Adv. Appl. Prob. **49,** 549–580 (2017)
*doi:*10.1017/apr.2017.12

LIMIT THEOREMS FOR SOME BRANCHING MEASURE-VALUED PROCESSES

BERTRAND CLOEZ,* *INRA*

Abstract

We consider a particle system in continuous time, a discrete population, with spatial motion, and nonlocal branching. The offspring's positions and their number may depend on the mother's position. Our setting captures, for instance, the processes indexed by a Galton–Watson tree. Using a size-biased auxiliary process for the empirical measure, we determine the asymptotic behaviour of the particle system. We also obtain a large population approximation as a weak solution of a growth-fragmentation equation. Several examples illustrate our results. The main one describes the behaviour of a mitosis model; the population is size structured. In this example, the sizes of the cells grow linearly and if a cell dies then it divides into two descendants.

Keywords: Branching Markov process; limit theorem; many-to-one formula; size-biased reproduction distribution; eigenproblem; deterministic macroscopic approximation; growth-fragmentation equation

2010 Mathematics Subject Classification: Primary 60J80
Secondary 60J75; 60F15; 60F05; 45C05; 92D25

1. Introduction

In this paper we study the evolution of a Markov process indexed by a tree in continuous time. The tree can represent a population of cells, polymers, or particles. On this population, we consider the evolution of an individual characteristic. This characteristic can represent the size, the age, or the rate of a nutrient. During the life of an individual, its characteristic evolves according to an underlying Markov process. At nonhomogeneous time, the individuals die and divide. The offspring's characteristics depend on the mother's and on the number of children. This model was studied in [1], [4]–[6], [17], and [23]. Here, we study the asymptotic behaviour of the empirical measure which describes the population. Following [5], we begin to prove a many-to-one formula (also known as, for example, spinal decomposition or tagged fragment) and then deduce its long-time behaviour. This formula looks like the Wald formula and reduces the problem to the study of a 'typical' individual. Closely related, we can find a limit theorem in discrete time as in [11], in continuous time with a continuous population as in [16] and for a space-structured population model as in [17]. Our approach is closer to [5] and extends their law of large number to a variable rate of division. This extension is essential in application [4]. In our model, the population is discrete. It is the microscopic version of some deterministic equations studied in [31], [39], and [40]. Following [20] and [44], we scale our empirical measure and prove that these partial differential equations are macroscopic versions

Received 22 September 2015; revision received 15 July 2016.
* Postal address: INRA Montpellier, UMR MISTEA, Bâtiment 29, 2 Place Pierre Viala, 34060 Montpellier Cedex 1, France. Email address: bertrand.cloez@supagro.inra.fr

of our model. Before expressing our main results, we begin with some notation. If we startwith one individual then we will use the Ulam–Harris–Neveu notation [5]:

- the first individual is labelled by $\varnothing$; when the individual u divides then it has a random number of descendants, denoted by K^u, who are labelled by $u1, \dots, uK^u$;
- set $\mathcal{U} = \bigcup_{m\geq 0}(\mathbb{N}^*)^m$, where $\mathbb{N} = \{0, 1, \dots\}$, $\mathbb{N}^* = \{1, 2, \dots\}$, and $(\mathbb{N}^*)^0 = \{\varnothing\}$; we denote by $\mathcal{V}_t \subset \mathcal{U}$ the set of individuals which are alive at time t; we denote by N_t the number of individuals alive at time t;
- we denote by $\mathcal{T} \subset \mathcal{U}$ the random set of individuals which are dead, alive, or will be alive; for each $u \in \mathcal{T}$, $\alpha(u)$, $\beta(u)$, and $(X_t^u)_{t\in[\alpha(u),\, \beta(u))}$ denote respectively the birth date and the death date of the individual u.

The dynamics of our model are then as follows.

- The characteristic of the first individual, $(X_t^{\varnothing})_{t\in[0,\beta(\varnothing))}$ is distributed according to an underlying càdlàg strong Markov process $(X_t)_{t\geq 0}$ starting from $X_0^{\varnothing}$. For the sake of simplicity, we will assume that $X = (X_t)_{t\geq 0}$ is a Feller process, takes values in a subset E of $\mathbb{R}^d$, and is generated by

$$Gf(x) = \phi(x) \cdot \nabla f(x) + \sigma \Delta f(x) \tag{1.1}$$

  for every f in the domain $\mathcal{D}(G)$ of G, where $d \in \mathbb{N}^*$, $\phi\colon \mathbb{R}^d \to \mathbb{R}^d$ is a C^∞ and Lipschitz function, and $\sigma \in \mathbb{R}_+$. It is well known that $\mathcal{D}(G)$ contains the set $C_c^2(E)$ of C^2 functions with compact support. Note that our approach is available for another underlying dynamic.

- The death time $\beta(\varnothing)$ of the first individual satisfies

$$\mathbb{P}(\beta(\varnothing) > t \mid X_s^{\varnothing},\ s \leq t) = \exp\biggl(-\int_0^t r(X_s^{\varnothing})\, \mathrm{d}s\biggr),$$

  where r is a nonnegative, measurable, and locally bounded function. Note that $\alpha(\varnothing) = 0$. We assume that for any starting distribution, we have

$$\int_0^\infty r(X_s^{\varnothing})\, \mathrm{d}s = +\infty \quad \text{almost surely.} \tag{1.2}$$

  This ensures that whatever the initial condition, the death time is almost surely finite.

- At time $\beta(\varnothing)$, the first individual splits into a random number of children given by an independent random variable $K^{\varnothing}$ of law $(p_k(X_{\beta(\varnothing)-}^{\varnothing}))_{k\in\mathbb{N}^*}$. For every $k \in \mathbb{N}$, the mapping $x \mapsto p_k(x)$ is continuous and, for every $x \in E$, $(p_k(x))_{k\geq 0}$ is a vector whose coordinates are nonnegative and sum to 1. We have $\alpha(1) = \dots = \alpha(K^{\varnothing}) = \beta(\varnothing)$.
- We assume that the mean offspring number, which is defined by $m\colon x \mapsto \sum_{k\geq 0} kp_k(x)$, is locally bounded on E.
- The characteristics of the new individuals are given by $(F_j^{(K^{\varnothing})}(X_{\beta(\varnothing)-}^{\varnothing}, \Theta))_{1\leq j\leq K^{\varnothing}}$, where Θ is a uniform variable on $[0, 1]$. The sequence $(F_j^{(k)})_{j\leq k, k\in\mathbb{N}^*}$ is supposed to be a family of continuous functions.
- Finally, the children evolve independently from each other as the first individual.

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

The last point is the branching property. In [5], the cell's death rate $x \mapsto r(x)$ and the law of the number of descendants $x \mapsto (p_k(x))_{k\geq 1}$ are constant (that is, do not depend on x) and a many-to-one formula is proved: for every continuous and bounded function f, we have

$$\frac{1}{\mathbb{E}[N_t]}\mathbb{E}\bigg[\sum_{u\in\mathcal{V}_t} f(X_t^u)\bigg] = \mathbb{E}[f(Y_t)], \tag{1.3}$$

where Y is generated, for every function $f \in \mathcal{D}(G)$ and $x \in E$, by

$$A_0 f(x) = Gf(x) + rm\sum_{k\geq 1}\frac{kp_k}{m}\int_0^1 \frac{1}{k}\sum_{j=1}^{k}(f(F_j^{(k)}(x,\theta)) - f(x))\,\mathrm{d}\theta. \tag{1.4}$$

This process evolves as X, until it jumps, at an exponential time with mean $1/rm$. We observe that r is not the jump rate of the auxiliary process. There is a biased phenomenon; note that the children distribution is size biased; see [5], [23], and the references therein. We can interpret it by the fact that the faster the cells divide, the more descendants they have; also, the more prolific the cells are, the more representative they are. That is why a uniformly chosen individual has an accelerated rate of division and a size-biased reproduction law. A possible generalisation of (1.3) is a Feynman–Kac formula as in [23]: for every continuous and bounded function f, we have

$$\mathbb{E}\bigg[\sum_{u\in\mathcal{V}_t} f(X_t^u)\bigg] = \mathbb{E}\bigg[f(Y_t)\exp\bigg(\int_0^t r(Y_s)(m(Y_s)-1)\,\mathrm{d}s\bigg)\bigg],$$

where Y is an auxiliary process generated by (1.4). In [4], another representation of the empirical measure was used to prove the extinction of a parasite population. However, it is difficult to exploit these formulae. Inspired by [16], [31], [39], [40], we follow an alternative approach. In (1.3), Y can be understood as a uniformly chosen individual. The problem is: if r is not constant then a uniformly chosen individual does not follow a homogeneous Markovian dynamic. Our solution is to choose this individual with an appropriate weight. This weight is a positive eigenvector V of the operator $\mathcal{G}$ defined, for every $f \in \mathcal{D}(G)$ and $x \in E$, by

$$\mathcal{G}f(x) = Gf(x) + r(x)\bigg[\bigg(\sum_{k\geq 0}\sum_{j=1}^{k}\int_0^1 f(F_j^{(k)}(x,\theta))\,\mathrm{d}\theta\; p_k(x)\bigg) - f(x)\bigg].$$

It is not the generator of a (conservative) Markov process on E but it is related to the branching mechanism; see Lemma 2.2. Under some assumptions, we are able to prove that the following weighted many-to-one formula holds:

$$\frac{1}{\mathbb{E}[\sum_{u\in\mathcal{V}_t} V(X_t^u)]}\mathbb{E}\bigg[\sum_{u\in\mathcal{V}_t} f(X_t^u)V(X_t^u)\bigg] = \mathbb{E}[f(Y_t)], \tag{1.5}$$

where Y is an auxiliary Markov process generated by

$$A = M + J, \tag{1.6}$$

where M describes the motion between the jumps and is defined by

$$Mf(x) = \frac{G(f\times V)(x) - f(x)GV(x)}{V(x)} = Gf(x) + 2\sigma\frac{\nabla V(x)\cdot\nabla f(x)}{V(x)},$$

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

and J describes the jump dynamics and is given by

$$Jf(x) = \Lambda(x)\Bigg[\frac{\sum_{k\in\mathbb{N}}\sum_{j=1}^{k}\int_0^1 V(F_j^{(k)}(x,\theta))f(F_j^{(k)}(x,\theta))\,\mathrm{d}\theta\; p_k(x)}{\sum_{k\in\mathbb{N}}\sum_{j=1}^{k}\int_0^1 V(F_j^{(k)}(x,\theta))\,\mathrm{d}\theta\; p_k(x)} - f(x)\Bigg],$$

where

$$\Lambda(x) = \Bigg[\sum_{k\in\mathbb{N}}\sum_{j=1}^{k}\int_0^1 V(F_j^{(k)}(x,\theta))\,\mathrm{d}\theta p_k(x)\Bigg] \times \frac{r(x)}{V(x)}$$

for every $f \in C_c^2(E)$ and $x \in E$. These formulae seem to be complicated but they are very simple when applied. Contrary to the previous bias, this one is present in the motion and the branching mechanism. It has already been observed in another context; see [16]. Also note that we do not assume that V is associated to the first eigenvalue. It is then possible to have different many-to-one formulae as can be seen in Remark 3.4. Some criteria for the existence of eigenelements can be found in [3], [12], [37], [41], and the references therein.

If Y is ergodic with invariant measure π then from (1.5), we have

$$\lim_{t\to+\infty}\frac{1}{\mathbb{E}[\sum_{u\in\mathcal{V}_t} V(X_t^u)]}\mathbb{E}\Bigg[\sum_{u\in\mathcal{V}_t} f(X_t^u)V(X_t^u)\Bigg] = \int f\,\mathrm{d}\pi$$

for all continuous and bounded function f. We improve this result.

Theorem 1.1. (Long-time behaviour of the empirical measure.) *If the following assumptions hold:*

(i) $X_0^{\varnothing} = x \in E$ *is deterministic;*

(ii) *the system is nonexplosive; namely,* $N_t < +\infty$ *almost surely for all* $t \geq 0$*;*

(iii) *there exists* (V, λ_0) *such that* $\mathcal{G}V = \lambda_0 V$, $V > 0$, *and* V *is* C^2*;*

(iv) Y *is a Feller process and is ergodic with invariant measure* π*;*

(v) *there exists* $\alpha < \lambda_0$ *and* C_x, *such that* $\mathbb{E}_x[V^2(Y_t)] \leq C_x \mathrm{e}^{\alpha t}$ *and*

$$\mathbb{E}_x\Bigg[\frac{r(Y_t)}{V(Y_t)}\int_0^1 \sum_{a,b\in\mathbb{N}^*,\, a\neq b}\ \sum_{k\geq\max(a,b)} p_k(Y_t)V(F_a^{(k)}(Y_t,\theta))V(F_b^{(k)}(Y_t,\theta))\,\mathrm{d}\theta\Bigg] \leq C_x\mathrm{e}^{\alpha t};$$

then for any continuous function g *such that* g/V *is bounded, we have*

$$\lim_{t\to+\infty}\mathrm{e}^{-\lambda_0 t}\sum_{u\in\mathcal{V}_t} g(X_t^u) = W\int\frac{g}{V}\,\mathrm{d}\pi,$$

where $W = \lim_{t\to+\infty}\mathrm{e}^{-\lambda_0 t}V(x_0)^{-1}\sum_{u\in\mathcal{V}_t}V(X_t^u)$ *and the convergences hold in probability. If, furthermore,* V *is lower bounded by a positive constant then*

$$\lim_{t\to+\infty}\frac{1_{\{W\neq 0\}}}{N_t}\sum_{u\in\mathcal{V}_t} g(X_t^u) = 1_{\{W\neq 0\}}\int\frac{g}{V}\,\mathrm{d}\pi\Big/\int\frac{1}{V}\,\mathrm{d}\pi \quad \textit{in probability},$$

where $1_{\{\cdot\}}$ *is the indicator function.*

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

Assumption (1.1) seems difficult to apply but if r, V, and $(p_k)_{k\geq 0}$ are polynomials then it is enough to prove the finiteness of the moments of $(Y_t)_{t\geq 0}$. Unfortunately, in general, it is difficult to derive such properties on V from assumptions on the branching mechanism.

As a direct application, if the maps $x \mapsto r(x)$ and $x \mapsto p_k(x)$ are constant, then $V \equiv 1$ is an eigenvector, and so this theorem generalises [5, Theorem 1.1]. Some inhomogeneous examples are developed in Sections 5 and 6.

On the other hand, our model can be seen as a microscopic version of some deterministic models. More precisely, let $(\boldsymbol{Z}_t)_{t\geq 0}$ be the empirical measure. It is defined, for all $t \geq 0$, by

$$\boldsymbol{Z}_t = \sum_{u\in\mathcal{V}_t} \delta_{X_t^u}.$$

Now let $\boldsymbol{Z}^{(n)}$ evolve as $\boldsymbol{Z}$, but with $\boldsymbol{Z}_0^{(n)}$ dependent on n and set $\boldsymbol{X}^{(n)} = \boldsymbol{Z}^{(n)}/n$.

Theorem 1.2. (Law of large number for the large population.) *If the following assumptions hold:*

(i) $T > 0$, *r is upper bounded, and there exist $\bar{k} \geq 0$ such that $p_k \equiv 0$ for all $k \geq \bar{k}$;*

(ii) *either E is compact or $E \subset \mathbb{R}$, $F_j^{(k)}(x,\theta) \leq x$ for all $j \leq k$ and $\theta \in [0,1]$, and $\sup_{x\in E}\phi(x) < +\infty$;*

(iii) *then (1.7) below admits a unique solution;*

(iv) *the starting distribution $\boldsymbol{X}_0^{(n)}$ converges in distribution to $\boldsymbol{X}_0 \in \mathcal{M}(E)$, embedded with the weak topology;*

(v) *we have $\sup_{n\geq 0}\mathbb{E}[\boldsymbol{X}_0^{(n)}(E)] < +\infty$; then $\boldsymbol{X}^{(n)}$ converges in distribution in $\mathbb{D}([0,T],\mathcal{M}(E))$ to $\boldsymbol{X}$, which satisfies*

$$\int_E f(x)\boldsymbol{X}_t(\mathrm{d}x) = \int_E f(x)\boldsymbol{X}_0(\mathrm{d}x) + \int_0^t \int_E \mathcal{G}f(x)\boldsymbol{X}_s(\mathrm{d}x)\,\mathrm{d}s. \tag{1.7}$$

Here, $\mathbb{D}([0,T],\mathcal{M}(E))$ is the space of càdlàg functions, with values in the set $\mathcal{M}(E)$ of finite measures on E, embedded with the Skorokhod topology [7], [27]. We observe that if $\boldsymbol{X}_0$ is deterministic then $\boldsymbol{X}_t$ is deterministic for any time $t \geq 0$. In a weak sense, (1.7) can be written as

$$\begin{aligned}\partial_t n(t,x) &+ \nabla(\phi(x)n(t,x)) + r(x)n(t,x)\\ &= \sigma\partial_{xx}n(t,x) + \sum_{k\geq 0}\sum_{j=1}^{k} K_j^k(r \times p_k \times n(t,\cdot)),\end{aligned} \tag{1.8}$$

where $\boldsymbol{X}_t = n(t,x)\,\mathrm{d}x$ and K_j^k is the adjoint operator of $f \mapsto \int_0^1 f(F_j^{(k)}(x,\theta))\,\mathrm{d}\theta$. Note that, in contrast with a classical parabolic partial differential equation, in general, the previous equation has no regularisation properties. In particular, if $\sigma = 0$ and $\boldsymbol{X}_0$ has no density, then nor does $\boldsymbol{X}_t$. This equation was studied in [31], [39], [40] and Theorem 1.1 is relatively close to their (long-time) limit theorems. We will see in the next section that it is also the Kolmogorov equation associated to $\boldsymbol{Z}$. So, we observe that $\boldsymbol{X}$ evolves as the mean measure of $\boldsymbol{Z}$; that is, $f \mapsto \mathbb{E}[\int_E f(x)\boldsymbol{Z}_t(\mathrm{d}x)]$. This average phenomenon comes from the branching property. After a branching event, each cell evolves independently from the others, there is no interaction. Another interpretation is the linearity of the operator $\mathcal{G}$. From the many-to-one formula, we also deduce that, in a large population, the empirical measure behaves as the auxiliary process. The proof is based on the Aldous–Rebolledo criterion [27], [43], and is inspired by [20], [35], and [44].

At the end of the paper, these two theorems are applied to some structured population models. Our main example is a size-structured population; the size of cells grows linearly and if they divide into two descendants. Thus, there is a motion between the branching events and discontinuity at division times. This model is a branching version of the well-known TCP (transmission control protocol) windows size process [8], [22], [32], [38]. We are able to give some explicit formulae of the invariant distribution, the moments, and the rate of convergence. Also, we prove a central limit theorem for the limit in a large population.

Outline. In the next section we introduce some properties of the empirical measure. In Section 3 we focus our interest on the long-time behaviour. We prove some many-to-one formula and deduce a general limit theorem which implies Theorem 1.1. Section 4 is devoted to the study of large populations. In this section, we prove Theorem 1.2. Note that Section 3 and Section 4 are independent. In Section 5 we give our main example, which describes the cell mitosis. Moreover, we give two theorems for the long-time behaviour of our empirical measure in addition to some explicit formula. We also give a central limit theorem for asymmetric cell division for the large population limit. In Section 6 we conclude with two classical examples which are branching diffusions and self-similar fragmentation.

2. Preliminaries

In this section we describe a little more the empirical measure $(\mathbf{Z}_t)_{t\geq 0}$. We recall that, for all $t \geq 0$, $\mathbf{Z}_t = \sum_{u\in\mathcal{V}_t} \delta_{X_t^u}$. Let us add the following notation:

$$\mathbf{Z}_t(f) = \int_E f(x)\mathbf{Z}_t(\mathrm{d}x) = \sum_{u\in\mathcal{V}_t} f(X_t^u), \qquad \mathbf{Z}_t(1+\|x\|^p) = \int_E (1+\|x\|^p)\mathbf{Z}_t(\mathrm{d}x),$$

for every continuous and bounded function f and for every $p \geq 0$. We can describe the dynamics of the population with a stochastic differential equation [25]. Let $C_c^{2,1}(E,\mathbb{R}_+)$ be the set of functions $f\colon (x,t)\mapsto f(x,t) = f_t(x)$ that are C^1 in time, with bounded derivative, such that $f_t \in C_c^2(E)$. For any function f belonging to $C_c^{1,2}(E\times\mathbb{R}_+)$, we have

$$\begin{aligned}\mathbf{Z}_t(f_t) = \mathbf{Z}_0(f_0) &+ \int_0^t \int_E \mathcal{G} f_s(x) + \partial_s f_s(x)\mathbf{Z}_s(\mathrm{d}x)\,\mathrm{d}s \\ &+ \int_0^t \sum_{u\in\mathcal{V}_s} \sqrt{2\sigma}\,\partial_x f_s(X_s^u)\,\mathrm{d}B_s^u \\ &+ \int_0^t \int_{\mathcal{U}\times\mathbb{R}_+\times\mathbb{N}\times[0,1]} \bigg[1_{\{u\in\mathcal{V}_{s-},\, l\leq r(X_{s-}^u)\}} \\ &\qquad \times \bigg(\sum_{j=1}^k f_s(F_j^{(k)}(X_{s-}^u,\theta))\bigg) - f_s(X_{s-}^u)\bigg]\rho(\mathrm{d}s,\mathrm{d}u,\mathrm{d}l,\mathrm{d}k,\mathrm{d}\theta),\end{aligned}$$

where $(B^u)_{u\in\mathcal{U}}$ is a family of independent standard Brownian motions and $\rho(\mathrm{d}s,\mathrm{d}u,\mathrm{d}l,\mathrm{d}k,\mathrm{d}\theta)$ is Poisson point process on $\mathbb{R}_+\times\mathcal{U}\times\mathbb{R}_+\times\mathbb{N}\times[0,1]$ of intensity

$$\bar{\rho}(\mathrm{d}s,\mathrm{d}u,\mathrm{d}l,\mathrm{d}k,\mathrm{d}\theta) = \mathrm{d}s\, n(\mathrm{d}u)\mathrm{d}l\mathrm{d}p_k\mathrm{d}\theta.$$

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

It is also independent from the Brownian motions. We have denoted by $n(\mathrm{d}u)$ the counting measure on $\mathcal{U}$ and $\mathrm{d}s, \mathrm{d}l$, and $\mathrm{d}\theta$ are Lebesgue measures. The fist hypothesis ensures the nonexplosion (and is a little bit stronger) of our model.

Assumption 2.1. (Nonexplosion.) *For all* $t \geq 0$, $\sup_{s \leq t} \mathbb{E}[N_s] < +\infty$.

Lemma 2.1. (Sufficient condition to nonexplosion.) *If there exists* $\bar{r}, \bar{k} > 0$ *such that, for every* $x \in E$, *we have*

$$r(x) \leq \bar{r} \quad \textit{and} \quad p_k(x) = 0 \quad \text{for all } k \geq \bar{k},$$

then Assumption 2.1 holds. Moreover, for any $T > 0$, *we have, for all* $t \leq T$,

$$\mathbb{E}[N_t] \leq \mathbb{E}[N_0]\mathrm{e}^{(\bar{k}-1)\bar{r}T}.$$

Proof. We can easily couple the number of particles $(N_t)_{t\geq 0}$ with a branching process $(W_t)_{t\geq 0}$, which does not depend on the underlying dynamics, to obtain

$$N_t \leq W_t \quad \text{for all } t \geq 0.$$

With rate $\bar{r}$, the process $(W_t)_{t\geq 0}$ produces $\bar{k}-1$ individuals, hence, its mean is

$$\mathbb{E}[W_t] = \mathrm{e}^{(\bar{k}-1)\bar{r}t} \quad \text{for all } t \geq 0.$$

See [2] for more details. □

If there is no explosion then we have a semimartingale decomposition.

Lemma 2.2. (Semimartingale decomposition.) *If Assumption 2.1 holds then, for all* $f \in \mathcal{D}(G)$, *we have*

$$\boldsymbol{Z}_t(f) = \boldsymbol{Z}_0(f) + \boldsymbol{M}_t(f) + \boldsymbol{V}_t(f),$$

where $(\boldsymbol{M}_t(f))_{t\geq 0}$ *is a semimartingale and*

$$\boldsymbol{V}_t(f) = \int_0^t \boldsymbol{Z}_s(\mathcal{G}f)\,\mathrm{d}s,$$

and if, furthermore, $f \in C_c^2(E)$ *then the bracket of* $\boldsymbol{M}_t(f)$ *is given by*

$$\begin{aligned}\langle \boldsymbol{M}(f)\rangle_t = \int_0^t &(G(f_s^2)(x) - 2f_s(x)Gf_s(x))\boldsymbol{Z}_s(\mathrm{d}x)\\ &+ \int_E r(x)\sum_{k\in\mathbb{N}}\int_0^1\bigg(\sum_{j=1}^k f_s(F_j^{(k)}(x,\theta)) - f_s(x)\bigg)^2 p_k(x)\,\mathrm{d}\theta\,\boldsymbol{Z}_s(\mathrm{d}x)\,\mathrm{d}s\end{aligned}$$

Proof. This follows by an application of the Dynkin and Itô formulae; see, for instance, [26, Lemma 3.68, p. 487] and [25, Theorem 5.1, p. 67]. If $f \in C_c^2(E)$ then, by assumption 2.1, $(\boldsymbol{M}_t(f))_{t\geq 0}$ is a square-integrable martingale. Indeed, if $f \in C_c^2(E)$ and $\tau_n = \inf\{s \geq 0 \mid \boldsymbol{M}_t(f)^2 \geq n\}$, then there exists $C > 0$ (only depending on f) such that $\langle \boldsymbol{M}(f)\rangle_{t\wedge\tau_n} \leq C\int_0^{t\wedge\tau_n} N_s\,\mathrm{d}s \leq \int_0^t N_s\,\mathrm{d}s$. Using the Fatou lemma, Fubini–Tonelli theorem, and the optional stopping theorem, we have

$$\mathbb{E}[\boldsymbol{M}_t(f)^2] \leq \liminf_{n\to\infty}\mathbb{E}[\boldsymbol{M}_{t\wedge\tau_n}(f)^2] = \liminf_{n\to\infty}\mathbb{E}[\langle \boldsymbol{M}(f)\rangle_{t\wedge\tau_n}] \leq C\int_0^t \mathbb{E}[N_s]\,\mathrm{d}s. \qquad \square$$

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

Let us fix $\mu = \sum_{i=1}^{m} \delta_{x_i}$, where $m \in \mathbb{N}^*$ and $x_i \in E$ for all $i \in \{1, \dots, m\}$. We define the mean measure $(z_t)_{t \geq 0}$, for any continuous and bounded function f on E, by, for all $t \geq 0$,

$$z_t(f, \mu) = \mathbb{E}(\mathbf{Z}_t(f) \mid \mathbf{Z}_0 = \mu) = \mathbb{E}\bigg[\sum_{u \in \mathcal{V}_t} f(X_t^u) \mid \mathbf{Z}_0 = \mu\bigg].$$

The measure $z_t(\mathrm{d}x, \mu)$ is defined by $\int_E f(x) z_t(\mathrm{d}x, \mu) = z_t(f, \mu)$. Moreover, the branching property ensures that $z_t(f, \mu) = \sum_{i=1}^{m} z_t(f, \delta_{x_i})$. In application, we have $z_t(f, \mu) = \int_E z_t(f, \delta_x)\mu(\mathrm{d}x)$. If μ is a probability measure then $z_t(f, \mu) = \mathbb{E}[\mathbf{Z}_t(f)]$, where $\mathbf{Z}_0 = \delta_{X_0^\varnothing}$ and $X_0^\varnothing$ is a random variable distributed by μ.

Corollary 2.1. (Evolution equation for the mean measure.) *Under Assumption 2.1, if* $f \in C_c^2$, $\mu \in \mathcal{M}(E)$, *and* $t \geq 0$, *then we have*

$$z_t(f, \mu) = \mu(f) + \int_0^t \bigg(z_s(Gf, \mu) + \int_E r(x) \bigg(\sum_{k \geq 0} \sum_{j=1}^{k} \int_0^1 f(F_j^{(k)}(x, \theta)) \, \mathrm{d}\theta p_k(x) \bigg) - f(x) z_s(\mathrm{d}x, \mu) \bigg) \mathrm{d}s. \tag{2.1}$$

The previous equation can be written as (1.8) in a weak sense; namely, n is defined by $n(t, x)\,\mathrm{d}x = z_t(\mathrm{d}x, \mu), t \geq 0$.

Remark 2.1. (*Uniqueness of (2.1).*) As illustrated in [15] for instance, the question of uniqueness for evolution equations is generally nontrivial. Think, for instance, of the reflected and absorbed Brownian motions on $(0, 1)$ whose semigroups have the same generator Δ when applied on smooth functions. Nevertheless, if (2.1) holds for $f = 1$ then using the following norm (on the space of signed measure):

$$\|z\|_0 = \sup\{|z(f)| \mid \|f\|_\infty + \|Gf\|_\infty \leq 1\},$$

and the Gronwall lemma in (2.1) ensures the uniqueness.

3. Long-time behaviour

Let us recall that

$$\mathcal{G} f(x) = Gf(x) + r(x) \bigg[\bigg(\sum_{k \geq 0} \sum_{j=1}^{k} \int_0^1 f(F_j^{(k)}(x, \theta)) \, \mathrm{d}\theta p_k(x) \bigg) - f(x) \bigg]$$

for every $f \in \mathcal{D}(G)$ and $x \in E$. In the following, we will prove some formula which characterise the mean behaviour of our model. Then we will use them to prove our limit theorems.

3.1. Eigenelements and auxiliary process

As said in introduction, the existence of eigenelements is fundamental in our approach. Nevertheless the eigenvector does not still belong to the domain of the generator. Henceforth, we assume the following.

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

Assumption 3.1. (Existence of eigenelements.) *Assumption 2.1 holds, and there exist* $\lambda_0 > 0$ *and a continuous and positive function* V *such that there exists a sequence* $(V_n)_{n\geq 0}$ *of functions belonging to* $C_c^2(E)$ *satisfying, for all* $x \in E$,

$$\lim_{n\to\infty} V_n(x) = V(x), \qquad \lim_{n\to\infty} \mathcal{G} V_n(x) = \lambda_0 V(x)$$

and the mappings $x \mapsto \sup_{n\geq 0} V_n(x)$ *and* $x \mapsto \sup_{n\geq 0} \mathcal{G} V_n(x)$ *are integrable with respect to* $z_t(\mathrm{d}x, \delta_{x_0})$ *for every* $t \geq 0$ *and* $x_0 \in E$.

Remark 3.1. (*Smooth eigenvector.*) If there exists a smooth eigenvector $V \in C_c^2(E)$ then we can choose $V_n = V$ for every n. Also if V is C^2 then a trivial truncation argument ensures the previous assumption. This assumption enables us to consider less regular eigenvectors. The integrability condition is essentially a consequence of V being an eigenfunction. Indeed, it can be proved using Lemma 2.2 and the suitable sequence of stopping times $(\tau_n)_{n\geq 0}$ defined by $\tau_n = \inf\{t \geq 0,\ \mathbf{Z}_t(V) \geq n\}$. See, for instance, Lemma 5.1, where a similar computation is done.

Under Assumption 3.1, we introduce the martingale $(\mathbf{Z}_t(V)\mathrm{e}^{-\lambda_0 t})_{t\geq 0}$ which plays an important role in the proof of Theorem 1.1.

Lemma 3.1. (Martingale properties.) *If Assumption 3.1 holds and*

$$\mathbb{E}[\mathbf{Z}_0(V)] < +\infty,$$

then the process $(\mathbf{Z}_t(V)\mathrm{e}^{-\lambda_0 t})_{t\geq 0}$ *is a martingale. Moreover, it converges almost surely to a random variable* W.

Proof. By corollary 2.1 and the dominated convergence theorem, we have

$$z_t(V, \mathbf{Z}_0) = \lim_{n\to\infty} z_0(V_n, \mathbf{Z}_0) + \int_0^t z_s(\mathcal{G} V_n, \mathbf{Z}_0)\,\mathrm{d}s = z_0(V, \mathbf{Z}_0) + \lambda_0 \int_0^t z_s(V, \mathbf{Z}_0)\,\mathrm{d}s.$$

Hence, for all $t \geq 0$, we have $z_t(V, \mathbf{Z}_0) = z_0(V, \mathbf{Z}_0)\mathrm{e}^{\lambda_0 t} = \mathbf{Z}_0(V)\mathrm{e}^{\lambda_0 t}$. Then if $\mathcal{F}_t = \sigma\{\mathbf{Z}_s \mid s \leq t\}$ then the Markov property, applied on $\mathbf{Z}$, yields

$$\mathbb{E}[\mathbf{Z}_{t+s}(V)\mathrm{e}^{-\lambda_0(t+s)} \mid \mathcal{F}_s] = \mathrm{e}^{-\lambda_0(t+s)} z_t(V, \mathbf{Z}_s) = \mathbf{Z}_s(V)\mathrm{e}^{-\lambda_0 s},$$

and so the process $(\mathbf{Z}_t(V)\mathrm{e}^{-\lambda_0 t})_{t\geq 0}$ is a martingale. Since it is a positive, it converges almost surely. □

To have our many-to-one formula, we add the following natural assumption.

Assumption 3.2. (Auxiliary process.) *The operator* A *defined in (1.6) is the generator of a Feller process.*

Lemma 3.2. (Weighted many-to-one formula.) *Under Assumptions 3.1 and 3.2, if* μ *is a probability measure satisfying* $\mu(V) < +\infty$ *then we have*

$$\frac{1}{z_t(V, \mu)} z_t(f_t \times V, \mu) = \int_E \mathbb{E}[f(Y_t, t) \mid Y_0 = x]\mu(\mathrm{d}x) \tag{3.1}$$

for any nonnegative function $f\colon (x, t) \mapsto f(x, t) = f_t(x)$ *on* $E \times \mathbb{R}_+$ *and* $t \geq 0$, *where* Y *is a Markov process generated by* A *starting from* x_0.

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

Proof. As the time variable t is deterministic, it is enough to prove that (3.1) holds for any function that are not time dependent. Let $(\gamma_t)_{t\geq 0}$ be the family of operators defined by

$$\gamma_t(f) = z_t(f \times V, \mu)\mathrm{e}^{-\lambda_0 t}\mu(V)^{-1}$$

for every $f \in \mathcal{D}$ and $t \geq 0$. Using Lemma 2.2, we have, for all $t \geq 0$ and $f \in \mathcal{D}(G)$,

$$\partial_t \gamma_t(f) = z_t(\mathcal{G}(Vf) - f\mathcal{G}V, \mu)\mathrm{e}^{-\lambda_0 t}\mu(V)^{-1} = \gamma_t(Af).$$

Now, by the Itô–Dynkin formula, the right-hand side of (3.1) satisfies the same equation. Uniqueness comes from classical arguments. Indeed, by Assumption 3.2, operator A is the generator of a Feller semigroup and satisfies the positive maximum principle; see [29, Theorem 3.6.6]. As a consequence, if $(P_t)_{t\geq 0}$ is the semigroup of the auxiliary process then $P_t = \gamma_t$ by [18, Proposition 9.19, Chapter 10]; see also [29, Theorem 4.1.2]. □

Remark 3.2. (*Schrödinger operator and h-transform.*) The operator $\mathcal{G}$ is not a (conservative) Markov generator. Indeed, for all $f \in C_c^2(E)$, $\mathcal{G}f = Bf - r(m-1)f$, where B is a Markov generator. Operator $\mathcal{G}$ is sometimes called a Schrödinger operator. Its study is connected to the Feynman–Kac formula. Our weighted many-to-one formula can be seen as an h-transform (Girsanov type transformation) of the Feynman–Kac semigroup as in [41]. This transformation is usual in the superprocesses study [16].

Remark 3.3. (*Galton–Watson tree and Malthus parameter.*) If the maps $x \mapsto r(x)$ and $x \mapsto p_k(x)$ are constant then $V \equiv 1$ is an eigenvector with respect to the eigenvalue $\lambda_0 = r(m-1)$, where $m = \sum_{k\geq 0} kp_k$ denotes the mean offspring number. So, $\boldsymbol{Z}_t(V) = N_t$ and the size of the population grows exponentially when it survives. This result is already know for N_t; see [2], [5]. Furthermore, since Malthus introduced the following simple model to describe the population evolution: $\partial_t N_t = (b_0 - d_0)N_t$, namely $N_t = \mathrm{e}^{(b_0-d_0)t}$, where b_0 and d_0 represent, respectively, the individual birth rate and death rate, in biology and genetic population study, $\lambda_0 = b_0 - d_0$ is sometimes called the Malthus parameter.

Remark 3.4. (*Many eigenelements are possible.*) In the previous lemmas, λ_0 was not required to be the first eigenvalue. So, it is possible to have different eigenelements and auxiliary processes. Consider the example of [4], where some eigenelements are explicit; that is,

$$Gf(x) = a_1 x f'(x) + b_1 x f''(x)$$

for every $f \in C_c^2(E)$ and $x \in E = \mathbb{R}_+$, where a_1, b_1 are two nonnegative numbers. We also consider that of $p_2 = 1$ and, for all $j \in \{1, 2\}$, $\mathbb{E}[f(F_j^{(2)}(x, \Theta))] = \mathbb{E}[f(Hx)]$, where H is a symmetric random variable on $[0, 1]$, that is, $H \overset{\mathrm{D}}{=} 1 - H$, where '$\overset{\mathrm{D}}{=}$' denotes equality in distribution. This example models cell division with parasite infection. In this case,

$$\mathcal{G}f(x) = a_1 x f'(x) + b_1 x f''(x) + r(x)(2\mathbb{E}[f(Hx)] - f(x))$$

for every continuous and bounded function f. Here a_1 is an eigenvalue of $\mathcal{G}$ and $V_1(x) = x$ is its eigenvector. So, we should have

$$\mathbb{E}\bigg[\sum_{u\in\mathcal{V}_t} X_t^u f(X_t^u) \mid X_0^{\varnothing} = x_0\bigg] = \mathbb{E}_{x_0}[f(Y_t)]\mathrm{e}^{a_1 t} x_0$$

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

for every continuous and bounded function f and $x_0 \in E$, where Y is a Markov process generated by G_Y defined by

$$G_Y f(x) = (a_1 x + 2b_1) f'(x) + b_1 x f''(x) + r(x)(2\mathbb{E}[H f(Hx)] - f(x))$$

for every $f \in C_c^2(E)$ and $x \in E$. We can see a bias in the drift terms and jumps mechanism which is not observed in [4], [5]. When r is affine, we obtain a second formula. Indeed, if $r(x) = c_1 x + d_1$ with $c_1 \geq 0$ and $d_1 > a_1$ (or $d_1 > 0$ and $c_1 = 0$), then $V_2(x) = x(c_1/(d_1 - a_1)) + 1$ is an eigenvector with respect to the eigenvalue $\lambda_2 = d_1$ (which implies $\lambda_2 > \lambda_1 = a_1$). A rapid calculation then gives a different many-to-one formula with another auxiliary process.

3.2. Many-to-one formulae

To compute our limit theorem, we need to control the second moment. As in [5], we describe the population over the whole tree and then give a many-to-one formula for forks. Recall that $\mathcal{T} = \{u \in \mathcal{U} \mid \text{there exists } t > 0,\, u \in \mathcal{V}_t\}$. Lemmas 3.3 and 3.4 that follow are, respectively, the generalisation of [5, Proposition 3.5] and [5, Proposition 3.9].

Lemma 3.3. (Many-to-one formula over the whole tree.) *Under Assumption 3.1, if* $\mathbf{Z}_0 = \delta_{x_0}$, *where* $x_0 \in E$, *then for any nonnegative measurable function* $f \colon E \times \mathbb{R}_+ \to \mathbb{R}$, *we have*

$$\mathbb{E}\bigg[\sum_{u \in \mathcal{T}} f(X^u_{\beta(u)-}, \beta(u))\bigg] = \int_0^{+\infty} \mathbb{E}\bigg[f(Y_s, s)\frac{r(Y_s)}{V(Y_s)}\bigg] V(x_0) \mathrm{e}^{\lambda_0 s} \,\mathrm{d}s.$$

Proof. First we have, for all $u \in \mathcal{U}$,

$$\mathbb{E}[1_{\{u \in \mathcal{T}\}} f(X^u_{\beta(u)-}, \beta(u))] = \mathbb{E}\bigg[1_{\{u \in \mathcal{T}\}} \int_{\alpha(u)}^{\beta(u)} f(X^u_s, s) r(X^u_s) \,\mathrm{d}s\bigg]$$

since, by (1.2) and the Fubini theorem,

$$\begin{aligned}
\mathbb{E}&\bigg[1_{\{u \in \mathcal{T}\}} \int_{\alpha(u)}^{\beta(u)} f(X^u_s, s) r(X^u_s) \,\mathrm{d}s\bigg] \\
&= \mathbb{E}\bigg[1_{\{u \in \mathcal{T}\}} \int_0^{+\infty} \int_{\alpha(u)}^{\tau} f(X^u_s, s) r(X^u_s) \,\mathrm{d}s\, r(X^u_\tau) \exp\bigg(-\int_{\alpha(u)}^{\tau} r(X^u_t) \,\mathrm{d}t\bigg) \mathrm{d}\tau\bigg] \\
&= \mathbb{E}\bigg[1_{\{u \in \mathcal{T}\}} \int_{\alpha(u)}^{+\infty} \int_s^{+\infty} r(X^u_\tau) \exp\bigg(-\int_{\alpha(u)}^{\tau} r(X^u_t) \,\mathrm{d}t\bigg) \mathrm{d}\tau\, f(X^u_s, s) r(X^u_s) \,\mathrm{d}s\bigg] \\
&= \mathbb{E}\bigg[1_{\{u \in \mathcal{T}\}} \int_{\alpha(u)}^{+\infty} \exp\bigg(-\int_{\alpha(u)}^{s} r(X^u_t) \,\mathrm{d}t\bigg) f(X^u_s, s) r(X^u_s) \,\mathrm{d}s\bigg] \\
&= \mathbb{E}[1_{\{u \in \mathcal{T}\}} f(X^u_{\beta(u)-}, \beta(u))].
\end{aligned}$$

Thus,

$$\mathbb{E}[1_{\{u \in \mathcal{T}\}} f(X^u_{\beta(u)-}, \beta(u))] = \mathbb{E}\bigg[\int_0^{+\infty} 1_{\{u \in V_s\}} f(X^u_s, s) r(X^u_s) \,\mathrm{d}s\bigg],$$

and finally,

$$\begin{aligned}
\mathbb{E}\bigg[\sum_{u \in \mathcal{T}} f(X^u_{\beta(u)-}, \beta(u))\bigg] &= \int_0^{+\infty} \mathbb{E}\bigg[\sum_{u \in V_s} f(X^u_s, s) r(X^u_s)\bigg] \mathrm{d}s \\
&= \int_0^{+\infty} \mathbb{E}\bigg[f(Y_s, s)\frac{r(Y_s)}{V(Y_s)}\bigg] V(x_0) \mathrm{e}^{\lambda_0 s} \,\mathrm{d}s. \qquad \square
\end{aligned}$$

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

If we set $g(x, s) = f(x, s)/V(x)$ then we have

$$\mathbb{E}\bigg[\sum_{u\in\mathcal{T}} g(X^u_{\beta(u)-}, \beta(u))V(X^u_{\beta(u)-})\bigg] = \int_0^{+\infty} \mathbb{E}[g(Y_s, s)r(Y_s)] \times \mathbb{E}[\mathbf{Z}_s(V)]\,\mathrm{d}s.$$

This equality means that adding the contributions over all the individuals corresponds to integrating the contribution of the auxiliary process over the average number of living individuals at time s. Let $(P_t)_{t\geq 0}$ be the semigroup of the auxiliary process; it is defined for any continuous and bounded f by $P_t f(x) = \mathbb{E}[f(Y_t) \mid Y_0 = x]$.

Lemma 3.4. (Many-to-one formula for forks.) *Under Assumption 3.1, if $\mathbf{Z}_0 = \delta_{x_0}$, where $x_0 \in E$, then for all nonnegative and measurable functions f, g on E, we have*

$$\mathbb{E}\bigg[\sum_{u,v\in\mathcal{V}_t,\, u\neq v} f(X^u_t)V(X^u_t)g(X^v_t)V(X^v_t)\bigg] = \mathbb{E}[\mathbf{Z}_t(V)]^2 \int_0^t \frac{1}{\mathbb{E}[\mathbf{Z}_s(V)]}\mathbb{E}\bigg[J_2(VP_{t-s}f, VP_{t-s}g)(Y_s)\frac{r(Y_s)}{V(Y_s)}\bigg]\,\mathrm{d}s,$$

where J_2 is defined by

$$J_2(f, g)(x) = \int_0^1 \sum_{a\neq b} \sum_{k\geq \max(a,b)} p_k(x) f(F_a^{(k)}(x, \theta))g(F_b^{(k)}(x, \theta))\,\mathrm{d}\theta.$$

The operator J_2 describes the starting positions of two siblings picked at random.

Proof. Let $u, v \in \mathcal{V}_t$, be such that $u \neq v$, then there exists $(w, \tilde{u}, \tilde{v}) \in \mathcal{U}^3$ and $a, b \in \mathbb{N}^*$, $a \neq b$, such that $u = wa\tilde{u}$ and $v = wb\tilde{v}$. The cell w is sometimes called the most recent common ancestor. We have

$$\mathbb{E}\bigg[\sum_{u,v\in\mathcal{V}_t,\, u\neq v} f(X^u_t)V(X^u_t)g(X^v_t)V(X^v_t)\bigg] = \sum_{w\in\mathcal{U}}\sum_{a\neq b}\sum_{\tilde{u},\tilde{v}\in\mathcal{U}} \mathbb{E}[1_{\{u\in\mathcal{V}_t\}} f(X^u_t)V(X^u_t)1_{\{v\in\mathcal{V}_t\}}g(X^v_t)V(X^v_t)].$$

Let $\mathcal{F}_t = \sigma\{\mathbf{Z}_s \mid s \leq t\}$. By the conditional independence between descendants, we have

$$\mathbb{E}\bigg[\sum_{u,v\in\mathcal{V}_t,\, u\neq v} f(X^u_t)V(X^u_t)g(X^v_t)V(X^v_t)\bigg] = \sum_{w\in\mathcal{U}}\sum_{a\neq b}\mathbb{E}\bigg[\mathbb{E}\bigg[\sum_{\tilde{u}\in\mathcal{U}} 1_{\{u\in\mathcal{V}_t\}} f(X^u_t)V(X^u_t) \mid \mathcal{F}_{\beta(w)}\bigg] \times \mathbb{E}\bigg[\sum_{\tilde{v}\in\mathcal{U}} 1_{\{v\in\mathcal{V}_t\}} g(X^v_t)V(X^v_t) \mid \mathcal{F}_{\beta(w)}\bigg]\bigg].$$

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

Therefore, as $\beta(w)$ is a stopping time, then using the strong Markov property, (3.1), and the previous lemma, we have

$$\begin{aligned}\mathbb{E}\bigg[\sum_{u,v\in\mathcal{V}_t,\,u\neq v} f(X_t^u)V(X_t^u)g(X_t^v)V(X_t^v)\bigg]\\ &= \sum_{w\in\mathcal{U}}\sum_{a\neq b}\mathbb{E}[1_{\{wa,\,wb\in\mathcal{T},\,t\geq\beta(w)\}}P_{t-\beta(w)}f(X^{wa}_{\beta(w)})V(X^{wa}_{\beta(w)})\\ &\qquad\times P_{t-\beta(w)}g(X^{wb}_{\beta(w)})V(X^{wb}_{\beta(w)})\mathrm{e}^{2\lambda_0(t-\beta(w))}]\\ &= \mathbb{E}\bigg[\sum_{w\in\mathcal{T}} 1_{\{t\geq\beta(w)\}}J_2(VP_{t-\beta(w)}f,\,VP_{t-\beta(w)}g)(X^w_{\beta(w)-})\mathrm{e}^{2\lambda_0(t-\beta(w))}\bigg]\\ &= \mathrm{e}^{2\lambda_0 t}V(x_0)\int_0^t \mathbb{E}_{x_0}\bigg[J_2(VP_{t-s}f,\,VP_{t-s}g)(Y_s)\frac{r(Y_s)}{V(Y_s)}\bigg]\mathrm{e}^{-\lambda_0 s}\,\mathrm{d}s.\end{aligned}$$

□

3.3. Proof of Theorem 1.1

In this section we give the main limit theorem which implies Theorem 1.1.

Theorem 3.1. (General condition for the convergence of the empirical measure.) *Under Assumption 3.1, if f is a measurable function defined on E and μ a probability measure such that there exists a probability measure π, two constants $\alpha < \lambda_0$ and $C > 0$, and a measurable function h such that, for all $t > 0$,*

(H1) $\pi(|f|) < +\infty$ *and for all* $x \in E$, $\lim_{s\to+\infty} P_s f(x) = \pi(f)$,

(H2) $\mu(V) < +\infty$ *and* $\mu P_t(f^2 \times V) \leq C\mathrm{e}^{\alpha t}$,

(H3) $P_t|f| \leq h$ *and* $\mu P_t(J_2(Vh, Vh)(r/V)) \leq C\mathrm{e}^{\alpha t}$,

and $\mathbf{Z}_0 = \delta_{X_0^\varnothing}$, where $X_0^\varnothing \sim \mu$, then we have

$$\lim_{t\to+\infty}\frac{1}{\mathbb{E}[\mathbf{Z}_t(V)]}\sum_{u\in\mathcal{V}_t} f(X_t^u)V(X_t^u) = W\times\pi(f),$$

where the convergence holds in probability. If, furthermore, $(\mathbf{Z}_t(V)\mathrm{e}^{-\lambda_0 t})_{t\geq 0}$ is bounded in L^2 then the convergence holds in L^2.

Note that the constants may depend on f and μ. Also note that λ_0 is not assumed to be the largest eigenvalue. The condition $\alpha < \lambda_0$ is not restrictive at all, even if λ_0 is unknown. Indeed, in all our examples, $\alpha = 0$.

Proof. As in [5, Theorem 4.2], we first prove the convergence for f such that $\pi(f) = 0$. From Lemma 3.1, we have $\mathbb{E}[\mathbf{Z}_t(V)] = \mu(V)\mathrm{e}^{\lambda_0 t}$ and so

$$\mathbb{E}\bigg[\bigg(\frac{1}{\mathbb{E}[\mathbf{Z}_t(V)]}\sum_{u\in\mathcal{V}_t} f(X_t^u)V(X_t^u)\bigg)^2\bigg] = \mathbb{E}[\mathbf{Z}_t(f\times V)^2\mathrm{e}^{-2\lambda_0 t}\mu(V)^{-2}] = A_t + B_t,$$

where, by Lemma 3.2,

$$A_t = \mathrm{e}^{-2\lambda_0 t}\mu(V)^{-2}\mathbb{E}\bigg[\sum_{u\in\mathcal{V}_t} f^2(X_t^u)V^2(X_t^u)\bigg] = \mathrm{e}^{-\lambda_0 t}\mu(V)^{-1}\mathbb{E}[f^2(Y_t)V(Y_t)],$$

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

and, by Lemma 3.4,

$$B_t = \mathrm{e}^{-2\lambda_0 t}\mu(V)^{-2}\mathbb{E}\bigg[\sum_{u,v\in\mathcal{V}_t,\,u\neq v} f(X_t^u)V(X_t^u)f(X_t^v)V(X_t^v)\bigg]$$
$$= \mu(V)^{-1}\int_0^t \mathbb{E}\bigg[J_2(VP_{t-s}f, VP_{t-s}f)(Y_s)\,\frac{r(Y_s)}{V(Y_s)}\bigg]\mathrm{e}^{-\lambda_0 s}\,\mathrm{d}s.$$

From (H2), we obtain $\lim_{t\to+\infty} A_t = 0$. Since $\pi(f)=0$, from (H1), we obtain $\lim_{t\to\infty} P_t f = 0$. Then, by (H3) and Lebesgue's theorem, we obtain, for all $s \geq 0$ and $x \in E$,

$$\lim_{t\to+\infty} J_2(VP_{t-s}f, VP_{t-s}f)(x) = 0.$$

And again by (H3) and Lebesgue's theorem, we obtain $\lim_{t\to+\infty} B_t = 0$. Now, if $\pi(f) \neq 0$ then we have

$$\mathbf{Z}_t(fV)\mathrm{e}^{-\lambda_0 t}\mu(V)^{-1} - W\pi(f)$$
$$= \mathbf{Z}_t((f-\pi(f))V)\mathrm{e}^{-\lambda_0 t}\mu(V)^{-1} + \pi(f)(\mathbf{Z}_t(V)\mathrm{e}^{-\lambda_0 t}\mu(V)^{-1} - W).$$

Then, thanks to the first part of the proof, the first term of the sum, in the right-hand side, converges to 0 in L^2. Moreover, the second term converges to 0 in probability thanks to Lemma 3.1. □

Proof of Theorem 1.1. If $f = g/V$ then it is a continuous and bounded function. If $h \equiv 1$ then all assumptions of the previous theorem hold. In particular, assumption (H3) of the previous theorem is exactly assumption (v) of Theorem 1.1. We then have the first convergence. Now if V is lower bounded, we can use the same argument with $g = 1$ and $f = 1/V$ which is also a continuous and bounded function. □

4. Macroscopic approximation

Let $(\mathcal{M}(E), d_v)$ (respectively $(\mathcal{M}(E), d_w)$) be the set of finite measures embedded with the vague (respectively weak) topology. These topologies are defined as follow:

$$\lim_{n\to+\infty} d_v(X_n, X_\infty) = 0 \iff \text{for all } f \in C_0, \quad \lim_{n\to+\infty} X_n(f) = X_\infty(f),$$
$$\lim_{n\to+\infty} d_w(X_n, X_\infty) = 0 \iff \text{for all } f \in C_b, \quad \lim_{n\to+\infty} X_n(f) = X_\infty(f),$$

where $(X_n)_{\geq 1}$ is a sequence on $\mathcal{M}(E)$ and $X_\infty \in \mathcal{M}(E)$. Here, C_0 is the set of continuous functions which vanish at ∞, and C_b is the set of continuous and bounded functions.

Note that vague convergence can also be defined as the weak* convergence with C_c^∞ test functions; see [28, Chapter 4] but we use the latter definition (used in [19, Section 7.3] for instance). These two definitions are generally not equivalent but, under the additional condition $\lim_{n\to\infty} X_n(1) = X_\infty(1)$ or if the family $(X_n)_{n\geq 1}$ is tight, they are.

Let $\mathbb{D}([0,T], E)$ and $C([0,T], E)$ be, respectively, the sets of càdlàg functions embedded with the Skorokhod topology and continuous functions embedded with the uniform topology [7].

4.1. Proof of Theorem 1.2

Let $(\mathbf{Z}^{(n)})_{n\geq 1}$ be a sequence of random measure-valued evolving as $\mathbf{Z}$ with starting distributions depending on n. In this section we consider the following scaling: $\boldsymbol{X}^{(n)} = (1/n)\mathbf{Z}^{(n)}$, and

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

we describe the behaviour of this scaled process for n tending to ∞. To understand the behaviour of our model in a large population, we can consider that it starts from a deterministic probability measure $\boldsymbol{X}_0$, and approach it by the interesting sequence defined by $\boldsymbol{X}_0^{(n)} = (1/n)\sum_{k=0}^n \delta_{Y_k}$, where $(Y_k)_{k\geq 1}$ is a sequence of independent and identically distributed (i.i.d.) random variable, distributed according to $\boldsymbol{X}_0$. In other words, we set $\boldsymbol{Z}_0^{(n)} = \sum_{k=0}^n \delta_{Y_k}$. The sequence $\boldsymbol{X}^{(n)}$ converges. Indeed, by the branching property, we have $\boldsymbol{Z}^{(n)} \overset{\mathrm{D}}{=} \sum_{k=0}^n \boldsymbol{Z}^{Y_k}$, where $\boldsymbol{Z}_t^{Y_k}$ are i.i.d., distributed as $\boldsymbol{Z}$, and starting from $\boldsymbol{Z}_0^{Y_k} = \delta_{Y_k}$. Henceforth, if f is a continuous and bounded function then from the classical law of large number, for all $t \geq 0$,

$$\lim_{n\to\infty} \boldsymbol{X}_t^{(n)}(f) = \mathbb{E}[\boldsymbol{Z}_t^{Y_1}(f)] \quad \text{almost surely.}$$

So, by corollary 2.1, and under a uniqueness assumption, it implies that $\boldsymbol{X}^{(n)}$ (point-wisely) converges to the solution $(\mu_t)_{t\geq 0}$ of the following integro-differential equation:

$$\begin{aligned}\mu_t(f) = \mu_0(f) + \int_0^t \mu_s(Gf) \\ + \int_E r(x)\bigg(\bigg(\sum_{k\geq 0} p_k(x)\int_0^1 \sum_{j=1}^k f(F_j^{(k)}(x,\theta))\,\mathrm{d}\theta\bigg) - f(x)\bigg)\mu_s(\mathrm{d}x)\,\mathrm{d}s.\end{aligned} \tag{4.1}$$

Theorem 1.2 gives a stronger convergence.

Lemma 4.1. (Semimartingale decomposition.) *If Assumption 2.1 holds then for all* $f \in C_c^2(E)$ *and* $t \geq 0$,

$$\boldsymbol{X}_t^{(n)}(f) = \boldsymbol{X}_0^{(n)}(f) + \boldsymbol{M}_t^{(n)}(f) + \boldsymbol{V}_t^{(n)}(f),$$

where $\boldsymbol{V}_t^{(n)}(f)$ *is equal to*

$$\int_0^t \int_E \bigg(Gf(x) + r(x)\bigg(\bigg(\int_0^1 \sum_{k\in\mathbb{N}}\sum_{j=1}^k f(F_j^{(k)}(x,\theta))p_k(x)\,\mathrm{d}\theta\bigg) - f(x)\bigg)\bigg)\boldsymbol{X}_s^{(n)}(\mathrm{d}x)\,\mathrm{d}s,$$

and $\boldsymbol{M}_t^{(n)}(f)$ *is a square-integrable and càdlàg martingale. Its bracket is defined by*

$$\begin{aligned}\langle \boldsymbol{M}^{(n)}(f)\rangle_t = \frac{1}{n}\int_0^t 2\boldsymbol{X}_s^{(n)}(Gf^2) - 2\boldsymbol{X}_s^{(n)}(f\times Gf) \\ + \int_E r(x)\int_0^1 \sum_{k\in\mathbb{N}}\bigg(\sum_{j=1}^k f(F_j^{(k)}(x,\theta)) - f(x)\bigg)^2 p_k(x)\,\mathrm{d}\theta\,\boldsymbol{X}_s^{(n)}(\mathrm{d}x)\,\mathrm{d}s.\end{aligned}$$

Proof. This is a direct consequence of Lemma 2.2. Indeed, if $\mathbb{L}^{(n)}$ is the generator of $\boldsymbol{X}^{(n)}$ then it satisfies

$$\begin{aligned}\mathbb{L}^{(n)}F_f(\mu) &= \restriction \partial_t \mathbb{E}[F_f(\boldsymbol{X}^{(n)}) \mid \boldsymbol{X}_0^{(n)} = \mu]t \\ &= 0 \\ &= \restriction \partial_t \mathbb{E}[F_{f/n}(\boldsymbol{Z}^{(n)}) \mid \boldsymbol{Z}_0^{(n)} = n\mu]t \\ &= 0 \\ &= \mathbb{L}F_{f/n}(n\mu),\end{aligned}$$

where $F_f(\mu) = F(\mu(f))$, F, f are two test functions, and $\mathbb{L}$ is the generator of $\boldsymbol{Z}$. □

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

Remark 4.1. (*Nonexplosion.*) Let us recall that, by Lemma 2.1, if the assumptions of Theorem 1.2 hold then Assumption 2.1 holds. In particular, there is no explosion.

Let us denote by $\mathcal{L}(U)$ the law of any random variable U.

Lemma 4.2. *Under the assumptions of Theorem 1.2, the sequence $(\mathcal{L}(\boldsymbol{X}^{(n)}))_{n\geq 1}$ is uniformly tight in the space of probability measures on $\mathbb{D}([0,T], (\mathcal{M}(E), d_v))$.*

Proof. We follow the approach of [20]. According to [43], it is enough to show that, for any continuous bounded function f, the sequence of laws of $\boldsymbol{X}^{(n)}(f)$ is tight in $\mathbb{D}([0,T],\mathbb{R})$. To prove it, we will use the Aldous–Rebolledo criterion. Let C_c^∞ be the set of functions of class C^∞ with finite support, we set $S = C_c^\infty \cup \{\mathbf{1}\}$, where $\mathbf{1}$ is the mapping $x \mapsto 1$. We have to prove that, for any function $f \in S$, we have

(i) for all $t \geq 0$, $(\boldsymbol{X}_t^{(n)}(f))_{n\geq 0}$ is tight;

(ii) for all $n \in \mathbb{N}$, and $\varepsilon, \eta > 0$, there exists $\delta > 0$ such that, for each stopping time S_n bounded by T,

$$\limsup_{n\to+\infty} \sup_{0\leq u\leq\delta} \mathbb{P}(|\boldsymbol{V}_{S_n+u}^{(n)}(f) - \boldsymbol{V}_{S_n}^{(n)}(f)| \geq \eta) \leq \varepsilon,$$

$$\limsup_{n\to+\infty} \sup_{0\leq u\leq\delta} \mathbb{P}(|\langle \boldsymbol{M}^{(n)}(f)\rangle_{S_n+u} - \langle \boldsymbol{M}^{(n)}(f)\rangle_{S_n}| \geq \eta) \leq \varepsilon.$$

The first point is the tightness of the family of time marginals $(\boldsymbol{X}_t^{(n)}(f))_{n\geq 1}$ and the second point, called the Aldous condition, gives a 'stochastic continuity'. It looks like the Arzelà–Ascoli theorem. Using Lemma 2.1, there exists $C > 0$ such that

$$\mathbb{P}(|\boldsymbol{X}_t^{(n)}(f)| > k) \leq \frac{\|f\|_\infty \mathbb{E}[\boldsymbol{X}_t^{(n)}(\mathbf{1})]}{k} \leq \frac{\|f\|_\infty C\mathbb{E}[\boldsymbol{X}_0^{(n)}(\mathbf{1})]}{k},$$

which tends to 0 as k tends to ∞. This proves the first point. Note that here and in all the proofs, constants may depend on T. Let $\delta > 0$, we have, for all stopping times $S_n \leq T_n \leq (S_n+\delta) \leq T$, that there exist $C', C_f > 0$ such that

$$\begin{aligned}
&\mathbb{E}[|\boldsymbol{V}_{T_n}^{(n)}(f) - \boldsymbol{V}_{S_n}^{(n)}(f)|] \\
&\quad= \mathbb{E}\bigg[\bigg|\int_{S_n}^{T_n} \boldsymbol{X}_s^{(n)}(Gf) \\
&\qquad\quad + \int_E \bigg(r(x)\bigg(\int_0^1 \sum_{k\in\mathbb{N}}\sum_{j=1}^k f(F_j^{(k)}(x,\theta) p_k(x)\,\mathrm{d}\theta)\bigg) - f(x)\bigg)\boldsymbol{X}_s^{(n)}(\mathrm{d}x)\,\mathrm{d}s\bigg|\bigg] \\
&\quad\leq C'[\|Gf\|_\infty + \|f\|_\infty] \times \mathbb{E}[|T_n - S_n|] \\
&\quad\leq C_f\delta.
\end{aligned}$$

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

On the other hand, there exists $C'_f > 0$ such that

$$\begin{aligned}\mathbb{E}[|\langle \boldsymbol{M}^{(n)}(f)\rangle_{T_n} - \langle \boldsymbol{M}^{(n)}(f)\rangle_{S_n}|] &= \frac{1}{n}\mathbb{E}\bigg[\bigg|\int_{S_n}^{T_n} \boldsymbol{X}_s^{(n)}(Gf^2) - 2\boldsymbol{X}_s^{(n)}(fGf) \\ &\qquad + \int_E r(x)\int_0^1 \sum_{k\in\mathbb{N}}\sum_{j=1}^{k}(f(F_j^{(k)}(x,\theta)) - f(x))^2 p_k(x)\,\mathrm{d}\theta\, \boldsymbol{X}_s^{(n)}(\mathrm{d}x)\,\mathrm{d}s\bigg|\bigg] \\ &\le \frac{C'_f\delta}{n}.\end{aligned}$$

Then, for a sufficiently small δ, the second point is verified and we conclude that $(\boldsymbol{X}^{(n)})_{n\ge1}$ is uniformly tight in $\mathbb{D}([0,T],(\mathcal{M}(E),d_v))$. □

Proof of Theorem 1.2. Let us denote by $\boldsymbol{X}$ a limit process of $(\boldsymbol{X}^{(n)})_{n\ge1}$; namely, there exists an increasing sequence $(u_n)_{n\ge1}$, on $\mathbb{N}^*$, such that $(\boldsymbol{X}^{(u_n)})_{n\ge1}$ converges to $\boldsymbol{X}$. It is almost surely continuous in $(\mathcal{M}(E), v)$ since

$$\sup_{t\ge0}\sup_{\|f\|_\infty\le1}|\boldsymbol{X}_{t-}^{(n)}(f) - \boldsymbol{X}_t^{(n)}(f)| \le \frac{\bar{k}}{n}. \tag{4.2}$$

In the case where E is compact, the vague and weak topologies coincide. By the Cauchy–Schwarz equation and Doob's inequality, there exists $C > 0$ such that

$$\sup_f \mathbb{E}\Big[\sup_{t\le T}|\boldsymbol{M}_t^{(n)}(f)|\Big] \le 2\sup_f \mathbb{E}[\langle \boldsymbol{M}^{(n)}(f)\rangle_T]^{1/2} \le \frac{C}{\sqrt{n}},$$

where the supremum is taken over all the function $f\in C_c^2(E)$ such that $\|f\|_\infty\le1$. Hence,

$$\lim_{n\to+\infty}\sup_f \mathbb{E}\Big[\sup_{t\le T}|\boldsymbol{M}_t^{(n)}(f)|\Big] = 0.$$

However, since

$$\begin{aligned}\boldsymbol{M}_t^{(n)}(f) &= \boldsymbol{X}_t^{(n)}(f) - \boldsymbol{X}_0^{(n)}(f) \\ &\quad - \int_0^t\int_E\bigg(Gf(x) + r(x)\bigg(\bigg(\int_0^1\sum_{k\in\mathbb{N}}\sum_{j=1}^{k} f(F_j^{(k)}(x,\theta))p_k(x)\,\mathrm{d}\theta\bigg) \\ &\qquad - f(x)\bigg)\bigg)\boldsymbol{X}_s^{(n)}(\mathrm{d}x)\,\mathrm{d}s,\end{aligned}$$

we have

$$\begin{aligned}0 &= \boldsymbol{X}_t(f) - \boldsymbol{X}_0(f) - \int_0^t \boldsymbol{X}_s(Gf) \\ &\quad + \int_E r(x)\bigg(\bigg(\sum_{j=1}^{k} f(F_j^{(K)}(x,\theta))p_k(x)\,\mathrm{d}\theta\bigg) - f(x)\bigg)\boldsymbol{X}_s(\mathrm{d}x)\,\mathrm{d}s.\end{aligned}$$

Since this equation has a unique solution, it ends the proof when E is compact. This approach fails in the noncompact case. Nevertheless, we can use the Méléard–Roelly criterion [34]. We have to prove that $\boldsymbol{X}$ is in $C([0,T],(\mathcal{M}(E),w))$ and $\boldsymbol{X}^{(n)}(\mathbf{1})$ converges to $\boldsymbol{X}(\mathbf{1})$. By (4.2), $\boldsymbol{X}$ is continuous. To prove that $\boldsymbol{X}^{(n)}(\mathbf{1})$ converges to $\boldsymbol{X}(\mathbf{1})$, we use the following lemmas. □

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

Lemma 4.3. (Approximation of indicator functions.) *Under the assumptions of Theorem 1.2, for each $k \in \mathbb{Z}$, there exists $\psi_k \in C^2(E)$ such that, for all $x \in E$,*

$$1_{[k;+\infty)}(x) \leq \psi_k(x) \leq 1_{[k-1;+\infty)}(x) \quad \textit{and there exists} \quad C > 0, \quad G\psi_k \leq C\psi_{k-1}.$$

Lemma 4.4. (Commutation of limits.) *Under the assumptions of Theorem 1.2,*

$$\lim_{k\to+\infty} \limsup_{n\to+\infty} \mathbb{E}\Big[\sup_{t\leq T} X_t^{(n)}(\psi_k)\Big] = 0,$$

where $(\psi_k)_{k\geq 0}$ are defined as in the previous lemma.

Proof. The proofs are very similar to [35] and we omit the details here. As a consequence, the same computation as [35] gives us the convergence in $\mathbb{D}([0,T], (\mathcal{M}(E), w))$. Thus, each subsequence converges to (4.1). The end of the proof follows with the same argument of the compact case. □

We can give another argument, which does not use the Méléard–Roelly criterion [34]. By (4.2), $\boldsymbol{X}$ is continuous from $[0, T]$ to $(\mathcal{M}(E), d_w)$, let $\boldsymbol{G}$ be a Lipschitz function on $C([0, T], (\mathcal{M}(E), d_w))$ with a Lipschitz constant equal to 1. We obtain

$$\begin{aligned}
|\mathbb{E}[\boldsymbol{G}(\boldsymbol{X}^{(u_n)})] - \boldsymbol{G}(\boldsymbol{X})| &\leq \mathbb{E}\Big[\sup_{t\in[0,T]} d_w(\boldsymbol{X}_t^{(u_n)}, \boldsymbol{X}_t)\Big] \\
&\leq \mathbb{E}\Big[\sup_{t\in[0,T]} d_w(\boldsymbol{X}_t^{(u_n)}, \boldsymbol{X}_t^{(u_n)}(\cdot \times (1-\psi_k)))\Big] \\
&\quad + \mathbb{E}\Big[\sup_{t\in[0,T]} d_w(\boldsymbol{X}_t^{(u_n)}(\cdot \times (1-\psi_k)), \boldsymbol{X}_t(\cdot \times (1-\psi_k)))\Big] \\
&\quad + \sup_{t\in[0,T]} d_w(\boldsymbol{X}_t(\cdot \times (1-\psi_k)), \boldsymbol{X}_t).
\end{aligned}$$

According to Lemma 4.4, we have

$$\lim_{k\to+\infty} \limsup_{n\to+\infty} \mathbb{E}\Big[\sup_{t\in[0,T]} d_w(\boldsymbol{X}_t^{(u_n)}, \boldsymbol{X}_t^{(u_n)}(\cdot \times (1-\psi_k)))\Big] = 0$$

and

$$\lim_{k\to+\infty} \sup_{t\in[0,T]} d_w(\boldsymbol{X}_t(\cdot \times (1-\psi_k)), \boldsymbol{X}_t) = 0.$$

Then, we have

$$d_w(\boldsymbol{X}_t^{(u_n)}(\cdot \times (1-\psi_k)), \boldsymbol{X}_t(\cdot \times (1-\psi_k))) = d_v(\boldsymbol{X}_t^{(u_n)}(\cdot \times (1-\psi_k)), \boldsymbol{X}_t(\cdot \times (1-\psi_k))).$$

Thus,

$$\lim_{k\to+\infty} \limsup_{n\to+\infty} \mathbb{E}\Big[\sup_{t\in[0,T]} d_w(\boldsymbol{X}_t^{(u_n)}(\cdot \times (1-\psi_k)), \boldsymbol{X}_t(\cdot \times (1-\psi_k)))\Big] = 0,$$

by the continuity of $\nu \mapsto \nu(1-\psi_k)$ in $\mathbb{D}(\mathcal{M}(E), d_v)$. Finally, $\lim_{n\to+\infty}\mathbb{E}[\boldsymbol{G}(\boldsymbol{X}^{(u_n)})] = \mathbb{E}[\boldsymbol{G}(\boldsymbol{X})] = \boldsymbol{G}(\boldsymbol{X})$, which completes the proof. □

Note that in the previous proof, we used Lipschitz functions instead of bounded continuous functions in order to prove weak convergence. This can be justified from (a slight modification of) the classical Portmanteau theorem. Indeed, see [7, Theorem 2.1] and its proof.

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

5. Main example: a size-structured population model

Let us introduce a model which represents the cell mitosis. It is described as follows: the underlying process X is deterministic and linear and when a cell dies, it divides in two parts. Formally and with our notation, we have

$$E = (0, +\infty), \qquad p_2 = 1, \quad \text{and} \quad Gf = f' \tag{5.1}$$

for every $f \in C^1(E)$, and, for all $x \in E$, for all $\theta \in [0, 1]$,

$$F_1^{(2)}(x, \theta) = F^{-1}(\theta)x \quad \text{and} \quad F_2^{(2)}(x, \theta) = (1 - F^{-1}(\theta))x, \tag{5.2}$$

where F is the cumulative distribution function of a random variable on $[0, 1]$. It satisfies $F(x) = 1 - F(1 - x)$. In this case, when r is smooth enough, one cell lineage is generated by, for all $x \geq 0$,

$$Lf = f'(x) + r(x)[\mathbb{E}[f(Hx)] - f(x)]$$

for every $f \in C^1(E)$, where H is distributed according to F. This (one cell lineage) process is sometimes called the TCP process in computer science [8], [22], [32], [38]. First, we prove the nonexplosion even if r is not bounded.

Lemma 5.1. (Nonexplosion.) *Let $p \geq 1$. If (5.1) and (5.2) hold, r is continuous and, for all $x \in \mathbb{R}_+^*$, $r(x) \leq C_0(1 + x^p)$, and $\mathbb{E}[\mathbf{Z}_0(1 + x^p)] < +\infty$, then our process is nonexplosive. Moreover,*

$$\mathbb{E}\Big[\sup_{s\in[0,T]} \mathbf{Z}_s(1 + x^p)\Big] \leq \mathbb{E}[\mathbf{Z}_0(1 + x^p)]\mathrm{e}^{C_p T},$$

where C_p is constant and $T > 0$.

Proof. Recall that, for every $f \in C_c^2(E)$, we have

$$\begin{aligned}
\mathbf{Z}_t(f) = \mathbf{Z}_0(f) &+ \int_0^t \int_E f'(x)\mathbf{Z}_s(\mathrm{d}x)\,\mathrm{d}s \\
&+ \int_0^t \int_{\mathcal{U}\times\mathbb{R}_+\times[0,1]} 1_{\{u\in V_{s-},\, l\leq r(X_{s-}^u)\}}(f(\theta X_{s-}^u) + f((1-\theta)X_{s-}^u) \\
&\qquad\qquad - f(X_{s-}^u))\rho(\mathrm{d}s, \mathrm{d}u, \mathrm{d}l, \mathrm{d}\theta).
\end{aligned}$$

Using the same argument as in [20, Theorem 3.1], we introduce $\tau_n = \inf\{t \geq 0 \mid \mathbf{Z}_t(1+x^p) > n\}$; and we have

$$\begin{aligned}
&\sup_{u\in[0,t\wedge\tau_n]} \mathbf{Z}_u(1 + x^p) \\
&\quad\leq \mathbf{Z}_0(1 + x^p) + \int_0^{t\wedge\tau_n} \mathbf{Z}_s(px^{p-1})\,\mathrm{d}s \\
&\qquad + \int_0^{t\wedge\tau_n} \int_{\mathcal{U}\times\mathbb{R}_+\times[0,1]} 1_{\{u\in V_{s-},\, l\leq r(X_{s-}^u)\}} \\
&\qquad\qquad \times (1 + (\theta^p + (1-\theta)^p - 1)(X_{s-}^u)^p)\rho(\mathrm{d}s, \mathrm{d}u, \mathrm{d}l, \mathrm{d}\theta) \\
&\quad\leq \mathbf{Z}_0(1 + x^p) + \int_0^{t\wedge\tau_n} p \times \sup_{u\in[0,s\wedge\tau_n]} \mathbf{Z}_u(1 + x^p)\,\mathrm{d}s \\
&\qquad + \int_0^t \int_{\mathcal{U}\times\mathbb{R}_+\times[0,1]} 1_{\{u\in V_{s-},\, l\leq r(X_{s-}^u)\}}\rho(\mathrm{d}s, \mathrm{d}u, \mathrm{d}l, \mathrm{d}\theta),
\end{aligned}$$

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

since $(\theta^p + (1-\theta)^p - 1) \le 0$. Thus, there exist $C > 0$ such that

$$\mathbb{E}\Big[\sup_{u\in[0,t\wedge\tau_n]} \mathbf{Z}_u(1+x^p)\Big] \le \mathbb{E}[\mathbf{Z}_0(1+x^p)] + \int_0^t C\mathbb{E}\Big[\sup_{u\in[0,s\wedge\tau_n]} \mathbf{Z}_u(1+x^p)\Big]\,\mathrm{d}s.$$

Finally, the Gronwall lemma implies the existence of C_p such that

$$\mathbb{E}\Big[\sup_{s\in[0,t\wedge\tau_n]} \mathbf{Z}_s(1+x^p)\Big] \le \mathbb{E}[\mathbf{Z}_0(1+x^p)]\mathrm{e}^{C_p t}.$$

We deduce that τ_n tends almost surely to ∞ and that there is nonexplosion. □

5.1. Equal mitosis: long-time behaviour

In this subsection we establish the long-time behaviour of $\mathbf{Z}$. We assume that, for all $x \ge 0$, for all $\theta \in [0, 1]$,

$$F_1^{(2)}(x,\theta) = F_2^{(2)}(x,\theta) = \tfrac{1}{2}x. \tag{5.3}$$

That is, the cells divide in two equal parts. In short, we have, for all $x \ge 0$,

$$\mathcal{G}f(x) = f'(x) + r(x)(2f(\tfrac{1}{2}x) - f(x)) \quad \text{for every } f \in C^1(E).$$

In order to give a many-to-one formula, we recall a theorem of [40].

Theorem 5.1. (Sufficient condition for the existence of eigenelements.) *If (5.1) and (5.3) hold, there exist $\underline{r}, \bar{r} > 0$ such that, for all $x \in E$, $\underline{r} \le r(x) \le \bar{r}$, r is continuous and $r(x)$ is a constant equal to some r_∞ for x large enough, then there exist $V \in C^1(\mathbb{R}_+)$ and $\lambda_0 > 0$ such that $\mathcal{G}V = \lambda_0 V$ and, for all $x \ge 0$,*

$$c(1+x^k) \le V(x) \le C(1+x^k),$$

where C, c are two constants and $2^k = 2r_\infty/(\lambda_0 + r_\infty)$.

So, we obtain a many-to-one formula with an auxiliary process generated by A, defined for every $f \in C^1(E)$ and $x \in E$, by

$$Af(x) = f'(x) + r(x)\frac{2V(x/2)}{V(x)}\bigg(f\Big(\frac{x}{2}\Big) - f(x)\bigg).$$

Our main result gives the two following limit theorems.

Corollary 5.1. (Convergence of the empirical measure for a mitosis model.) *Under the assumptions of Theorem 5.1, there exists a probability measure π_1 such that, for any continuous and bounded function g, we have*

$$\lim_{t\to+\infty} \frac{1}{N_t} \sum_{u\in\mathcal{V}_t} g(X_t^u) = \int g\,\mathrm{d}\pi_1 \quad \textit{in probability}.$$

In particular, for a constant rate r, π_1 has Lebesgue density

$$x \mapsto \frac{2r}{\prod_{n=1}^{+\infty}(1-2^{-n})} \sum_{n=0}^{+\infty}\bigg(\prod_{k=1}^{n} \frac{2}{1-2^k}\bigg)\mathrm{e}^{-2^{n+1}rx}. \tag{5.4}$$

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

The explicit formula in (5.4) is not new (see [39], [40]), but here, the empirical measure converges in probability, while, in the mentioned papers, the mean measure or the macroscopic process converges (see Theorem 1.2).

Proof of Corollary 5.1. By Theorem 5.1, the mapping $x \mapsto V(x/2)/V(x)$ is upper and lower bounded. Thus, the auxiliary process is ergodic and admits a unique invariant law π, as can be checked using a suitable Foster–Lyapunov function [36, Theorem 6.1]. Indeed, if $\psi_p\colon x \mapsto 1 + x^p$ then, for $p \geq 1$,

$$A\psi_p(x) \leq px^{p-1} - x^p \frac{rc}{2^p C} \leq K_p - \alpha_p \psi(x) \quad \text{for some } K_p, \alpha_p > 0.$$

See also [21, Theorem 1] which gives several details. As a second consequence, all the moments of Y are finite. Assumption (1.1) of Theorem 1.1 then holds because the left-hand side is bounded. Indeed r is bounded, V is bounded by polynomials, and Y has finite moments. Now, applying Theorem 1.1, we have the convergence on the set $\{W \neq 0\}$, where $W = \lim_{t\to\infty} \boldsymbol{Z}_t(V)\mathrm{e}^{-\lambda_0}$ almost surely. It remains to prove that $W > 0$ almost surely. We begin by proving that the martingale $(\boldsymbol{Z}_t(V)\mathrm{e}^{-\lambda_0 t})_{t\geq 0}$ converges to W in L^1. Let $p > 1$, by the Burkholder–Davis–Gundy inequality [10, Theorem 92, p. 304], there exists $C > 0$ such that

$$\begin{aligned}\mathbb{E}[|\boldsymbol{Z}_t(V)\mathrm{e}^{-\lambda_0 t} - \boldsymbol{Z}_0(V)|^p] &\leq C\mathbb{E}\bigg[\sum_{t\geq 0} |\boldsymbol{Z}_{t+}(V)\mathrm{e}^{-\lambda_0 t} - \boldsymbol{Z}_{t-}(V)\mathrm{e}^{-\lambda_0 t}|^p\bigg] \\ &\leq C\mathbb{E}\bigg[\sum_{u\in\mathcal{T}} \mathrm{e}^{-\lambda_0 p\beta(u)} \bigg|2V\bigg(\frac{X^u_{\beta(u)-}}{2}\bigg) - V(X^u_{\beta(u)-})\bigg|^p\bigg].\end{aligned}$$

Now by Lemma 3.3, we have

$$\begin{aligned}&\mathbb{E}[|\boldsymbol{Z}_t(V)\mathrm{e}^{-\lambda_0 t} - \boldsymbol{Z}_0(V)|^p] \\ &\quad\leq C\int_0^\infty \mathbb{E}[\boldsymbol{Z}_0(V)]\mathrm{e}^{-(p-1)\lambda_0 s}\mathbb{E}\bigg[r(Y_s)\frac{|2V(Y_s/2) - V(Y_s)|^p}{V(Y_s)}\bigg]\,\mathrm{d}s.\end{aligned}$$

Finally, using the fact that r is bounded, the conclusion of Theorem 5.1, and that all moments of Y are bounded, it holds that the right-hand side of the previous equation is bounded. As a consequence, the martingale $(\boldsymbol{Z}_t(V)\mathrm{e}^{-\lambda_0 t})_{t\geq 0}$ has a bounded second moment and converges to W in L^1. We deduce that $\mathbb{E}[W] > 0$ and $\varrho = \mathbb{P}(W = 0) < 1$. But, conditioning to the time of the first division and taking the limit $t \to +\infty$ shows that $\varrho^2 = \varrho$. Finally, $\varrho = 0$ and this completes the proof. The measure π_1 is then $1/(V(x)\int 1/V\,\mathrm{d}\pi)\pi(\mathrm{d}x)$. In particular, when r is constant, V is constant, and π_1 corresponds to the invariant distribution of a TCP process with rate $2r$; the explicit formula is then an application of [38, Theorem 1] and [22, Proposition 5]. □

We can see that the assumptions of Theorem 5.1 are strong and not necessary.

Corollary 5.2. (Convergence of the empirical measure when r is affine.) *If (5.1) and (5.3) hold and for all $x \geq 0$, $r(x) = u_1 x + v_1$, where $u_1 > 0$ and $v_1 \geq 0$, then there exists a measure π such that*

$$\lim_{t\to+\infty} \frac{1}{N_t} \sum_{u\in\mathcal{V}_t} g(X^u_t) = \int g\,\mathrm{d}\pi.$$

The convergence holds in probability and for any continuous function g on E such that, for all $x \in E$, $|g(x)| \leq C(1 + x)$.

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

Proof. If $r(x) = u_1 x + v_1$ then $V(x) = x(\sqrt{v_1^2 + 4u_1} - v_1)/2 + 1$ is an eigenvector and $2u_1/(\sqrt{v_1^2 + 4u_1} - v_1)$ is its corresponding eigenvalue. Henceforth, the proof is very similar to the previous corollary. □

Remark 5.1. (*Malthus parameter.*) Under the assumptions of the previous corollary, we also deduce that

$$\lim_{t\to+\infty} N_t \mathrm{e}^{-\lambda_0 t} = W \int_E \frac{1}{V}\,\mathrm{d}\pi,$$

where $\lambda_0 = 2u_1/(\sqrt{v_1^2 + 4u_1} - v_1)$ is the Malthus parameter (see Remark 3.3).

Remark 5.2. (*Estimation of r.*) We can find some estimates of the division rate in the literature. An inverse problem was developed and applied with experimental data in [13] (see also [30]). More recently, in [14] the author gave a nonparametric estimation of the division rate.

5.2. Homogeneous case: moment and rate of convergence

When r is constant, the process is easier to study since the auxiliary process has already been studied [8], [32], [38]. Here, we give the moments and a first approach to estimate the rate of convergence.

Lemma 5.2. (Moments of the empirical measure.) *If (5.1) and (5.2) hold and r is constant, then, for all $m \in \mathbb{N}$, and $t \geq 0$, we have*

$$\begin{aligned}\mathbb{E}[\mathbf{Z}_t(x^m)] &= \mathbb{E}\bigg[\sum_{u\in\mathcal{V}_t}(X_t^u)^m\bigg]\\ &= \int_0^{+\infty} \mathrm{e}^{rt}\bigg[\frac{m!}{\prod_{i=1}^m \theta_i} + m!\sum_{i=1}^m\bigg(\sum_{k=0}^i \frac{x^k}{k!}\prod_{j=k,\, j\neq i}^m \frac{1}{\theta_j - \theta_i}\bigg)\mathrm{e}^{-\theta_i t}\bigg] z_0(\mathrm{d}x),\end{aligned}$$

where $\theta_i = 2r(1 - 2^{-i})$.

Proof. Since r is constant, we have $\mathcal{G}\mathbf{1} = r\mathbf{1}$, where $\mathbf{1}$ is the constant mapping, which is equal to 1. From Lemma 3.2, we have

$$\frac{1}{\mathbb{E}[N_t]}\mathbb{E}\bigg[\sum_{u\in\mathcal{V}_t} f(X_t^u)\bigg] = \mathbb{E}[f(Y_t)]$$

for every continuous and bounded function f, where Y is generated by A, defined, for every $f \in C^1(E)$ and $x \in E$, by

$$Af(x) = f'(x) + 2r\big(f\big(\tfrac{1}{2}x\big) - f(x)\big).$$

Finally, we complete the proof using [32, Theorem 4]. □

Now, let us talk about the rate of convergence. To estimate the distance between two random measures, we will use the Wasserstein distance [9], [45].

Definition 5.1. (*Wasserstein distance.*) Let μ_1 and μ_2 two finite measures on a Polish space (F, d_F), the Wasserstein distance between μ_1 and μ_2 is defined by

$$W_{d_F}(\mu_1, \mu_2) = \inf \int_{F\times F} d_F(x_1, x_2)\Pi(\mathrm{d}x_1, \mathrm{d}x_2),$$

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

where the infimum runs over all the measures Π on $F \times F$ with marginals μ_1 and μ_2. In particular, if μ_1 and μ_2 are two probability measures, we have

$$W_{d_F}(\mu_1, \mu_2) = \inf \mathbb{E}[d_F(X_1, X_2)],$$

where the infimum runs over all two random variables X_1, X_2, which are distributed according to μ_1, μ_2.

So, if M_1, M_2 are two random measures then

$$W_d(\mathcal{L}(M_1), \mathcal{L}(M_2)) = \inf \mathbb{E}[d(M_1, M_2)],$$

where the infimum is taken over all the couples of random variables (M_1, M_2) such that $M_1 \sim \mathcal{L}(M_1)$ and $M_2 \sim \mathcal{L}(M_2)$, and d is a distance on the measures space. Here, we consider $d = W_{|\cdot|}$. It is the Wasserstein distance on $(E, |\cdot|)$.

Theorem 5.2. (Quantitative bounds.) *If (5.1) and (5.2) hold and r is constant, then we have, for all $t \geq 0$,*

$$W_{W_{|\cdot|}}\bigg(\mathcal{L}\bigg(\frac{\mathbf{Z}_t^x}{\mathbb{E}[N_t]}\bigg), \mathcal{L}\bigg(\frac{\mathbf{Z}_t^y}{\mathbb{E}[N_t]}\bigg)\bigg) \leq |x-y|\mathrm{e}^{-rt},$$
$$W_{W_{|\cdot|}}\bigg(\mathcal{L}\bigg(\frac{\mathbf{Z}_t^x}{N_t}\bigg), \mathcal{L}\bigg(\frac{\mathbf{Z}_t^y}{N_t}\bigg)\bigg) \leq |x-y|\frac{rt\mathrm{e}^{-rt}}{1-\mathrm{e}^{-rt}},$$

where $\mathbf{Z}^x$ and $\mathbf{Z}^y$ are distributed as $\mathbf{Z}$ and start from δ_x and δ_y.

This result does not give a bound for $W_{W_{|\cdot|}}(\mathcal{L}(\mathbf{Z}_t/\mathbb{E}[N_t]), \mathcal{L}(W\pi))$, or $W_{W_{|\cdot|}}(\mathcal{L}(\mathbf{Z}_t/N_t), \mathcal{L}(\pi))$, where π is the limit measure of 5.1.

Proof of Theorem 5.2. By homogeneity, we can see our branching measure $\mathbf{Z}$ as a process indexed by a Galton–Watson tree [5]. For our coupling, we take two processes indexed by the same tree. More precisely, as the branching time does not depend on the position, we can set the same times to our two processes. Let $\mathcal{T} = \bigcup_{n\in\mathbb{N}}\{1,2\}^n$ represent cells that have lived at a certain moment. Let $(\tau_u)_{u\in\mathcal{U}}$ be a family of i.i.d. exponential variables with mean $1/r$, which model the lifetimes. We build $\mathbf{Z}^x$ and $\mathbf{Z}^y$ by induction. First, for all $t \in [0, \tau_\varnothing)$, $X_t^\varnothing = x+t$ and $Y_t^\varnothing = y+t$. We set $\alpha(\varnothing) = 0$. Then, for all $u \in \mathcal{T}$ and $k \in \{1,2\}$, we set $\alpha(uk) = \alpha(u) + \tau_u$, for all $t \in [\alpha(uk), \alpha(uk) + \tau_{uk})$,

$$X_t^{uk} = \tfrac{1}{2}X^u_{\alpha(uk)-} + t - \alpha(uk) \quad \text{and} \quad Y_t^{uk} = \tfrac{1}{2}Y^u_{\alpha(uk)-} + t - \alpha(uk).$$

Finally, we have $\mathcal{V}_t = \{u \in \mathcal{T} \mid \alpha(u) \leq t < \alpha(u) + \tau_u\}$, $\mathbf{Z}_t^x = \sum_{u\in\mathcal{V}_t}\delta_{X_t^u}$, and $\mathbf{Z}_t^y = \sum_{u\in\mathcal{V}_t}\delta_{Y_t^u}$.

We observe that, for any cell u, the trajectories of X^u and Y^u are parallel (because they are linear). When a branching occurs, $\sum_{u\in\mathcal{V}_t}|X_t^u - Y_t^u|$ is constant. Hence, we easily deduce that

$$\sum_{u\in\mathcal{V}_t}|X_t^u - Y_t^u| = |x-y|.$$

Finally, we have, for all $t \geq 0$,

$$W_{|\cdot|}(\mathbf{Z}_t^x, \mathbf{Z}_t^y) \leq \sum_{u\in\mathcal{V}_t}|X_t^u - Y_t^u| \leq |x-y|.$$

Dividing by $\mathbb{E}[N_t] = \mathrm{e}^{rt}$, we obtain the first bound. For the second bound, a similar computation yields

$$W_{W_{|\cdot|}}\left(\mathcal{L}\left(\frac{\boldsymbol{Z}_t^x}{N_t}\right), \mathcal{L}\left(\frac{\boldsymbol{Z}_t^y}{N_t}\right)\right) \le \mathbb{E}\left[\frac{1}{N_t}\right]|x-y|.$$

The process $(N_t)_{t\ge 0}$ is know to be the Yule process. It is geometrically distributed with parameter e^{-rt} (see Lemma 5.3 below). The following equality completes the proofs:

$$\mathbb{E}\left[\frac{1}{N_t}\right] = \frac{rt\mathrm{e}^{-rt}}{1-\mathrm{e}^{-rt}}. \qquad \square$$

Recall the following well-known result whose the proof is given for sake of completeness.

Lemma 5.3. (Properties of the Yule process.) *Let $(M_t)_{t\ge 0}$ be the process which gives the number of individuals alive at time t in a branching process in which each individual lives for an exponential time of constant parameter r and gives birth at its death to two children. For every $t \ge 0$, M_t follows a geometric law with parameter* e^{-rt}.

Proof. We have $M_t = \inf\{n \ge 0 \mid S_n \le t\}$, where $S_n = \sum_{k=1}^n F_k$ and F_k denotes the time of the kth birth; F_k is an exponential variable with parameter rk. A straightforward recurrence shows that

$$S_n \overset{\mathrm{D}}{=} \max(E_1, \dots, E_n),$$

where $(E_i)_{i\ge 0}$ is a sequence of i.i.d. exponential random variable with parameter r and then

$$\mathbb{P}(N_t - 1 \ge n) = \mathbb{P}(\max(E_1, \dots, E_n) \le t) = (1-\mathrm{e}^{-rt})^n. \qquad \square$$

Remark 5.3. (*Generalisation of Theorem 5.2.*) In the proof of Theorem 5.2, we only need that, for all $n \in \mathbb{N}^*$, $\theta \in [0,1]$, $t \ge 0$, and $x, y \in E$,

$$\sum_{j=1}^n |F_j^{(k)}(X_t, \theta) - F_j^{(k)}(Y_t, \theta)| \le |x-y|,$$

where X and Y are generated by G and start respectively from x, y. For instance, we can consider that X is a continuous Lévy process and the division is a subcritical fragmentation; namely, $F_j^{(k)}(x, \Theta) = \Theta_j^k x$, where $(\Theta_j^k)_{j,k}$ is a family of random variables satisfying

$$\sum_{j=1}^k \Theta_j^k \le 1 \quad \text{and,} \quad \text{for all } j \in \{1, \dots, k\}, \qquad \Theta_j^k \in [0,1].$$

5.3. Asymmetric mitosis: macroscopic approximation

Now, we do not assume that size is divided by 2 at each division. We assume that $F_1^{(2)}(x,\theta) = F^{-1}(\theta)x$ and $F_2^{(2)}(x,\theta) = (1-F^{-1}(\theta))x$. We recall that $F(x) = 1 - F(1-x)$. In this case, (1.7) becomes

$$\partial_t n(t,x) + \partial_x n(t,x) + r(x)n(t,x) = 2\mathbb{E}\left[\frac{1}{\Theta} r\left(\frac{x}{\Theta}\right) n\left(t, \frac{x}{\Theta}\right)\right],$$

where $n(t,\cdot)$ is the density of $\boldsymbol{X}_t$. In particular, we deduce that the following partial differential equation has a weak solution:

$$\partial_t n(t,x) + \partial_x n(t,x) + r(x)n(t,x) = \int_x^{+\infty} b(x,y)n(t,y)\,\mathrm{d}y, \tag{5.5}$$

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

where b satisfies the following properties:

$$b(x, y) \geq 0, \qquad b(x, y) = 0 \quad \text{for } y < x, \tag{5.6}$$

$$\int_0^{+\infty} b(x, y)\,\mathrm{d}x = 2r(y), \tag{5.7}$$

$$\int_0^{+\infty} x b(x, y)\,\mathrm{d}x = yr(y), \tag{5.8}$$

$$b(x, y) = b(y - x, y). \tag{5.9}$$

Equation (5.5) was studied in [39]. Here,

$$b(x, y) = \frac{2}{y} r(y) g\left(\frac{x}{y}\right), \tag{5.10}$$

where g is the weak density of F. We easily prove the equivalence satisfying (5.10) and (5.6)–(5.9). Our aim in this section is to describe the limit of the fluctuation process. It is defined by, for all $t \in [0, T]$, for all $n \in \mathbb{N}^*$,

$$\eta_t^{(n)} = \sqrt{n}(\boldsymbol{X}_t^{(n)} - \boldsymbol{X}_t).$$

Theorem 5.3. (Central limit theorem for asymmetric size-structured population.) *Let $T > 0$. Assume that $\eta_0^{(n)}$ converges in distribution and that*

$$\mathbb{E}\left[\sup_{n\geq 1} \int_E (1 + x^2) \boldsymbol{X}_0^{(n)}(\mathrm{d}x)\right] < +\infty. \tag{5.11}$$

Then the sequence $(\eta^{(n)})_{n\geq 1}$ converges in $\mathbb{D}([0, T], C^{-2,0})$ to the unique solution of the evolution equation: for all $f \in C^{2,0}$,

$$\begin{aligned}\eta_t(f) &= \eta_0(f)\\ &\quad + \int_0^t \int_0^{+\infty} (f'(x) + r(x) \int_0^1 (f(qx) + f((1-q)x) - f(x)) F(\mathrm{d}q)) \eta_s(\mathrm{d}x)\,\mathrm{d}s\\ &\quad + \widetilde{\boldsymbol{M}}_t(f),\end{aligned} \tag{5.12}$$

where $\widetilde{\boldsymbol{M}}(f)$ is a martingale and a Gaussian process with bracket

$$\langle \widetilde{\boldsymbol{M}}(f)\rangle_t = \int_0^t \int_0^{+\infty} \left(2f'(x)f(x) + 2r(x)\int_0^1 (f(qx) - f(x))^2 F(\mathrm{d}q)\right) \boldsymbol{X}_s(\mathrm{d}x)\,\mathrm{d}s.$$

And $C^{2,0}$ is the set of C^2 functions, such that f, f', f'' vanish to 0 when x tends to ∞. Also $C^{-2,0}$ is its dual space.

From Lemma 4.1 we have, for all $t \geq 0$, $\eta_t^{(n)} = \eta_0^{(n)} + \widetilde{\boldsymbol{V}}_t^{(n)} + \widetilde{\boldsymbol{M}}_t^{(n)}$, where, for any $f \in C_c^2(E)$,

$$\widetilde{\boldsymbol{V}}_t^{(n)}(f) = \int_0^t \int_0^{+\infty} \left(f'(x) + r(x) \int_0^1 (f(qx) + f((1-q)x) - f(x)) F(\mathrm{d}q)\right) \eta_s^{(n)}(\mathrm{d}x)\,\mathrm{d}s,$$

and $\widetilde{\boldsymbol{M}}^{(n)}$ is a martingale with bracket

$$\langle \widetilde{\boldsymbol{M}}^{(n)}(f)\rangle_t = \int_0^t \int_0^{+\infty} r(x) \int_0^1 (f(qx) + f((1-q)x) - f(x))^2 F(\mathrm{d}q) \boldsymbol{X}_s^{(n)}(\mathrm{d}x)\,\mathrm{d}s. \tag{5.13}$$

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

As the set of signed measure is not metrisable, we cannot adapt the proof of Theorem 1.2. Inspired by [33, Section 3.2] and [5, Section 6], we consider $\eta^{(n)}$ as an operator in a Sobolev space, and use the Hilbertian properties of this space to prove tightness. Let us explain the Sobolev space that we will use. Let $p > 0$ and $j \in \mathbb{N}$. The set $W^{j,p}$ is the closure of C_c^∞, which is the set of functions of class C^∞ from $\mathbb{R}_+$ into $\mathbb{R}$ with compact support, embedded with the norm, for all $f \in W^{j,p}$,

$$\|f\|_{W^{j,p}}^2 = \sum_{k=0}^{j} \int_0^\infty \left(\frac{f^{(k)}(x)}{1+x^p}\right)^2 \mathrm{d}x.$$

The set $W^{j,p}$ is a Hilbert space and we denote by $W^{-j,p}$ its dual space. Let $C^{j,p}$ be the space of function f of class C^j such that, for all $k \le j$,

$$\lim_{x\to+\infty} \frac{f^{(k)}(x)}{1+x^p} = 0.$$

We embed it with the norm, for all $f \in C^{j,p}$,

$$\|f\|_{C^{j,p}} = \sum_{k=0}^{j} \sup_{x\ge 0} \frac{f^{(k)}(x)}{1+x^p}.$$

The set $C^{j,p}$ is also a Banach space and we denote by $C^{-j,p}$ its dual space. These spaces satisfy the following continuous injection [33, Section 3.2]:

$$C^{j,p} \subset W^{j,p+1} \quad \text{and} \quad W^{1+j,p} \subset C^{j,p}. \tag{5.14}$$

Or, equivalently, if for every function f, we have

$$\|f\|_{W^{j,p+1}} \le C\|f\|_{C^{j,p}} \quad \text{and} \quad \|f\|_{C^{j,p}} \le C\|f\|_{W^{j+1,p}}.$$

The first embedding/inequality prove that the tightness in $W^{j,p+1}$ implies the tightness in $C^{j,p}$. The second is useful for some upper bounds. For instance, we have the following lemma.

Lemma 5.4. *If $(e_k)_{k\ge1}$ is a basis of $W^{2,1}$ then we have, for all $k \ge 0$ and $x \in E$,*

$$\sum_{k\ge1} e_k(x)^2 \le C(1+x^2).$$

Proof. We have $\delta_x : f \mapsto f(x)$ is an operator on $W^{2,1}$. We have, for all $f \in W^{2,1}$,

$$|\delta_x f| \le (1+x)\|f\|_{C^{0,1}} \le C(1+x)\|f\|_{W^{1,1}} \le C(1+x)\|f\|_{W^{2,1}}.$$

But, by Parseval's identity, we obtain $\|\delta_x\|_{W^{-2,1}}^2 = \sum_{k\ge1} e_k(x)^2$, completing the proof. □

We introduce the trace $(\langle\langle \widetilde{\boldsymbol{M}}^{(n)}\rangle\rangle_t)_{t\ge0}$ of $(\widetilde{\boldsymbol{M}}_t^{(n)})_{t\ge0}$. It is defined such that

$$(\|\widetilde{\boldsymbol{M}}_t^{(n)}\|_{W^{-2,1}}^2 - \langle\widetilde{\boldsymbol{M}}^{(n)}\rangle\rangle_t)_{t\ge0}$$

is a local martingale. Then

$$\|\widetilde{\boldsymbol{M}}_t^{(n)}\|_{W^{-2,1}}^2 = \sum_{k\ge1} \widetilde{\boldsymbol{M}}_t^{(n)}(e_k)^2,$$

where $(e_k)_{k\geq 1}$ is a basis of $W^{2,1}$. Then, by (5.13), we obtain

$$\langle\langle \widetilde{\boldsymbol{M}}^{(n)}\rangle\rangle_t = \sum_{k\geq 1}\int_0^t \int_0^{+\infty} r(x)\int_0^1 (e_k(qx)+e_k((1-q)x)-e_k(x))^2 F(\mathrm{d}q)\boldsymbol{X}_s^{(n)}(\mathrm{d}x)\,\mathrm{d}s.$$

Now, we first prove the tightness of $(\eta^{(n)})_{n\geq 1}$ then Theorem 5.3.

Lemma 5.5. (Tightness.) *The sequence $(\eta^n)_{n\geq 1}$ is tight in $\mathbb{D}([0,T], W^{-2,1})$.*

Proof. By [27, Theorem 2.2.2] and [27, Theorem 2.3.2] (see also [33, Lemma C]), it is enough to prove that

(i) $\mathbb{E}[\sup_{s\leq t}\|\eta_s^n\|^2_{W^{-2,1}}] < +\infty$,

(ii) for all $n\in\mathbb{N}$, for all $\varepsilon,\rho>0$, there exists $\delta>0$ such that for each stopping times S_n bounded by T,

$$\limsup_{n\to+\infty}\sup_{0\leq u\leq\delta}\mathbb{P}(\|\widetilde{\boldsymbol{V}}^{(n)}_{S_n+u}-\widetilde{\boldsymbol{V}}_{S_n}\|_{W^{-2,1}}\geq\eta)\leq\varepsilon,$$

$$\limsup_{n\to+\infty}\sup_{0\leq u\leq\delta}\mathbb{P}(|\langle\langle\widetilde{\boldsymbol{M}}^{(n)}\rangle\rangle_{S_n+u}-\langle\langle\widetilde{\boldsymbol{M}}^{(n)}\rangle\rangle_{S_n}|\geq\eta)\leq\varepsilon.$$

For the first point, using lemma 5.1, there exists $C>0$ such that

$$\begin{aligned}\sum_{k\geq 1}\langle\widetilde{\boldsymbol{M}}_t^{(n)}(e_k)\rangle &\leq \int_0^t \bar{r}\int_0^1 3\sum_{k\geq 1}(e_k^2(qx)+e_k^2((1-q)x)+e_k^2(x))F(\mathrm{d}q)\boldsymbol{X}_s^{(n)}(\mathrm{d}x)\,\mathrm{d}s\\ &\leq C\boldsymbol{X}_0^{(n)}(1+x).\end{aligned}$$

Then, since $\|\widetilde{\boldsymbol{M}}_t^{(n)}\|^2_{W^{-2,1}}=\sum_{k\geq 1}(\widetilde{\boldsymbol{M}}_t^{(n)}(e_k))^2$, from Doob's inequality and (5.11), we have

$$\mathbb{E}\Big[\sup_{t\in[0,t]}\|\widetilde{\boldsymbol{M}}_t^{(n)}\|^2_{W^{-2,1}}\Big]\leq C',$$

where $C'>0$. Then there exits $C''>0$ such that

$$\|\eta_t^{(n)}\|^2_{W^{-2,1}}\leq\|\eta_0^{(n)}\|^2_{W^{-2,1}}+\|\widetilde{\boldsymbol{V}}_t^{(n)}\|^2_{W^{-2,1}}+\|\widetilde{\boldsymbol{M}}_t^{(n)}\|^2_{W^{-2,1}}\leq C''+\|\widetilde{\boldsymbol{V}}_t^{(n)}\|^2_{W^{-2,1}}.$$

And as

$$\|\widetilde{\boldsymbol{V}}_t^{(n)}\|^2_{W^{-2,1}}\leq C\int_0^t\sup_{w\leq s}\|\eta_s^{(n)}\|^2_{W^{-2,1}}\,\mathrm{d}s,$$

from the Gronwall lemma, we obtain

$$\mathbb{E}\Big[\sup_{s\leq t}\|\eta_s^{(n)}\|^2_{W^{-2,1}}\Big]\leq K$$

for a certain constant K. Finally, for the second point, we have

$$\mathbb{E}[\|\widetilde{\boldsymbol{V}}^{(n)}_{S_n+u}-\widetilde{\boldsymbol{V}}^{(n)}_{S_n}\|_{W^{-2,1}}]\leq\mathbb{E}\bigg[K'\int_{S_n}^{S_n+u}\sup_{s\leq T}\|\eta_s^{(n)}\|^2_{W^{-2,1}}\,\mathrm{d}s\bigg]\leq K''u.$$

Here, K' and K'' are two constants. Using the Markov–Chebyshev inequality, we prove the Aldous condition. We similarly prove that $\langle\langle\widetilde{\boldsymbol{M}}^{(n)}\rangle\rangle$ satisfies the Aldous condition. We deduce that $(\eta^{(n)})_{n\geq 1}$ is tight. □

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

Proof of Theorem 5.3. Let $\widetilde{\boldsymbol{M}}$ be a continuous Gaussian process with quadratic variation satisfying, for every $f \in C^{2,0}$ ($\subset W^{2,1}$) and $t \in [0, T]$,

$$\langle \widetilde{\boldsymbol{M}}(f) \rangle_t = \int_0^t \int_0^{+\infty} r(x) \int_0^1 (f(qx) + f((1-q)x) - f(x))^2 F(\mathrm{d}q) \boldsymbol{X}_s(\mathrm{d}x).$$

Since there exists C_f such that, for all $f \in C^{2,0}$, $\sup_{t\in[0,T]} |\widetilde{\boldsymbol{M}}^{(n)}(f)| \le C_f/\sqrt{n}$ and $\langle \widetilde{\boldsymbol{M}}^{(n)} \rangle_t$ converges in law to $\langle \widetilde{\boldsymbol{M}} \rangle_t$, then by [26, Theorem 3.11, p. 473], $\widetilde{\boldsymbol{M}}^{(n)}(f)$ converges to $\widetilde{\boldsymbol{M}}(f)$ in distribution, as n tends to ∞.

By Lemma 5.5 and (5.14) , the sequence $(\eta^{(n)})_{n\ge1}$ is also tight in $C^{-2,0}$. Let η be an accumulation point. Since its martingale part $\widetilde{\boldsymbol{M}}$ in its Doob's decomposition is almost surely continuous, then η is also almost surely continuous. Hence, η is a solution of (5.12). Using the Gronwall inequality, we obtain the uniqueness of this equation, in $C([0, T], C^{-2,0})$, up to a Gaussian white noise $\widetilde{\boldsymbol{M}}$. We deduce the announced result. □

6. Another two examples

6.1. Space-structured population model

Here, we study an example which can model the cell localisation. One cell moves following a diffusion on $E \subset \mathbb{R}^d$, $d \ge 1$, and when it dies, its offspring is localised at the same place. Hence, in all this section the branching is local; that is, for all $k \ge 0$, for all $j \le k$, for all $x \in E$, for all $\theta \in [0, 1]$, $F_j^{(k)}(x, \theta) = x$.

6.1.1. *Branching Ornstein–Uhlenbeck.* In this subsection we consider the model of [17, Example 10]. Assume that $E = \mathbb{R}^d$ and G is given by

$$Gf(x) = \tfrac{1}{2}\sigma^2 \Delta f(x) - gx \cdot \nabla f(x)$$

for every $f \in C_c^2(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$, where $d \in \mathbb{N}^*$ and $\sigma, g > 0$. Also assume that the division is dyadic; that is, $p_2 = 1$, with rate $r(x) = bx^2 + a$, where $a, b \ge 0$ and a or b is not null. Here $x^2 = \|x\|^2 = x.x$. If $g > \sqrt{2b}$ then we add the notation

$$\Gamma = \frac{g - \sqrt{g^2 - 2b\sigma^2}}{2\sigma^2} \quad \text{and} \quad \alpha = \sqrt{g^2 - 2b\sigma^2}.$$

We also denote by π_∞ the Gaussian measure whose density is defined by

$$x \mapsto \left(\frac{\alpha}{\pi\sigma^2}\right) \exp\left(-\frac{\alpha}{\sigma^2}x^2\right).$$

From our main theorem, we deduce the following corollary.

Corollary 6.1. (Limit theorem for an branching Ornstein–Uhlenbeck process.) *If $g > \sigma\sqrt{2b}$ and $X_0^\varnothing = x \in \mathbb{R}^d$ then, for any continuous and bounded f, we have*

$$\lim_{t\to+\infty} \frac{1}{N_t} \sum_{u\in\mathcal{V}_t} f(X_t^u) = \frac{\int_{\mathbb{R}^d} f(y) \mathrm{e}^{\Gamma y^2} \pi_\infty(\mathrm{d}y)}{\int_{\mathbb{R}^d} \mathrm{e}^{\Gamma y^2} \pi_\infty(\mathrm{d}y)} \quad \textit{in probability.}$$

In particular,

$$\mathbb{E}[N_t] = \mathrm{e}^{\lambda t + \Gamma x^2} \left(\frac{\alpha}{\pi\sigma^2}\right) \int_{\mathbb{R}^d} \mathrm{e}^{-\Gamma y^2} \exp\left(-\frac{\alpha(y - x\mathrm{e}^{-\alpha t/\sigma^2})^2}{\sigma^2(1 - \mathrm{e}^{-2\alpha t/\sigma^2})}\right) \mathrm{d}y,$$

where $\lambda = \frac{1}{2}(g - \sqrt{g^2 - 2b\sigma^2}) + a$ is the Malthus parameter.

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

The assumption $g > \sigma\sqrt{2b}$ is quite natural. Indeed, the quantity g/σ represents the rate of convergence to equilibrium of the Ornstein–Uhlenbeck process. Also the larger it is, the closer from the origin the trajectories are. In contrast, in a certain sense, parameter b keeps the particles away from the origin. Indeed, the more distant from the origin the cell is, the more prolific it is. Condition $g > \sigma\sqrt{2b}$ is then a pay-off type assumption.

Proof. If $V : x \mapsto \mathrm{e}^{\lambda x^2}$ then it is an eigenvector of $\mathcal{G}$, which is defined for every $f \in C_c^2(E)$ by

$$\mathcal{G} f(x) = Gf(x) + r(x) f(x).$$

A computation similar to Lemma 5.1 ensures nonexplosion and

$$z_t(V, \delta_x) < +\infty.$$

Hence, Lemma 3.2 gives the second equality. The limit result comes from Theorem 1.1 and that V^2 is integrable with respect to the semigroup of the auxiliary process. □

Remark 6.1. (*Another eigenelement.*) Note that if $V_2 : x \mapsto \mathrm{e}^{\lambda_2 x^2}$ then it is an eigenvector of $\mathcal{G}$, associated to the eigenvalue

$$\lambda_2 = \tfrac{1}{2}\Big(g + \sqrt{g^2 - 2b\sigma^2}\Big) + a.$$

But in this case, the auxiliary process is not ergodic and we are not able to deduce any convergence from our main theorem.

General case. Let us assume that G is the generator of a diffusive Markov process. If the state space E is bounded then we can find sufficient conditions to the eigenproblem in [41, Section 3] and [41, Theorem 5.5]. For instance, under some assumptions, we have

$$\lambda_0 = \lim_{t\to+\infty} \ln\Big(\sup_{x\in E} \mathbb{E}[N_t \mid X_0^{\varnothing} = x]\Big).$$

If E is not bounded then we can see the results of [24], [42]. This example was developed in [17]. The authors proved a strong law of large numbers, closely related to Theorem 1.1.

6.2. Self-similar fragmentation

Self-similar mass fragmentation processes are characterised by:

- the index of self-similarity $\alpha \in \mathbb{R}$;
- a so-called dislocation measure ν on $\mathcal{S} = \{s = (s_i)_{i\in\mathbb{N}} \mid \lim_{i\to+\infty} s_i = 0,\ 1 \geq s_j \geq s_i \geq 0, \text{ for all } j \leq i\}$, which satisfies

$$\nu(1, 0, 0, \dots) = 0 \quad \text{and} \quad \int_{\mathcal{S}} (1 - s)\nu(\mathrm{d}s) < +\infty.$$

If $\nu(\mathcal{S}) < +\infty$ then the dynamics are as follows:

- a block of mass x remains unchanged for exponential periods of time with parameter $x^{\alpha}\nu(\mathcal{S})$;
- a block of mass x dislocates into a mass partition xs, where $s \in \mathcal{S}$, at rate $\nu(\mathrm{d}s)$;
- there are finitely many dislocations over any finite time horizon.

https://doi.org/10.1017/apr.2017.12 Published online by Cambridge University Press

The last point is not verified when $\nu(\mathcal{S}) = +\infty$. In this case, there is a countably infinite number of dislocations over any finite time horizon. So, when $\nu(\mathcal{S}) < +\infty$, our setting captures this model with the following parameters:

$$G = 0, \qquad r(x) = x^{\alpha}\nu(\mathcal{S}),$$

and, for every continuous and bounded function f,

$$\int_0^1 \sum_{k\geq 0} p_k(x) \sum_{j=1}^{k} f(F_j^k x, \theta)\,\mathrm{d}\theta = \int_{\mathcal{S}} \sum_{i\geq 0} f(s_i x)\frac{\nu(\mathrm{d}s)}{\nu(\mathcal{S})}.$$

Hence, in this case we have

$$\mathcal{G} f(x) = x^{\alpha}\nu(\mathcal{S})\biggl(\biggl(\int_{\mathcal{S}} \sum_{i\geq 0} f(s_i x)\frac{\nu(\mathrm{d}s)}{\nu(\mathcal{S})}\biggr) - f(x)\biggr)$$

for every continuous and bounded f, and $V : x \mapsto x^p$ is an eigenvector. See [6] for further details. Theorem 1.1 does not give a relevant result in this case since the auxiliary process have a trivial behaviour (convergence to 0).

Acknowledgements

I would like to express my gratitude to my PhD supervisor Djalil Chafaï for his encouragement and essential help on the form and the content of this paper, Viet Chi Tran for pointing out some references and for fruitful discussions, and, finally, the anonymous referee for valuable remarks and suggestions which helped improve the manuscript.

This work was financially supported by the ANR PIECE (grant number ANR-12-JS01-0006-01) and the Chair Modélisation Mathématique et Biodiversité.

References

[1] Aïdékon, E., Berestycki, J., Brunet, É. and Shi, Z. (2013). Branching Brownian motion seen from its tip. *Prob. Theory Relat. Fields* **157,** 405–451.

[2] Athreya, K. B. and Ney, P. E. (2004). *Branching Processes*. Dover Publications, Mineola, NY.

[3] Balagué, D., Cañizo, J. A. and Gabriel, P. (2013). Fine asymptotics of profiles and relaxation to equilibrium for growth-fragmentation equations with variable drift rates. *Kinet. Relat. Models* **6,** 219–243.

[4] Bansaye, V. and Tran, V. C. (2011). Branching Feller diffusion for cell division with parasite infection. *ALEA Lat. Amer. J. Prob. Math. Statist.* **8,** 95–127.

[5] Bansaye, V., Delmas, J.-F., Marsalle, L. and Tran, V. C. (2011). Limit theorems for Markov processes indexed by continuous time Galton–Watson trees. *Ann. App. Prob.* **21,** 2263–2314.

[6] Bertoin, J. (2006). *Random Fragmentation and Coagulation Processes* (Camb. Stud. Adv. Math. **102**). Cambridge University Press.

[7] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd edn. John Wiley, New York.

[8] Chafaï, D., Malrieu, F. and Paroux, K. (2010). On the long time behavior of the TCP window size process. *Stoch. Process. Appl.* **120,** 1518–1534.

[9] Chen, M.-F. (2004). *From Markov chains to Non-Equilibrium Particle Systems*, 2nd edn. World Scientific, River Edge, NJ.

[10] Dellacherie, C. and Meyer, P.-A. (1980). *Probabilités et Potentiel. Chapitres V à VIII* (Current Sci. Indust. Topics **1385**). Hermann, Paris.

[11] Delmas, J.-F. and Marsalle, L. (2010). Detection of cellular aging in a Galton-Watson process. *Stochastic Process. Appl.* **120,** 2495–2519.

[12] Doumic Jauffret, M. and Gabriel, P. (2010). Eigenelements of a general aggregation-fragmentation model. *Math. Models Methods Appl. Sci.* **20,** 757–783.

[13] Doumic, M., Maia, P. and Zubelli, J. P. (2010). On the calibration of a size-structured population model from experimental data. *Acta Biotheoret.* **58,** 405–413.

[14] Doumic, M., Hoffmann, M., Reynaud-Bouret, P. and Rivoirard, V. (2012). Nonparametric estimation of the division rate of a size-structured population. *SIAM J. Numer. Anal.* **50,** 925–950.

[15] Eberle, A. (1999). *Uniqueness and Non-Uniqueness of Semigroups Generated by Singular Diffusion Operators*. (Lecture Notes Math. **1718**). Springer, Berlin.

[16] Engländer, J. and Winter, A. (2006). Law of large numbers for a class of superdiffusions. *Ann. Inst. H. Poincaré Prob. Statist.* **42,** 171–185.

[17] Engländer, J., Harris, S. C. and Kyprianou, A. E. (2010). Strong law of large numbers for branching diffusions. *Ann. Inst. H. Poincaré Prob. Statist.* **46,** 279–298.

[18] Ethier, S. N. and Kurtz, T. G. (1986). *Markov Processes*. John Wiley, New York.

[19] Folland, G. B. (1999). *Real Analysis*, 2nd edn. John Wiley, New York.

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

[21] Grigorescu, I. and Kang, M. (2010). Steady state and scaling limit for a traffic congestion model. *ESAIM Prob. Statist.* **14,** 271–285.

[22] Guillemin, F., Robert, P. and Zwart, B. (2004). AIMD algorithms and exponential functionals. *Ann. Appl. Prob.* **14,** 90–117.

[23] Harris, S. C. and Williams, D. (1996). Large deviations and martingales for a typed branching diffusion. I. *Astérisque* **236,** 133–154.

[24] Hislop, P. D. and Sigal, I. M. (1996). *Introduction to Spectral Theory* (Appl. Math. Sci. **113**). Springer, New York.

[25] Ikeda, N. and Watanabe, S. (1989). *Stochastic Differential Equations and Diffusion Processes* (North-Holland Math. Library **24**), 2nd edn. North-Holland, Amsterdam.

[26] Jacod, J. and Shiryaev, A. N. (2003). *Limit Theorems for Stochastic Processes* (Fundamental Principles Math. Sci. **288**), 2nd edn. Springer, Berlin.

[27] Joffe, A. and Métivier, M. (1986). Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Adv. Appl. Prob.* **18,** 20–65.

[28] Kallenberg, O. (2002). *Foundations of Modern Probability*, 2nd edn. Springer, New York.

[29] Kolokoltsov, V. N. (2011). *Markov Processes, Semigroups and Generators* (De Gruyter Stud. Math. **38**). Walter de Gruyter, Berlin.

[30] Kubitschek, H. E. (1969). Growth during the bacterial cell cycle: analysis of cell size distribution. *Biophys. J.* **9,** 792–809.

[31] Laurençot, P. and Perthame, B. (2009). Exponential decay for the growth-fragmentation/cell-division equation. *Commun. Math. Sci.* **7,** 503–510.

[32] Löpker, A. H. and van Leeuwaarden, J. S. H. (2008). Transient moments of the TCP window size process. *J. Appl. Prob.* **45,** 163–175.

[33] Méléard, S. (1998). Convergence of the fluctuations for interacting diffusions with jumps associated with Boltzmann equations. *Stoch. Stoch. Reports* **63,** 195–225.

[34] Méléard, S. and Roelly, S. (1993). Sur les convergences étroite ou vague de processus à valeurs mesures. *C. R. Acad. Sci. Paris* **317,** 785–788.

[35] Méléard, S. and Tran, V. C. (2012). Slow and fast scales for superprocess limits of age-structured populations. *Stoch. Process. Appl.* **122,** 250–276.

[36] Meyn, S. P. and Tweedie, R. L. (1993). Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. Appl. Prob.* **25,** 518–548.

[37] Michel, P. (2006). Existence of a solution to the cell division eigenproblem. *Math. Models Methods Appl. Sci.* **16,** 1125–1153.

[38] Ott, T., Kemperman, J. and Mathis, M. (1996). The stationary behavior of ideal TCP congestion avoidance. Unpublished manuscript. Available at http://www.teunisott.com/.

[39] Perthame, B. (2007). *Transport Equations in Biology*. Birkhäuser, Basel.

[40] Perthame, B. and Ryzhik, L. (2005). Exponential decay for the fragmentation or cell-division equation. *J. Differential Equat.* **210,** 155–177.

[41] Pinsky, R. G. (1995). *Positive Harmonic Functions and Diffusion* (Camb. Stud. Adv. Math. **45**). Cambridge University Press.

[42] Reed, M. and Simon, B. (1978). *Methods of Modern Mathematical Physics. IV. Analysis of Operators*. Academic Press, New York.

[43] ROELLY-COPPOLETTA, S. (1986). A criterion of convergence of measure-valued processes: application to measure branching processes. *Stochastics* **17,** 43–65.

[44] TRAN, V. C. (2006). Modèles particulaires stochastiques pour des problèmes d'évolution adaptative et pour l'approximation de solutions statistiques. Doctoral Thesis, Universités Paris X - Nanterre. Available at http://tel.archives-ouvertes.fr/tel-00125100.

[45] VILLANI, C. (2003). *Topics in Optimal Transportation* (Grad. Stud. Math. **58**). American Mathematical Society, Providence, RI.