dawson *et al.* [9], Dynkin [10], and Kaj and Sagitov [19], the age dependence is modelled through an additive functional of the motion process. In that way, the age "accumulates" along the lineage. In our case, the age is set to zero at each birth, inducing a renewal phenomenon. The life-length does not have a fixed probability distribution anymore, unless there is no interaction. In [4], the authors consider a particle system with a different scaling, which favors large reproduction events. The limit in this case is not a superprocess any more but behaves as the solution of a McKendrick-Forster equation perturbed by random immigration events created by the large rare birth events.

The proof of Theorem 3.1 is the aim of Section 3. We first establish the tightness of the sequence $(\bar{X}^n)_{n\in\mathbb{N}^*}$ (Section 3.1). We identify the limiting values as solutions of the martingale problem given in (3.3), (3.4), proving the convergence of the time marginals $(X_t^n(dx,da))_n$, for any fixed t, to a limit involving the age averaging (Section 3.2). Uniqueness is then showed to conclude.

3.1 Tightness of $(\bar{X}^n)_{n\in\mathbb{N}^*}$

In this subsection, we shall prove that:

Proposition 3.2. *Assume Hypotheses 2.1 and 2.2, and (2.8). If the sequence $(\bar{X}_0^n)_{n\in\mathbb{N}^*}$ is uniformly tight in $(\mathcal{M}_F(\mathcal{X}), w)$, then the sequence $(\mathcal{L}(\bar{X}^n))_{n\in\mathbb{N}^*}$ is uniformly tight in the space of probability measures on $\mathbb{D}([0,T], (\mathcal{M}_F(\mathcal{X}), w))$.*

Proof. Recall firstly that for a measurable and bounded function f on $\mathcal{X}$, the process

$$M_t^{n,f} = \langle X_t^n, f\rangle - \langle X_0^n, f\rangle - \int_0^t \int_{\mathcal{X}\times\mathbb{R}_+} \Big(\big(nr(x,a)+b(x,a)\big)\int_{\mathbb{R}^d} f(x+h)K^n(x,a,dh) \\ - \big(nr(x,a)+d(x,a)+X_s^nU(x,a)\big)f(x)\Big)X_s^n(dx,da)\,ds \quad (3.5)$$

is a square integrable martingale started at 0 with quadratic variation

$$\langle M^{n,f}\rangle_t = \frac{1}{n}\int_0^t \int_{\mathcal{X}\times\mathbb{R}_+} \Big(\big(nr(x,a)+b(x,a)\big)\int_{\mathbb{R}^d} f^2(x+h)K^n(x,dh) \\ + \big(nr(x,a)+d(x,a)+X_s^nU(x,a)\big)f^2(x)\Big)X_s^n(dx,da)\,ds. \quad (3.6)$$

We divide the proof into several steps.

Step 1 Firstly, we prove the uniform tightness of $(\mathcal{L}(\bar{X}^n))_{n\in\mathbb{N}^*}$ in the space of probability measures on $\mathbb{D}([0,T], (\mathcal{M}_F(\mathcal{X}), v))$.
Let us consider a continuous bounded function $f \in \mathcal{D}(A)$, and show the uniform tightness of the sequence $(\langle \bar{X}_.^n, f\rangle)_{n\in\mathbb{N}^*}$ in $\mathbb{D}([0,T],\mathbb{R})$. We remark that for every fixed $t\in[0,T]$ and $n>0$,

$$\mathbb{P}(|\langle \bar{X}_t^n, f\rangle| > k) \le \frac{\|f\|_\infty\,\mathbb{E}(\sup_{t\in[0,T]}\langle \bar{X}_t^n, 1\rangle)}{k}, \quad (3.7)$$

which tends to 0 as k tends to infinity (cf. (2.9)). This proves the tightness of the family of time-marginals $(\langle \bar{X}_t^n, f\rangle)_{n\in\mathbb{N}^*}$. Denoting by $A^{n,f}$ the finite variation process in the r.h.s. of (3.5) and thanks to Assumption 2.2, we get for all stopping times $S_n < T_n < (S_n+\delta)\wedge T$, that

$$\mathbb{E}(|A_{T_n}^{n,f} - A_{S_n}^{n,f}|) \le \delta\,\Big[\Big((\|Af(s,.)\|_\infty + 1)\frac{\bar{r}}{2} + \|f\|_\infty(\bar{b}+\bar{d})\Big)\sup_{n\in\mathbb{N}^*}\mathbb{E}(\sup_{t\in[0,T]}\langle X_t^n, 1\rangle) \\ + \bar{U}\sup_{n\in\mathbb{N}^*}\mathbb{E}(\sup_{t\in[0,T]}\langle X_t^n, 1\rangle^2)\Big]. \quad (3.8)$$

The quadratic variation process (3.6) satisfies a similar inequality:

$$\begin{aligned}
&\mathbb{E}(|\langle M^{n,f}\rangle_{T_n} - \langle M^{n,f}\rangle_{S_n}|)\\
\le&\|f\|_\infty^2\mathbb{E}\Big(\int_{S_n}^{T_n}\Big[2\bar{r}\langle \bar{X}_s^n, 1\rangle + \frac{\bar{b}\langle X_s^n, 1\rangle}{n} + \frac{\bar{d}\langle \bar{X}_s^n, 1\rangle + \bar{U}\langle \bar{X}_s^n, 1\rangle^2}{n}\Big]ds\Big)\\
\le&\|f\|_\infty^2\delta\,\Big[\Big(2\bar{r}+\frac{\bar{b}+\bar{d}}{n}\Big)\sup_{n\in\mathbb{N}^*}\mathbb{E}(\sup_{t\in[0,T]}\langle \bar{X}_t^n, 1\rangle) + \frac{\bar{U}}{n}\sup_{n\in\mathbb{N}^*}\mathbb{E}(\sup_{t\in[0,T]}\langle \bar{X}_t^n, 1\rangle^2)\Big].
\end{aligned} \quad (3.9)$$

Then, for $\varepsilon>0$, $\eta>0$, a sufficiently large n and small δ, we have using (2.9) and Assumption 2.2, that

$$\mathbb{P}(|A_{T_n}^{n,f} - A_{S_n}^{n,f}| > \eta) \le \varepsilon \quad \text{and} \quad \mathbb{P}(|\langle M^{n,f}\rangle_{T_n} - \langle M^{n,f}\rangle_{S_n}| > \eta) \le \varepsilon. \quad (3.10)$$

From (3.7), (3.8), (3.9) and the Aldous-Rebolledo criterion (see *e.g.* [17] or [11, Th. 1.17]), we obtain the uniform tightness of the sequence $(\langle \bar{X}^n_{\cdot}, f\rangle)_{n\in\mathbb{N}^*}$ in $\mathbb{D}([0,T],\mathbb{R})$. Thanks to Roelly's criterion [27], we conclude that $(\bar{X}^n)_{n\in\mathbb{N}^*}$ is uniformly tight in $\mathbb{D}([0,T],(\mathcal{M}_F(\mathcal{X}),v))$.

Let us denote by $\bar{X}$ a limiting process of $(\bar{X}^n)_{n\in\mathbb{N}^*}$. It is almost surely (a.s.) continuous in $(\mathcal{M}_F(\mathcal{X}),v)$ and $(\mathcal{M}_F(\mathcal{X}),w)$ since

$$\sup_{t\in\mathbb{R}_+}\sup_{f,\,\|f\|_\infty\le 1}|\langle \bar{X}^n_t, f\rangle - \langle \bar{X}^n_{t_-}, f\rangle| \le \frac{1}{n}. \tag{3.11}$$

In the case where $\mathcal{X}$ is a compact subset of $\mathbb{R}^d$, the vague and weak topologies coincide, which fails in the non-compact case. Nevertheless the tightness in $\mathbb{D}([0,T],(\mathcal{M}_F(\mathcal{X}),w))$ is needed to identify the limiting values of $(\bar{X}^n)_{n\in\mathbb{N}^*}$.

Step 2 Let us now concentrate on the case where $\mathcal{X}$ is unbounded and let us show the tightness of $(\bar{X}^n)_{n\in\mathbb{N}^*}$ in $\mathbb{D}([0,T],(\mathcal{M}_F(\mathcal{X}),w))$. The same computation as in Step 1 for $f(x)=1$ implies that the sequence $(\langle \bar{X}^n, 1\rangle)_{n\in\mathbb{N}^*}$ is uniformly tight in $\mathbb{D}([0,T],\mathbb{R}_+)$. As a consequence, it is possible to extract from $(\bar{X}^n)_{n\in\mathbb{N}^*}$ a subsequence $(\bar{X}^{u_n})_{n\in\mathbb{N}^*}$ such that:

- $(\bar{X}^{u_n})_{n\in\mathbb{N}^*}$ converges in distribution to $\bar{X}$ in $\mathbb{D}([0,T],(\mathcal{M}_F(\mathcal{X}),v))$,
- $(\langle \bar{X}^{u_n}, 1\rangle)_{n\in\mathbb{N}^*}$ converges in distribution in $\mathbb{D}([0,T],\mathbb{R}_+)$.

Let us now show that the limit of $(\langle \bar{X}^{u_n}, 1\rangle)_{n\in\mathbb{N}^*}$ is $\langle \bar{X}, 1\rangle$, empedding a loss of mass in the limit. Indeed, as a consequence, a criterion in Méléard and Roelly [22] will prove that $(\bar{X}^{u_n})_{n\in\mathbb{N}^*}$ converges in distribution to $\bar{X}$ in $\mathbb{D}([0,T],(\mathcal{M}_F(\mathcal{X}),w))$.

By simplicity, we will again denote u_n by n.

As in Jourdain-Méléard [18], we introduce a sequence of smooth functions ψ_k defined on $\mathbb{R}_+$ and approximating $\mathbf{1}_{\{u\ge k\}}$. For $k\in\mathbb{N}$, let $\psi_k(u)=\psi(0\vee(u-(k-1))\wedge 1)$ where $\psi(y)=6y^5-15y^4+10y^3$ is a nondecreasing function such that $\psi(0)=\psi'(0)=\psi''(0)=1-\psi(1)=\psi'(1)=\psi''(1)=0$. The function $u\mapsto\psi_k(u)$ is nondecreasing on $\mathbb{R}_+$, equals 0 on $[0,k-1]$ and 1 on the complement of $[0,k)$. In particular $\psi_0\equiv 1$. Moreover the sequence $(\psi_k)_{k\in\mathbb{N}^*}$ is nonincreasing, and satisfies for $u\ge 0$ and $p\ge 1$ that

$$\begin{aligned}\mathbf{1}_{\{u\ge k\}} &\le \psi_k(u)\le \mathbf{1}_{\{u\ge k-1\}};\\ \psi_k^{(p)}(u) &\le \sup_{u\in[k-1,k]}|\psi_k^{(p)}(u)|\,\mathbf{1}_{\{u\ge k-1\}}\le \sup_{u\in[k-1,k]}|\psi_k^{(p)}(u)|\,\psi_{k-1}(u).\end{aligned} \tag{3.12}$$

The proof of the following lemma is postponed at the end of Proposition 3.2's proof. We define $f_k(x)=\psi_k(\|x\|)$, for all $x\in\mathcal{X}$.

Lemma 3.3. *Under the assumptions of Proposition 3.2,*

$$\lim_{k\to+\infty}\limsup_{n\to+\infty}\mathbb{E}\left(\sup_{t\le T}\langle \bar{X}^n_t, f_k\rangle\right)=0.$$

From Lemma 3.3, we can deduce that

$$\lim_{k\to+\infty}\mathbb{E}\Big(\sup_{t\le T}\langle \bar{X}_t, f_k\rangle\Big)=0. \tag{3.13}$$

Indeed, for $k \in \mathbb{N}$, the continuous and compactly supported functions $(f_{k,\ell} \overset{\text{def}}{=} f_k(1-f_\ell))_{\ell\in\mathbb{N}}$ increase to f_k as $\ell \to +\infty$. By continuity of $\nu \mapsto \sup_{t\leq T}\langle \nu_t, f_{k,\ell}\rangle$ on $\mathbb{D}([0,T],(M_F,v))$ and uniform integrability deduced from the uniform square moment estimates (2.9), one has

$$\mathbb{E}\Big(\sup_{t\leq T}\langle \bar{X}_t, f_{k,\ell}\rangle\Big) = \lim_{n\to+\infty}\mathbb{E}\Big(\sup_{t\leq T}\langle \bar{X}^n_t, f_{k,\ell}\rangle\Big) \leq \liminf_{n\to+\infty}\mathbb{E}\Big(\sup_{t\leq T}\langle \bar{X}^n_t, f_k\rangle\Big).$$

Taking the limit $\ell \to +\infty$ in the left-hand-side by the monotone convergence theorem, one concludes that for $k=0$

$$\mathbb{E}\Big(\sup_{t\leq T}\langle \bar{X}_t, 1\rangle\Big) = \mathbb{E}\Big(\sup_{t\leq T}\langle \bar{X}_t, f_0\rangle\Big) < +\infty, \tag{3.14}$$

and from Lemma 3.3, that (3.13) holds for any k.
As a consequence one may extract a subsequence of $(\sup_{t\leq T}\langle \bar{X}_t, f_k\rangle)_k$ that converges to 0 a.s.
We can now prove the convergence of $\langle X^{u_n}, 1\rangle$ to $\langle \bar{X}, 1\rangle$. For F a Lipschitz continuous and bounded function on $\mathbb{D}([0,T],\mathbb{R})$, we have

$$\begin{aligned}\limsup_{n\to+\infty}|\mathbb{E}\big(F(\langle \bar{X}^n,1\rangle)-F(\langle \bar{X},1\rangle)\big)| &\leq \limsup_{k\to+\infty}\limsup_{n\to+\infty}|\mathbb{E}\big(F(\langle \bar{X}^n,1\rangle)-F(\langle \bar{X}^n,1-f_k\rangle)\big)|\\ &+\limsup_{k\to+\infty}\limsup_{n\to+\infty}|\mathbb{E}\big(F(\langle \bar{X}^n,1-f_k\rangle)-F(\langle \bar{X},1-f_k\rangle)\big)|\\ &+\limsup_{k\to+\infty}|\mathbb{E}\big(F(\langle \bar{X},1-f_k\rangle)-F(\langle \bar{X},1\rangle)\big)|.\end{aligned}$$

Since $|F(\langle \nu, 1-f_k\rangle)-F(\nu,1)\rangle| \leq C\sup_{t\leq T}\langle \nu_t, f_k\rangle$ by Lipschitz property, the first and the third terms in the r.h.s. are equal to 0 respectively according to Lemma 3.3 and to (3.13). The second term is 0 by continuity of $\nu \mapsto \langle \nu, 1-f_k\rangle$ in $\mathbb{D}([0,T],(\mathcal{M}_F(\mathcal{X}),v))$.
This ends the proof of Proposition 3.2. ∎

Proof of Lemma 3.3. Firstly, let us show that for each $t\in[0,T]$,

$$\lim_{k\to+\infty}\limsup_{n\to+\infty}\mathbb{E}\left(\langle \bar{X}^n_t, f_k\rangle\right) = 0. \tag{3.15}$$

The boundedness of r and Assumption 2.2-2 ensure the existence of a sequence $(\varepsilon_n)_{n\in\mathbb{N}^*}$ converging to 0 such that

$$\begin{aligned}\mathbb{E}(\langle \bar{X}^n_t, f_k\rangle) \leq \mathbb{E}(\langle \bar{X}^n_0, f_k\rangle) + \bar{b}\int_0^t \mathbb{E}(\langle \bar{X}^n_s, f_k\rangle)ds + \varepsilon_n\int_0^t \mathbb{E}(\langle \bar{X}^n_s, 1\rangle)ds\\ + \mathbb{E}\Big(\int_0^t\int_{\mathbb{R}^d\times\mathbb{R}_+} r(x,a)p(x,a)\,Af_k(x)X^n_s(dx,da)ds\Big),\end{aligned} \tag{3.16}$$

and we have by (2.4) and (3.12)

$$\Big|\int_{\mathbb{R}^d\times\mathbb{R}_+} r(x,a)p(x,a)\,Af_k(x)X^n_s(dx,da)\Big| \leq \bar{r}\sum_{|\ell|\leq\ell_0}\|D^\ell f_k\|_\infty\ \langle \bar{X}^n_s, f_{k-1}\rangle.$$

Since moreover, the sequence $(f_k)_{k\in\mathbb{N}^*}$ is non-increasing, $\langle \bar{X}^n_s, f_k\rangle \leq \langle \bar{X}^n_s, f_{k-1}\rangle$ and there is a constant $C>0$ independent of $k\geq 2$ such that

$$\mathbb{E}(\langle \bar{X}^n_t, f_k\rangle) \leq \mathbb{E}(\langle \bar{X}^n_0, f_k\rangle) + C\int_0^t \mathbb{E}(\langle \bar{X}^n_s, f_{k-1}\rangle)ds + \varepsilon_n\int_0^t \mathbb{E}(\langle \bar{X}^n_s, 1\rangle)ds. \tag{3.17}$$

Let $\mu_s^{n,k} = \mathbb{E}\left(\langle \bar{X}_s^n, f_k \rangle\right) \le \mu_s^n = \mathbb{E}\left(\langle \bar{X}_s^n, 1 \rangle\right)$ which is bounded uniformly in $n \in \mathbb{N}^*$ and $s \in [0,T]$ according to (2.9). There exist two positive constants C_1 and C_2 such that

$$\mu_t^{n,k} \le \mu_0^{n,k} + C_1 \int_0^t \mu_s^{n,k-1} ds + C_2 \varepsilon_n.$$

Iteration of this inequality yields

$$\begin{aligned} \mu_t^{n,k} &\le \sum_{\ell=0}^{k-1} \mu_0^{n,(k-\ell)} \frac{(C_1 t)^\ell}{\ell!} + \frac{(C_1 \int_0^t \mu_s^n ds)^k}{k!} + \varepsilon_n C_2 \sum_{\ell=0}^{k-1} \frac{(C_1 t)^\ell}{\ell!} \\ &\le \mu_0^{n,\lfloor k/2 \rfloor} e^{C_1 t} + \mu_0^n \sum_{\ell=\lfloor k/2 \rfloor + 1}^{+\infty} \frac{(C_1 t)^\ell}{\ell!} + \frac{(C_1' t)^k}{(k)!} + \varepsilon_n C_2 e^{C_1 t}. \end{aligned}$$

where we used the monotonicity of $\mu_0^{n,k}$ w.r.t. k for the second inequality. Given the moment condition (2.8), the assumption of tightness in $(\mathcal{M}_F(\mathcal{X}), w)$ of the initial conditions $(\bar{X}_0^n)_{n \in \mathbb{N}^*}$ is equivalent to

$$\lim_{k \to +\infty} \limsup_{n \to +\infty} \mu_0^{n,k} = 0. \tag{3.18}$$

Hence

$$\lim_{k \to +\infty} \limsup_{n \to +\infty} \mu_t^{n,k} \le \sup_{n \in \mathbb{N}^*} \mu_0^n \lim_{k \to +\infty} \sum_{\ell=\lfloor k/2 \rfloor + 1}^{+\infty} \frac{(C_1 t)^\ell}{\ell!} + \lim_{k \to +\infty} \frac{(C_1' t)^k}{(k)!}.$$

We deduce immediately that

$$\lim_{k \to +\infty} \limsup_{n \to +\infty} \mathbb{E}\left(\langle \bar{X}_t^n, f_k \rangle\right) = \lim_{k \to +\infty} \limsup_{n \to +\infty} \mu_t^{n,k} = 0. \tag{3.19}$$

Let us now consider the martingale $M_t^{n,k}$ defined by (3.5) with f_k instead of f, and with quadratic variation given in (3.6). Similar arguments as above allow us to prove that

$$\mathbb{E}(\langle M^{n,k} \rangle_t) \le C_1 \int_0^t \mathbb{E}(\langle X_s^n, f_{k-1} \rangle) ds + \varepsilon_n C_2 \int_0^t \mathbb{E}(\langle X_s^n, 1 \rangle) ds.$$

Thus, using that $f_k \le 1$, Doob's inequality, (3.15), (2.9) and the dominated convergence theorem, we get

$$\lim_{k \to +\infty} \limsup_{n \to +\infty} \mathbb{E}\Big(\sup_{t \le T} |M_t^{n,k}|\Big) = 0.$$

Let us now come back to the process $\langle \bar{X}^n, f_k \rangle$. As before, we can get

$$\begin{aligned} \langle \bar{X}_t^n, f_k \rangle &\le \langle \bar{X}_0^n, f_k \rangle + M_t^{n,k} + \bar{b} \int_0^t \langle \bar{X}_s^n, f_k \rangle ds + \varepsilon_n \int_0^t \langle \bar{X}_s^n, 1 \rangle ds \\ &\quad + \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}_+} r(x,a) p(x,a)\, A f_k(x) X_s^n(dx, da) ds \\ &\le \langle \bar{X}_0^n, f_k \rangle + M_t^{n,k} + C_1 \int_0^t \langle \bar{X}_s^n, f_{k-1} \rangle ds + \varepsilon_n C_2 \int_0^t \langle \bar{X}_s^n, 1 \rangle ds, \end{aligned} \tag{3.20}$$

for constants C_1 and C_2. Let $\alpha_t^{n,k} = \mathbb{E}\left(\sup_{s \le t} \langle \bar{X}_s^n, f_k \rangle\right)$ and $\alpha_t^n = \mathbb{E}\left(\sup_{s \le t} \langle \bar{X}_s^n, 1 \rangle\right)$ which is bounded uniformly in $n \in \mathbb{N}^*$ and $t \in [0,T]$ according to (2.9). One deduces that

$$\alpha_t^{n,k} \le \alpha_0^{n,k} + C_1 \int_0^t \mu_s^{n,k-1} ds + C_2 \varepsilon_n + \mathbb{E}\Big(\sup_{t \le T} |M_t^{n,k}|\Big).$$

An iteration as before allows us to prove that

$$\lim_{k \to +\infty} \limsup_{n \to +\infty} \mathbb{E}\Big(\sup_{t \le T} \langle \bar{X}_t^n, f_k \rangle\Big) = \lim_{k \to +\infty} \limsup_{n \to +\infty} \alpha_t^{n,k} = 0,$$

which concludes the proof of Lemma 3.3 and thus the one of Proposition 3.2. ∎

3.2 Identification of the limiting values

To obtain the convergence stated in Theorem 3.1, we show that the limiting value $\bar{X}$ of the uniformly tight sequence $(\bar{X}^n)_{n\in\mathbb{N}^*}$ is unique. We establish a martingale problem satisfied by $\bar{X}$ in which there are integration terms with respect to the equilibrium (3.2) involved in the averaging phenomenon for the ages. The uniqueness of the solution to the martingale problem is then proved.

3.2.1 Averaging phenomenon

We begin with establishing the form of the limiting values of the time-marginal distributions $(X_t^n(dx,da))_{n\in\mathbb{N}^*}$ for $t\in[0,T]$. Since the sequence $(\bar{X}^n)_{n\in\mathbb{N}^*}$ is uniformly tight, there exists a subsequence of $(X_t^n(dx,da))_{n\in\mathbb{N}^*}$, with trait-marginals converging to a limiting value $\bar{X}$, that by simplicity, we denote again by $(X_t^n(dx,da))_{n\in\mathbb{N}^*}$.

We have already explained why the uniform tightness of the sequence $(t\mapsto X_t^n(dx,da))_{n\in\mathbb{N}^*}$ in $\mathbb{D}([0,T],\mathcal{M}_F(\mathcal{X}\times\mathbb{R}_+))$ cannot hold. However, following Kurtz [20], we will prove the uniform tightness of the sequence of random measures

$$\Gamma^n(dt,dx,da)=X_t^n(dx,da)dt \tag{3.21}$$

on $\mathcal{M}_F([0,T]\times\mathcal{X}\times\mathbb{R}_+)$. Proceeding in this way allows us to escape the difficulties created by the degeneracies due to the rapid time scale for age, when one tries to follow individual paths.

Proposition 3.4. *We have:*
(i) For dt-almost every (a.e.) $t\in[0,T]$, the sequence $(X_t^n(dx,da))_{n\in\mathbb{N}^*}$ converges weakly to $\widehat{m}(x,da)\bar{X}_t(dx)$, with $\widehat{m}$ defined in (3.2).
(ii) The sequence $(\Gamma^n)_{n\in\mathbb{N}^*}$ converges in law to $\bar{X}_t(dx)\widehat{m}(x,da)dt$ in $\mathcal{M}_F([0,T]\times\mathcal{X}\times\mathbb{R}_+)$.

The proof of Proposition 3.4 is inspired by Kurtz [20]. Points (i) and (ii) are proved jointly. Point (i) enables us to identify the limiting finite-dimensional distributions of $(X_t^n(dx,da))_{n\in\mathbb{N}^*}$. To establish this result, we need to consider (3.21) which provides the point (ii).
Firstly, we prove the uniform tightness of the sequence $(X_t^n(dx,da))_{n\in\mathbb{N}^*}$, for fixed $t\in[0,T]$ (Lemma 3.5), as well as the one of the sequence of measures $(\Gamma^n)_{n\in\mathbb{N}^*}$ (Lemma 3.6), where the pathwise and individual points of view have been forgotten. The techniques to disentangle the traits and individuals' time scales appear strikingly in the proof of Lemma 3.5, where different treatments are used for the trait marginal and for the ages, with the introduction of the individuals' lifelengths. Then, in the proof of Proposition 3.4, a factor n appears in (3.37), when changing from the macroscopic scale to the microscopic scale. The next part of the proof consists in identifying the limiting martingale problem.

Lemma 3.5. *For dt-a.e. $t\in[0,T]$, the sequence $(X_t^n)_{n\in\mathbb{N}^*}$ is uniformly tight on $\mathcal{M}_F(\mathcal{X}\times\mathbb{R}_+)$.*

Proof. Let $\varepsilon>0$. Since the family $(\bar{X}_t^n)_{n\in\mathbb{N}^*}$ is tight, there exists a compact set $K\subset\mathbb{R}^d$ such that

$$\sup_{n\in\mathbb{N}^*}\mathbb{P}\big(\bar{X}_t^n(K^c)>\varepsilon\big)<\varepsilon. \tag{3.22}$$

Moreover, as r is bounded below by $\underline{r}$, we deduce that uniformly in $x\in\mathcal{X}$, the life-lengths $D_i(t)$ of the individuals in the population are stochastically dominated by independent exponential random variables of parameter $n\underline{r}$. Thus, for $A>0$ and $n_0, N\in\mathbb{N}^*$,

$$\begin{aligned}\sup_{n\geq n_0}\mathbb{P}\big(X^n_t((K\times[0,A])^c)>2\varepsilon\big)\leq&\sup_{n\geq n_0}\mathbb{P}\big(\bar{X}^n_t(K^c)>\varepsilon\big)+\sup_{n\geq n_0}\mathbb{P}\Big(\frac{1}{n}\sum_{i=1}^{N^n_t}\mathbb{1}_{\{nD_i(t)>A\}}>\varepsilon\Big)\\ \leq&\varepsilon+\sup_{n\geq n_0}\mathbb{P}\Big(\sum_{i=1}^{nN}\mathbb{1}_{\{E_i>A\}}>n\varepsilon\Big)+\sup_{n\geq n_0}\mathbb{P}\big(N^n_t>nN\big)\end{aligned}\tag{3.23}$$

where E_i are independent exponential variables of parameter $\underline{r}$. By (2.8), it is possible to find N such that:

$$\sup_{n\geq n_0}\mathbb{P}\big(N^n_t>nN\big)=\sup_{n\geq n_0}\mathbb{P}\big(\langle X^n_t,1\rangle>N\big)\leq\frac{\sup_{n\geq n_0}\mathbb{E}\big(\sup_{t\in[0,T]}\langle X^n_t,1\rangle\big)}{N}\leq\varepsilon.\tag{3.24}$$

For such an N, we choose A such that $\exp(-\underline{r}A)<\varepsilon/2N$. Then

$$\begin{aligned}\mathbb{P}\Big(\sum_{i=1}^{nN}\mathbb{1}_{\{E_i>A\}}>n\varepsilon\Big)=&\mathbb{P}\Big(\sum_{i=1}^{nN}(\mathbb{1}_{\{E_i>A\}}-e^{-\underline{r}A})>n(\varepsilon-Ne^{-\underline{r}A})\Big)\\ \leq&\mathbb{P}\Big(\sum_{i=1}^{nN}(\mathbb{1}_{\{E_i>A\}}-e^{-\underline{r}A})>n\varepsilon/2\Big)\leq\exp\Big(-\frac{n\varepsilon^2}{8(Ne^{-\underline{r}A}(1-e^{-\underline{r}A})+\varepsilon/6)}\Big)\end{aligned}\tag{3.25}$$

by Bernstein's inequality (*e.g.* [29] p.855). For a sufficiently large n_0, the r.h.s. of (3.25) is smaller than ε for $n\geq n_0$. The tightness of the sequence $(X^n_t)_{n\in\mathbb{N}^*}$ is thus a consequence of (3.23), (3.24) and (3.25). ■

Lemma 3.6. *The family $(\Gamma^n)_{n\in\mathbb{N}}$ is tight in $\mathcal{M}_F([0,T]\times\mathcal{X}\times\mathbb{R}_+)$.*

Proof. Following Kurtz [20, Lemma 1.3], a sufficient condition for the tightness of the family $(\Gamma^n)_{n\in\mathbb{N}}$ is that for all $\varepsilon>0$, there exists a compact set Ξ of $\mathcal{X}\times\mathbb{R}_+$ such that

$$\sup_{n\in\mathbb{N}^*}\mathbb{E}\Big(\Gamma^n\big([0,T]\times\Xi^c\big)\Big)\leq C(T)\varepsilon.\tag{3.26}$$

Let us establish (3.26). From the proof of Lemma 3.5, it appears that the upperbounds (3.22), (3.23) and (3.25) are uniform in $t\in[0,T]$ so that:

$$\sup_{t\in[0,T]}\sup_{n\in\mathbb{N}^*}\mathbb{P}\Big(X^n_t\big((K\times[0,A])^c\big)>2\varepsilon\Big)<3\varepsilon.\tag{3.27}$$

We are now ready to upperbound

$$\mathbb{E}\Big(\Gamma^n\big([0,T]\times(K\times[0,A])^c\big)\Big)=\mathbb{E}\Big(\int_0^T\langle X^n_t,\mathbb{1}_{(K\times[0,A])^c}\rangle dt\Big)=\int_0^T\mathbb{E}\Big(X^n_t\big((K\times[0,A])^c\big)\Big)dt.$$

Indeed:

$$\begin{aligned}\mathbb{E}\Big(X^n_t\big((K\times[0,A])^c\big)\Big)\leq&2\varepsilon\;\mathbb{P}\Big(X^n_t\big((K\times[0,A])^c\big)\leq2\varepsilon\Big)+\mathbb{E}\Big(\langle X^n_t,1\rangle\mathbb{1}_{X^n_t\big((K\times[0,A])^c\big)>2\varepsilon}\Big)\\ \leq&2\varepsilon+\sqrt{\mathbb{E}\Big(\langle X^n_t,1\rangle^2\Big)}\sqrt{\mathbb{P}\Big(X^n_t\big((K\times[0,A])^c\big)>2\varepsilon\Big)}\leq C(T)(\varepsilon+\sqrt{\varepsilon}),\end{aligned}\tag{3.28}$$

by Cauchy-Schwarz inequality and (2.9). This proves (3.26) and finishes the proof. ■

Before proving Proposition 3.4, we provide a lemma characterizing $\widehat{m}(x,a)$.

Lemma 3.7. *Let $x \in \mathcal{X}$ be fixed. There exists a unique probability measure $\widehat{m}(x,da)$ on $\mathbb{R}_+$, solution of the following equation: For $\psi \in \mathcal{C}_b^1(\mathbb{R}_+,\mathbb{R})$,*

$$\int_{\mathbb{R}_+} \partial_a\psi(a)\widehat{m}(x,da) = \int_{\mathbb{R}_+} \psi(a)r(x,a)\widehat{m}(x,da) - \psi(0)\int_{\mathbb{R}_+} r(x,a)\widehat{m}(x,da). \tag{3.29}$$

The probability measure $\widehat{m}(x,da)$ is absolutely continuous with respect to the Lebesgue measure and its density is given in (3.2).

Proof. Let us consider the test function $\psi(a) = f(0) + \int_0^a f(\alpha)d\alpha$, where $f \in \mathcal{C}(\mathbb{R}_+,\mathbb{R}_+)$ is positive. Then $\partial_a\psi(a) = f(a)$ and $\psi(0) = f(0)$. Equation (3.29) gives by Fubini's theorem:

$$\begin{aligned}\int_{\mathbb{R}_+} f(a)\widehat{m}(x,da) &= \int_{\mathbb{R}_+} \Big[\Big(f(0) + \int_0^{+\infty} \mathbb{1}_{\alpha<a} f(\alpha)d\alpha\Big) - f(0)\Big] r(x,a)\widehat{m}(x,da)\\ &= \int_{\mathbb{R}_+} f(\alpha)\int_\alpha^{+\infty} r(x,a)\widehat{m}(x,da)\,d\alpha. \end{aligned}\tag{3.30}$$

This entails that $\widehat{m}(x,da)$ is absolutely continuous with respect to the Lebesgue measure with density $\widehat{m}(x,a) = \int_a^{+\infty} r(x,\alpha)\widehat{m}(x,\alpha)d\alpha$. The latter implies that $a \mapsto \widehat{m}(x,a)$ is a function of class $\mathcal{C}^\infty$. Using further an integration by part in (3.29), we get for all $\psi \in \mathcal{C}_b^1(\mathbb{R}_+,\mathbb{R})$

$$-\psi(0)\widehat{m}(x,0) - \int_{\mathbb{R}_+} \psi(a)\partial_a\widehat{m}(x,a)da = \int_{\mathbb{R}_+} \big(\psi(a) - \psi(0)\big)r(x,a)\widehat{m}(x,a)da. \tag{3.31}$$

By identification, we obtain that $\widehat{m}(x,a)$ is a solution of

$$\begin{aligned}\partial_a\widehat{m}(x,a) &= -r(x,a)\widehat{m}(x,a)\\ \widehat{m}(x,0) &= \int_{\mathbb{R}_+} r(x,a)\widehat{m}(x,a)da,\end{aligned}\tag{3.32}$$

which is solved by

$$\widehat{m}(x,a) = \widehat{m}(x,0)\exp\big(-\int_0^a r(x,\alpha)d\alpha\big). \tag{3.33}$$

Since $\widehat{m}(x,a)da$ is a probability measure, necessarily

$$\widehat{m}(x,0) = \int_{\mathbb{R}_+} r(x,a)\widehat{m}(x,a)da = \frac{1}{\int_{\mathbb{R}_+}\exp\big(-\int_0^a r(x,\alpha)d\alpha\big)da}. \tag{3.34}$$

This provides existence and uniqueness of the solution of (3.32) and hence of (3.29). ■

Remark 3.8. Notice that the system (3.32) defines the stable age equilibrium of the McKendrick-Von Foerster equation [21, 15] (see also [33]) when the birth and death rates equal to $r(x,a)$ and the trait x is fixed. □

We are now able to prove Proposition 3.4.

Proof of Proposition 3.4. From (3.21), we can see that the marginal measure of $\Gamma^n(ds,dx,da)$ on $[0,T]\times\mathcal{X}$ is $\bar{X}_s^n(dx)ds$. For any real bounded test function $\varphi : (s,x) \mapsto \varphi_s(x)$ on $[0,T]\times\mathcal{X}$,

$$\int_0^t\int_{\mathcal{X}\times\mathbb{R}_+} \varphi_s(x)\Gamma^n(ds,dx,da) = \int_0^t \langle \bar{X}_s^n, \varphi_s\rangle ds. \tag{3.35}$$

The sequence $(\bar{X}^n)_{n\in\mathbb{N}^*}$ is uniformly tight by Proposition 3.2, as well as $(\Gamma^n)_{n\in\mathbb{N}^*}$, by Lemma 3.6 (ii). Using Prohorov's theorem, we thus deduce that $(\Gamma^n(ds,dx,da), \bar{X}^n_s(dx)ds)_n$ is relatively compact and there exists a subsequence that converges in distribution to a limiting value, say $(\Gamma(ds,dx,da), \bar{X}_s(dx)ds)$. Taking (3.35) to the limit, we obtain that $\bar{X}_s(dx)ds$ is necessarily the marginal measure of $\Gamma(ds,dx,da)$ on $[0,T]\times\mathcal{X}$ up to a null-measure set. We deduce from this (*e.g.* Lemma 1.4 of Kurtz [20]) that there exists a (random) probability-valued process $(\gamma_{s,x}(da),\ s\in[0,T], x\in\mathcal{X})$ that is predictable in (ω,s) and such that for all bounded measurable function $\varphi(s,x,a)$ on $[0,T]\times\mathcal{X}\times\mathbb{R}_+$,

$$\int_0^t\int_{\mathcal{X}\times\mathbb{R}_+}\varphi(s,x,a)\Gamma(ds,dx,da)=\int_0^t\int_{\mathcal{X}}\int_{\mathbb{R}_+}\varphi(s,x,a)\gamma_{s,x}(da)\bar{X}_s(dx)ds. \tag{3.36}$$

We now want to characterize the limiting value $\Gamma(ds,dx,da)=\gamma_{s,x}(da)\bar{X}_s(dx)ds$. Applying (2.10) for a test function $\varphi(x,a)\in\mathcal{C}^{0,1}_b(\mathcal{X}\times\mathbb{R}_+,\mathbb{R})$ and dividing by n gives that:

$$\begin{aligned}\frac{M^{n,\varphi}_t}{n}=\frac{\langle X^n_t,\varphi\rangle-\langle X^n_0,\varphi\rangle}{n}-\int_0^t\int_{\mathcal{X}\times\mathbb{R}_+}\Big[&\partial_a\varphi(x,a)\\ &+\Big(r(x,a)+\frac{b(x,a)}{n}\Big)\int_{\mathbb{R}^d}\varphi(x+h,0)K^n(x,a,dh)\\ &-\Big(r(x,a)+\frac{d(x,a)+X^n_sU(x,a)}{n}\Big)\varphi(x,a)\Big]\Gamma^n(ds,dx,da)\end{aligned} \tag{3.37}$$

is a martingale. For each t, the process

$$\widetilde{M}^{n,\varphi}_t := \int_0^t\int_{\mathcal{X}\times\mathbb{R}_+}\Big[\partial_a\varphi(x,a)+r(x,a)\Big(\varphi(x,0)-\varphi(x,a)\Big)\Big]\Gamma^n(ds,dx,da) \tag{3.38}$$

converges in distribution to

$$\widetilde{M}^{\varphi}_t=\int_0^t\int_{\mathcal{X}\times\mathbb{R}_+}\Big[\partial_a\varphi(x,a)+r(x,a)\big(\varphi(x,0)-\varphi(x,a)\big)\Big]\gamma_{s,x}(da)\bar{X}_s(dx)ds, \tag{3.39}$$

since the integrand in (3.38) is continuous and bounded. Thanks to the moments estimates of Proposition 2.5,

$$\lim_{n\to+\infty}\mathbb{E}\Big(\Big|\frac{M^{n,\varphi}_t}{n}-\widetilde{M}^{n,\varphi}_t\Big|\Big)=0,$$

and hence, we deduce that the process $(\widetilde{M}^{\varphi}_t)_t$ defined in (3.39) is a martingale. As it is also a continuous and finite variation process, it must hence be almost surely zero. Since this holds for every $t\in\mathbb{R}_+$, we have proved that a.s., dt-a.e.

$$\int_{\mathcal{X}\times\mathbb{R}_+}\Big[\partial_a\varphi(x,a)+r(x,a)\big(\varphi(x,0)-\varphi(x,a)\big)\Big]\gamma_{t,x}(da)\bar{X}_t(dx)=0. \tag{3.40}$$

Choosing $\varphi(x,a)=\phi(x)\psi(a)$ with ϕ, respectively ψ, bounded, resp. C^1_b, real functions, (3.40) provides that

$$\int_{\mathcal{X}}\phi(x)H(t,x)\bar{X}_t(dx)=0$$

where $H(t,x)=\int_{\mathbb{R}_+}\Big[\partial_a\psi(a)+r(x,a)\big(\psi(0)-\psi(a)\big)\Big]\gamma_{t,x}(da)$. Almost surely, the function $H(t,x)$ is bounded and is thus dt-a.e. $\bar{X}_t(dx)$-integrable. We obtain that a.s., dt-a.e., $\bar{X}_t(dx)$-a.e.,

$$\int_{\mathbb{R}_+}\Big[\partial_a\psi(a)+r(x,a)\big(\psi(0)-\psi(a)\big)\Big]\gamma_{t,x}(da)=0. \tag{3.41}$$

By Lemma 3.7, we deduce that a.s., dt-a.e., and $\bar{X}_t(dx)$-a.e., $\gamma_{t,x}(da)=\widehat{m}(x,a)da$ and as a consequence, any limiting value of $(X^n_t)_{n\in\mathbb{N}^*}$ is of the form $\bar{X}_t(dx)\otimes\widehat{m}(x,a)da$. ∎

3.2.2 Characterization of the limiting values

In the previous sections, we have proved that the sequence $(\bar{X}^n)_{n\in\mathbb{N}^*}$ is tight and that for a given limiting value $\bar{X}$, the associated subsequence $(\Gamma^n(dt,dx,da) = X^n_t(dx,da)\ dt)_{n\in\mathbb{N}^*}$ converges in $(\mathcal{M}_F([0,T]\times\mathcal{X}\times\mathbb{R}_+),w)$ to $\bar{X}_t(dx)\widehat{m}(x,da)\ dt$. Now, we are ready to prove that:

Lemma 3.9. *The limiting values $\bar{X}$ of the sequence $(\bar{X}^n)_{n\in\mathbb{N}^*}$ are solution of the martingale problem (3.3)-(3.4).*

Proof. Let $0 < s_1 \leq \dots s_k < s < t$, and let us introduce for $Y \in \mathbb{D}(\mathbb{R}_+, \mathcal{M}_F(\mathcal{X}))$:

$$\Psi_{s,t}(Y) = \phi_1(Y_{s_1})\dots\phi_k(Y_{s_k})\Big\{\langle Y_t, f\rangle - \langle Y_s, f\rangle - \int_s^t du \int_{\mathcal{X}} Y_u(dx)\Big[\widehat{(p\,r)}(x)Af(x) + \big(\widehat{b}(x) - \widehat{d}(x) - Y_s\widehat{U}(x)\big)f(x)\Big]\Big\}, \quad (3.42)$$

where $\phi_1,\dots,\phi_k$ are bounded continuous functions on $\mathcal{M}_F(\mathcal{X})$ and $f\in\mathcal{D}(A)$. Our purpose is to prove that $\mathbb{E}\big(\Psi_{s,t}(\bar{X})\big)=0$ for any limiting value $\bar{X}$ of $(\bar{X}^n)_{n\in\mathbb{N}^*}$.
Let $\bar{X}$ be a limiting value of $(\bar{X}^n)_{n\in\mathbb{N}^*}$ and let $(\bar{X}^{u_n})_{n\in\mathbb{N}^*}$ be a subsequence converging to $\bar{X}$. On the one hand, thanks to Proposition 3.4 (ii) and (2.9):

$$\mathbb{E}\big(\Psi_{s,t}(\bar{X})\big) = \lim_{n\to+\infty}\mathbb{E}\Big(\phi_1(\bar{X}^{u_n}_{s_1})\dots\phi_k(\bar{X}^{u_n}_{s_k})\Big\{\langle \bar{X}^{u_n}_t, f\rangle - \langle \bar{X}^{u_n}_s, f\rangle - \int_s^t du\int_{\mathcal{X}\times\mathbb{R}_+} X^{u_n}_u(dx,da)\Big[p(x,a)\,r(x,a)\,Af(x) + \Big(b(x,a) - d(x,a) - \int_{\mathcal{X}\times\mathbb{R}_+} U((x,a),(y,\alpha))X^{u_n}_u(dy,d\alpha)\Big)f(x)\Big]\Big\}\Big). \quad (3.43)$$

On the other hand, the term under the expectation in the r.h.s. of (3.43) equals:

$$\phi_1(X^{u_n}_{s_1})\dots\phi_k(X^{u_n}_{s_k})\Big\{M^{u_n,f}_t - M^{u_n,f}_s + A_{u_n} + B_{u_n}\Big\}, \quad (3.44)$$

where $M^{n,f}$ has been defined in (3.5) and where:

$$\begin{aligned} A_n &= \int_s^t du\int_{\mathcal{X}\times\mathbb{R}_+} X^n_u(dx,da)\ r(x,a)\Big[n\int_{\mathcal{X}}\Big(f(x+h)-f(x)\Big)K^n(x,a,dh) - p(x,a)Af(x)\Big] \\ &= \int_s^t du\int_{\mathcal{X}\times\mathbb{R}_+} X^n_u(dx,da)\ r(x,a)\,p(x,a)\ \Big[n\int_{\mathcal{X}}\Big(f(x+h)-f(x)\Big)\pi^n(x,dh) - Af(x)\Big] \\ B_n &= \int_s^t du\int_{\mathcal{X}\times\mathbb{R}_+} X^n_u(dx,da)\ b(x,a)\Big[\int_{\mathbb{R}^d} f(x+h)K^n(x,a,dh) - f(x)\Big] \end{aligned}$$

Firstly, using (2.9) and the fact that the process $M^{n,f}$ is a martingale we obtain that:

$$\mathbb{E}\big(\phi_1(X^{u_n}_{s_1})\dots\phi_k(X^{u_n}_{s_k})\big[M^{u_n,f}_t - M^{u_n,f}_s\big]\big) = 0. \quad (3.45)$$

Secondly, from Assumption 2.2, $|\int_{\mathbb{R}^d} f(x+h)K^n(x,a,dh) - f(x)| = o(1/n)$ and using (2.9) again provides:

$$\lim_{n\to+\infty}\mathbb{E}\big(\phi_1(X^{u_n}_{s_1})\dots\phi_k(X^{u_n}_{s_k})\big[A_{u_n}+B_{u_n}\big]\big) = 0. \quad (3.46)$$

From (3.43), (3.44), (3.45) and (3.46), we deduce that $\mathbb{E}\big(\Psi_{s,t}(\bar{X})\big)=0$ which ends the proof. ∎

3.2.3 Uniqueness of the martingale problem

We have shown that the limiting values of the uniformly tight sequence $(\bar{X}^n)_{n\in\mathbb{N}^*}$ satisfy the martingale problem (3.3)-(3.4). To conclude the proof of Theorem 3.1, it remains to prove the uniqueness of the solution of this martingale problem.

Proposition 3.10. *There is a unique solution to the martingale problem defined in Theorem 3.1.*

Proof. We start with getting rid of the non-linearity by using Girsanov's formula (see Dawson [8] Theorem 7.2.2). There exists a probability measure $\mathbb{Q}$ on the path space such that for all $f \in \mathcal{D}(A)$:

$$\begin{aligned}\widetilde{M}_t^f =& M_t^f + \int_0^t \int_{\mathbb{R}^d} \Big(\widehat{b}(x) - \widehat{d}(x) - \bar{X}_s U(x)\Big) f(x) \bar{X}_s(dx)\, ds \\ =& \langle \bar{X}_t, f\rangle - \langle \bar{X}_0, f\rangle - \int_0^t \int_{\mathbb{R}^d} \widehat{(p\, r)}(x) Af(x) \bar{X}_s(dx)\, ds \end{aligned} \tag{3.47}$$

is a square integrable martingale with bracket (3.4).

The uniqueness of the martingale problem (3.47)-(3.4) is proved by Roelly and Rouault [28]. It is based on the branching property of $\bar{X}$ under $\mathbb{Q}$ which allows us to characterize the Laplace functional of X by its cumulant:

$$L_t(f) = \mathbb{E}\Big(e^{\langle \bar{X}, f\rangle}\Big) = \mathbb{E}\Big(e^{\langle \bar{X}_0, U_t f\rangle}\Big). \tag{3.48}$$

The latter is the unique positive solution of the following PDE:

$$\frac{\partial u}{\partial t}(t,x) = Au(t,x) - \widehat{r}(x) u^2(t,x), \tag{3.49}$$

(see *e.g.* Pazy [25, Th. 1.4 and 1.5 p.185 and 187]). ■

The proof of Theorem 3.1 is now complete.

4 Examples

Let us develop and compare two examples, which only differ by the function $r(x,a)$.

4.1 Logistic physical-age and size-structured population

In Méléard and Tran [23], the following example for a population structured by age $a \in \mathbb{R}_+$ and size $x \in \mathcal{X} = [0, x_0]$ is considered:

$$\begin{aligned} &b(x,a) = x(x_0 - x) e^{-a} \mathbb{1}_{[0,x_0]}(x) \text{ for } x_0 > 0, \\ &d(x,a) = d_0, \quad U((x,a),(y,\alpha)) = \eta(x_0 - x), \end{aligned} \tag{4.1}$$

with $x_0 = 4$, $d_0 = 1/4$ and $\eta = 1.7$. Because reproduction needs energy, and since this energy depends on the size of the created offspring, very small or big individuals are disadvantaged. Individuals of intermediate size $x = 2$ have the highest birth rate. The competition term in contrast favors bigger individuals. Hence there is a trade-off between competitiveness and reproduction. The decreasing exponential in age describes a senescence phenomenon: older individuals reproduce less than their young competitors. In [23], partial differential equation limits, Trait substitution sequence and Canonical equations are considered. Here we consider the superprocess approximation described in the above sections, with $r(x,a)=1$ and $\pi^n(x,dh)$ a centered Gaussian kernel with variance $\frac{\sigma^2}{n}$ conditioned on $[0,x_0]$, as in Example 2.4.

Computation gives $\widehat{m}(x,a) = e^{-a}$ so that $X_t(dx,da) = \bar{X}_t(dx) \otimes e^{-a}\,da$ becomes in this particular case a product measure. As soon as the population survives, the age distribution "stabilizes" around an exponential distribution with parameter 1, as seen on the simulations of Figure 1. With the age distribution $\widehat{m}(x,a)=e^{-a}$, we get

$$\widehat{b}(x) = x(x_0-x)\int_{\mathbb{R}_+} e^{-2a}da = \frac{x(x_0-x)}{2} \quad ; \quad \widehat{d}(x) = d_0 \quad \widehat{U}(x,y) = \eta(x_0-x). \tag{4.2}$$

The martingale problem (3.3) becomes here:

$$\begin{aligned} M_t^f = \langle \bar{X}_t, f\rangle - \langle \bar{X}_0, f\rangle - \int_0^t \int_{\mathcal{X}} \Big(& p\frac{\sigma^2}{2}\Delta f(x) \\ & + \big[\frac{x(x_0-x)}{2} - \big(d_0 + \eta(x_0-x)\langle \bar{X}_s, 1\rangle\big)\big] f(x)\Big) \bar{X}_s(dx)\,ds, \end{aligned} \tag{4.3}$$

$$\langle M^f\rangle_t = \int_0^t \int_{\mathcal{X}} 2f^2(x)\bar{X}_s(dx)\,ds.$$

(a) (b) (c)

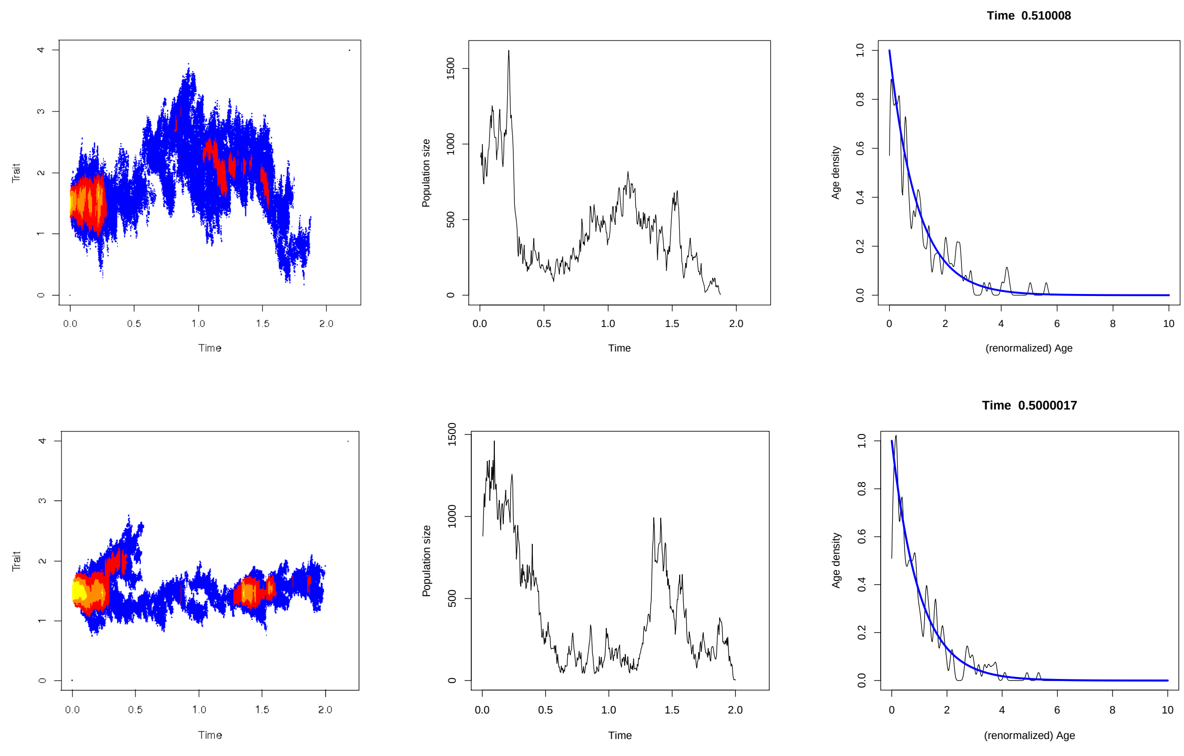

Figure 1: *Simulation of the individual-based process X^n, with $n=1000$ and discretization step $\Delta t = 0.005$. The system is started with 1000 particles of trait $x=1.5$. First line: $\sigma=1$. Second line: $\sigma=0.8$. (a): Support of the process $\bar{X}^n$ (with time in abscissa and trait in ordinate). (b): Evolution of the population size. (c): Age distribution for $t=0.5$. It can be checked that the age distribution converges to an exponential of parameter 1 (plain line).*

In Figure 1, two sets of simulations are presented, depending on two different mutation variances σ^2. As expected, when σ increases, the traits vary more rapidly, and the irregularity of the trait support appears more strikingly. On both simulations of Fig. 1, extinction happens in a fast time. Almost-sure extinction is due to the logistic interaction, as proved in the following proposition.

Proposition 4.1. *There is almost-sure extinction of the super-process (4.3).*

Proof. The mass of the super-process satisfies the following equation:

$$\begin{aligned}\langle \bar{X}_t, 1\rangle &= \langle \bar{X}_0, 1\rangle + \int_0^t \int_{[0,x_0]} A(x, \langle X_s, 1\rangle)\bar{X}_s(dx)\, ds + M^1_t\\ \langle M^1\rangle_t &= \int_0^t 2\langle \bar{X}_s, 1\rangle\, ds,\\ \text{where } A(x, Z) &= \frac{x(x_0 - x)}{2} - \big(d_0 + \eta(x_0 - x)Z\big).\end{aligned} \tag{4.4}$$

This equation is not closed for the mass process, since the drift depends on the trait distribution. Our purpose is to upper-bound $A(x, Z)$ so that $\langle \bar{X}_., 1\rangle$ can be stochastically dominated by a Feller diffusion with negative drift, that goes extinct almost surely.
In the case where $x > x_0 - (\frac{2d_0}{x_0} - \zeta)$ with $\zeta \in (0, \frac{2d_0}{x_0} \wedge 1)$ and since $x \in (0, x_0)$, one gets

$$\begin{aligned}A(x, Z) =& - d_0 + (x_0 - x)\Big(\frac{x}{2} - \eta Z\Big)\\ \leq& - d_0 + \Big(\frac{2d_0}{x_0} - \zeta\Big) \times \frac{x_0}{2} = -\frac{\zeta x_0}{2}.\end{aligned} \tag{4.5}$$

In the case where $x < x_0 - (\frac{2d_0}{x_0} - \zeta)$, then $0 < \frac{2d_0}{x_0} - \zeta \leq x_0 - x \leq x_0$ and depending on the sign of $x/2 - \eta Z$:

$$\begin{aligned}A(x, Z) \leq& - d_0 + \max\Big(x_0\big(\frac{x_0}{2} - \eta Z\big)\ ;\ \big(\frac{2d_0}{x_0} - \zeta\big)\big(\frac{x_0}{2} - \eta Z\big)\Big)\\ \leq& \frac{x_0^2}{2} - d_0 - \eta\big(\frac{2d_0}{x_0} - \zeta\big)Z.\end{aligned} \tag{4.6}$$

Since the upper bounds in (4.5) and (4.6) are equal when the mass Z equals m_0 defined by

$$m_0 = \frac{\frac{x_0(x_0+\zeta)}{2} - d_0}{\eta\big(\frac{2d_0}{x_0} - \zeta\big)}, \tag{4.7}$$

we thus get in any case that

$$A(x, Z) \leq -\frac{\zeta x_0}{2}\mathbb{1}_{Z\geq m_0} + \Big(\frac{x_0^2}{2} - d_0\Big)\mathbb{1}_{Z\leq m_0}. \tag{4.8}$$

Hence, the process $\langle \bar{X}_., 1\rangle$ can be stochastically dominated by the following positive process:

$$Z_t = \langle \bar{X}_0, 1\rangle + \int_0^t \Big(- \frac{\zeta x_0}{2} Z_s + m_0\big(\frac{x_0(x_0 + \zeta)}{2} - d_0\big)\mathbb{1}_{Z_s\leq m_0}\Big)ds + \int_0^t \sqrt{2Z_s}dB_s \tag{4.9}$$

where B is a standard Brownian motion.
We can adapt the results of Meyn and Tweedie [24] to prove almost sure extinction. For $u \in \mathbb{R}_+$, let us denote by $\tau_u = \inf\{t \geq 0,\ Z_t \leq u\}$ and let $z = \langle \bar{X}_0, 1\rangle > 0$. Either $z \leq m_0$

and then $\tau_{m_0} = 0$, or $z > m_0$. In the latter case, let us consider $M > z > m_0$ and let $\rho_M = \inf\{t \geq 0,\ Z_t \geq M\}$. We have

$$\mathbb{E}_z\Big(Z_{\tau_{m_0}\wedge\rho_M} - z + \frac{\zeta x_0}{2}\int_0^{\tau_{m_0}\wedge\rho_M} Z_s\,ds + \int_0^{\tau_{m_0}\wedge\rho_M}\sqrt{2Z_s}dB_s\Big) = 0. \tag{4.10}$$

By uniform integrability of the fourth term and optional stopping theorem, we deduce that

$$m_0\ \mathbb{E}_z\big(\tau_{m_0}\wedge\rho_M\big) \leq \mathbb{E}_z\Big(\int_0^{\tau_{m_0}\wedge\rho_M} Z_s\ ds\Big) \leq \frac{2z}{\zeta x_0}. \tag{4.11}$$

It can easily be proved that for all $T > 0$, $\mathbb{E}(\sup_{t\leq T}(Z_t)^2) < \infty$, implying that ρ_M tends to infinity with M. Thus, (4.11) provides that for all $z > 0$, $\mathbb{P}_z(\tau_{m_0} < +\infty) = 1$. By Girsanov's theorem, there exists a probability measure under which the process Z is a sub-critical Feller diffusion. It turns out that $\mathbb{P}_{m_0}(\tau_0 \wedge \rho_M < +\infty) = 1$. Standard computation using the strong Markov property yields $\mathbb{P}_z(\tau_0 < +\infty) = 1$. ■

4.2 Logistic biological-age and size-structured population

In this section, the trait x is linked to the rate of metabolism, which measures the energy expended by individuals, and is often an increasing function of the body size. Ageing may result from toxic by-products of the metabolism. In this example the ageing velocity equals x and $r(x,a) = xa$ represents the biological age, the other parameters are chosen as in Subsection 4.1. For a review on body size, energy metabolism and ageing, we refer the reader to [30]. In our example:

$$\widehat{m}(x,a) = \frac{2\sqrt{x}e^{-\frac{xa^2}{2}}}{\sqrt{2\pi}}\mathbb{1}_{[0,+\infty)}(a). \tag{4.12}$$

We recognize the Gaussian distribution conditioned on being positive. Then:

$$\begin{aligned}\widehat{b}(x) =&\frac{2x^{3/2}(x_0-x)}{\sqrt{2\pi}}\int_0^{+\infty} e^{-\frac{x(a^2+2a/x)}{2}}da = \frac{2x(x_0-x)e^{\frac{1}{2x}}}{\sqrt{2\pi}}\int_{1/\sqrt{x}}^{+\infty} e^{-\frac{\alpha^2}{2}}d\alpha\\ =&2x(x_0-x)e^{\frac{1}{2x}}\Phi\Big(-\frac{1}{\sqrt{x}}\Big), \qquad (4.13)\\ \widehat{r}(x) =&\frac{2x^{3/2}}{\sqrt{2\pi}}\int_0^{+\infty} ae^{-\frac{xa^2}{2}}da = \sqrt{\frac{2x}{\pi}}\int_0^{+\infty}\alpha e^{-\frac{\alpha^2}{2}}d\alpha = \sqrt{\frac{2x}{\pi}}, \qquad (4.14)\end{aligned}$$

where Φ is the cumulative distribution function of the standard Gaussian distribution. The functions $\widehat{d}(x)$ and $\widehat{U}(x,y)$ are unchanged and given by (4.2). The martingale problem (3.3) becomes here:

$$\begin{aligned}M_t^f &= \langle\bar{X}_t, f\rangle - \langle\bar{X}_0, f\rangle - \int_0^t\int_{\mathcal{X}}\Big(p\sqrt{\frac{x}{2\pi}}\sigma^2\Delta\varphi(x) + \big[2x(x_0-x)e^{\frac{1}{2x}}\Phi\Big(-\frac{1}{\sqrt{x}}\Big)\\ &\qquad - \big(d_0 + \eta(x_0-x)\langle\bar{X}_s,1\rangle\big)\big]f(x)\Big)\bar{X}_s(dx)\,ds \qquad (4.15)\\ \langle M^f\rangle_t &= \int_0^t\int_{\mathcal{X}} 2\sqrt{\frac{2x}{\pi}}f^2(x)\bar{X}_s(dx)\,ds.\end{aligned}$$

In this example, there is a higher senescence for individuals with trait $x > 1$, compared with Example 1. The new choice of $r(x,a)$ influences the age distribution: lifelengths are shortened. This can be seen on the smaller support of the age distribution (compare Fig. 2-(c) with Fig.

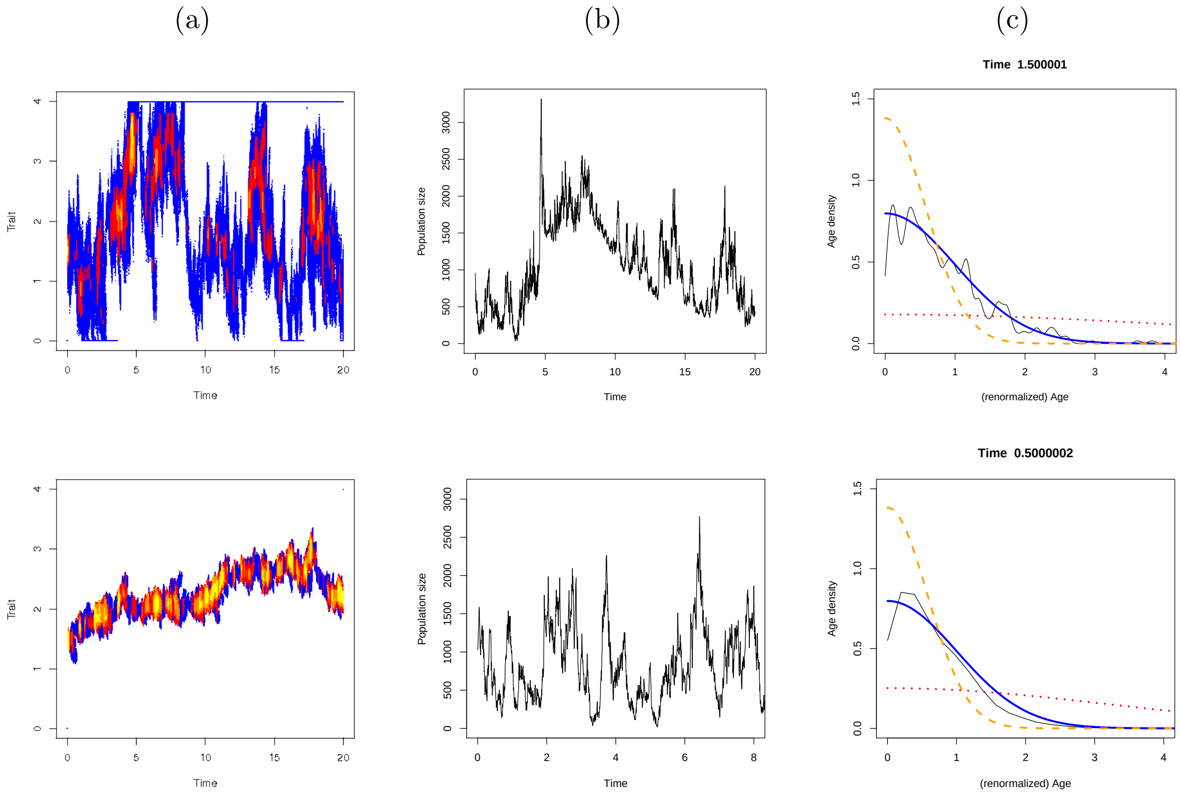

Figure 2: *Simulation of the individual-based process X^n, with $n = 1000$ and discretization step $\Delta t = 0.005$. The system is started with 1000 particles of trait $x = 1.5$. First line: $\sigma = 1$, Second line: $\sigma = 0.2$. (a): Support of the process $\bar{X}^n$ (with time in abscissa and trait in ordinate). (b): Evolution of the population size. (c): Marginal age distribution for $t = 0.5$. For comparison, we draw the density $\widehat{m}(1, a)$ (plain line), $\widehat{m}(0.5, a)$ (dotted line) and $\widehat{m}(3, a)$ (dashed line).*

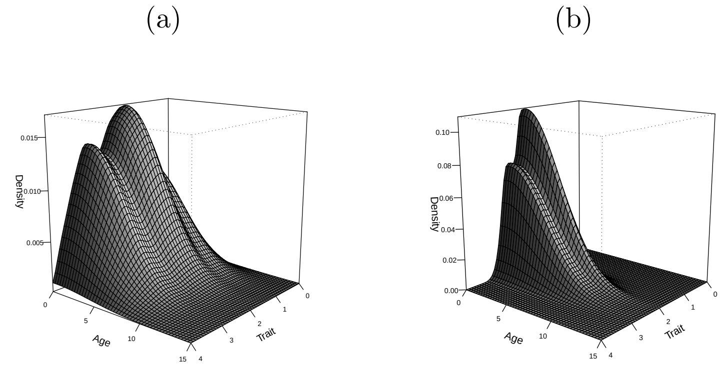

Figure 3: *Joint distributions in trait and age for the simulation of Fig. 2. We see that contrarily to Example 1, we do not have here independence of traits and age. (a): $\sigma = 1$. (b): $\sigma = 0.2$*

1-(c)).

However, the populations are more persistent in Example 2, although it can be proved similarly to Prop. 4.1 that there is almost sure extinction. Indeed, contrary to the populations in Ex-

ample 1 which are extinct at $t=2$, the population of Example 2 still survives at $t=20$. One reason is that the growth rate in the finite variation term of (4.15) is bigger than the one in Example 1. Indeed, for many values of x, the factor $2\exp(1/2x)\Phi(-1/\sqrt{x})$ in the birth rate $\widehat{b}(x)$ is bigger than the factor 1/2. For $x=1.5$, $2\exp(1/2x)\Phi(-1/\sqrt{x})=0.58>0.5$ and for $x=3$, $2\exp(1/2x)\Phi(-1/\sqrt{x})=0.67>0.5$.
When comparing Fig. 1-(b) and Fig. 2-(b), we observe more fluctuations of the population size in Example 2. The bracket of the martingale in (4.15) presents a multiplicative x term, compared to (4.3). As soon as $x>\frac{\pi}{2}$, this explains the increased variance. Notice however that this variance tends to zero when the population size tends to zero, which also explains why there is no decrease in the population persistence.
Finally, the multiplicative term $p\sqrt{x/2\pi}\sigma^2$ in front of the diffusion term $\Delta\varphi(x)$ explains the large variability of the trait support, which is observed in Fig. 2-(a). When the diffusion coefficient σ is small (second line of Fig. 2), the traits evolve towards a value between $x=2$ and $x=4$ where the trade-off between reproduction and competition is optimized.

Acknowledgements: This work benefited from the support of the ANR MANEGE (ANR-09-BLAN-0215) and from the "Chair Modélisation Mathématique et Biodiversité of Veolia Environnement-Ecole Polytechnique-Museum National d'Histoire Naturelle-Fondation X". The authors thank S. Billiard, T. Kurtz and L. Popovic for enriching discussions.

References

[1] K. Athreya, S. Athreya, and S. Iyer. Super-critical age dependent branching Markov processes and their scaling limits. Bernoulli, 2010. to appear.

[2] K. Ball, T.G. Kurtz, L. Popovic, and G. Rempala. Asymptotic analysis of multiscale approximations to reaction networks. Annals of Applied Probability, 16(4):1925–1961, 2006.

[3] A. Bose and I. Kaj. Diffusion approximation for an age-structured population. Annals of Applied Probabilities, 5(1):140–157, 1995.

[4] A. Bose and I. Kaj. A scaling limit process for the age-reproduction structure in a Markov population. Markov Processes and Related Fields, 6(3):397–428, 2000.

[5] N. Champagnat, R. Ferrière, and S. Méléard. Individual-based probabilistic models of adpatative evolution and various scaling approximations. In Proceedings of the 5th seminar on Stochastic Analysis, Random Fields and Applications, Probability in Progress Series, Ascona, Suisse, 2006. Birkhauser.

[6] N. Champagnat, R. Ferrière, and S. Méléard. Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models via timescale separation. Theoretical Population Biology, 69:297–321, 2006.

[7] B. Charlesworth. Evolution in Age structured Population. Cambridge University Press, 2 edition, 1994.

[8] D. A. Dawson. Mesure-valued markov processes. In Springer, editor, Ecole d'Eté de probabilités de Saint-Flour XXI, volume 1541 of Lectures Notes in Math., pages 1–260, New York, 1993.

[9] D.A. Dawson, L.G. Gorostiza, and Z. Li. Nonlocal branching superprocesses and some related models. Acta Applicandae Mathematicae, 74:93–112, 2002.

[10] E.B. Dynkin. Branching particle systems and superprocesses. Annals of Probability, 19:1157–1194, 1991.

[11] A. Etheridge. An introduction to superprocesses, volume 20 of University Lecture Series. Providence, American Mathematical Society edition, 2000.

[12] S.N. Evans and D. Steinsaltz. Damage segregation at fissioning may increase growth rates: A superprocess model. Theoretical Population Biology, 71:473–490, 2007.

[13] R. Ferrière and V.C. Tran. Stochastic and deterministic models for age-structured populations with genetically variable traits. ESAIM: Proceedings, 27:289–310, 2009.

[14] K. Fleischmann, V.A. Vatutin, and A. Wakolbinger. Branching systems with long-living particles at the critical dimension. Theoretical Probability and its Applications, 47(3):429–454, 2002.

[15] H. Von Foerster. Some remarks on changing populations. In Grune & Stratton, editor, The Kinetics of Cellular Proliferation, pages 382–407, New York 1959.

[16] N. Fournier and S. Méléard. A microscopic probabilistic description of a locally regulated population and macroscopic approximations. Ann. Appl. Probab., 14(4):1880–1919, 2004.

[17] A. Joffe and M. Métivier. Weak convergence of sequences of semimartingales with applications to multitype branching processes. Advances in Applied Probability, 18:20–65, 1986.

[18] B. Jourdain and S. Méléard. Lévy flights in ecology. 2010. work in progress.

[19] I. Kaj and S. Sagitov. Limit processes for age-dependent branching particle systems. Journal of Theoretical Probability, 11(1):225–257, 1998.

[20] T.G. Kurtz. Averaging for martingale problems and stochastic approximation. In Springer, editor, Applied stochastic analysis (New Brunswick, NJ, 1991), volume 177 of Lectures Notes in Control and Inform. Sci., pages 186–209, Berlin, 1992.

[21] A.G. McKendrick. Applications of mathematics to medical problems. Proc. Edin. Math.Soc., 54:98–130, 1926.

[22] S. Méléard and S. Roelly. Sur les convergences étroite ou vague de processus à valeurs mesures. C.R.Acad.Sci.Paris, Serie I, 317:785–788, 1993.

[23] S. Méléard and V.C. Tran. Trait substitution sequence process and canonical equation for age-structured populations. Journal of Mathematical Biology, 58(6):881–921, 2009.

[24] S.P. Meyn and R.L. Tweedie. Stability of Markovian processes III: Foster-Lyapunov criteria for continuous-time processes. Advances in Applied Probability, 25(3):518–548, 1993.

[25] A. Pazy. Semigroups of Linear Operators and Applications to Partial Differential Equations. Springer-verlag edition, 1995.

[26] S.T. Rachev. Probability Metrics and the Stability of Stochastic Models. John Wiley & Sons, 1991.

[27] S. Roelly. A criterion of convergence of measure-valued processes: Application to measure branching processes. Stochastics, 17:43–65, 1986.

[28] S. Roelly and A. Rouault. Construction et propriétés de martingales des branchements spatiaux interactifs. International Statistical Review, 58(2):173–189, 1990.

[29] G.R. Shorack and J.A. Wellner. Empirical Processes with Applications to Statistics. Wiley, New-York, 1986.

[30] J.R. Speakman. Body size, energy metabolism and lifespan. The Journal of Experimental Biology, 208:1717–1730, 2005.

[31] V.C. Tran. Large population limit and time behaviour of a stochastic particle model describing an age-structured population. ESAIM: P&S, 12:345–386, 2008.

[32] F.J.S. Wang. A central limit theorem for age- and density-dependent population processes. Stochastic Processes and their Applications, 5:173–193, 1977.

[33] G.F. Webb. Theory of Nonlinear Age-Dependent Population Dynamics, volume 89 of Monographs and Textbooks in Pure and Applied mathematics. Marcel Dekker, inc., New York - Basel, 1985.