ECOLE POLYTECHNIQUE

CENTRE DE MATHÉMATIQUES APPLIQUÉES

UMR CNRS 7641

91128 PALAISEAU CEDEX (FRANCE). Tél: 01 69 33 46 00. Fax: 01 69 33 46 46
http://www.cmap.polytechnique.fr/

Limit theorems for Markov processes indexed by continuous time Galton-Watson trees.

V. Bansaye, J-F Delmas, L. Marsalle, V. C. Tran

R.I. 679 *December 2009*

Abstract

We study the evolution of a particle system whose genealogy is given by a supercritical continuous time Galton-Watson tree. The particles move independently according to a Markov process and when a branching event occurs, the offspring locations depend on the position of the mother and the number of offspring. We prove a law of large numbers for the empirical measure of individuals alive at time t. This relies on a probabilistic interpretation of its intensity by mean of an auxiliary process. This latter has the same generator as the Markov process along the branches plus additional branching events, associated with jumps of accelerated rate and biased distribution. This comes from the fact that choosing an individual uniformly at time t favors lineages with more branching events and larger offspring number. The central limit theorem is considered on a special case. Several examples are developed, including applications to splitting diffusions, cellular aging, branching Lévy processes and ancestral lineages.

Key words. Branching Markov process. Branching diffusion. Limit theorems. Size biased reproduction distribution. Ancestral lineage.

A.M.S. Classification. 60J80, 60F17, 60F15, 60F05.

1 Introduction and main results

We consider a continuous time Galton-Watson tree $\mathbb{T}$, *i.e.* a tree where each branch lives during an independent exponential time of constant mean $1/r$, then splits into a random number of new branches given by an independent random variable (r.v.) ν of law $(p_k, k \in \mathbb{N})$. We are interested in the following process indexed by this tree. Along the edges of the tree, the process evolves as a càdlàg strong Markov process $(X_t)_{t\geq 0}$ with values in a Polish space E and with infinitesimal generator L of domain $D(L)$. The branching event is nonlocal and is described by a Markov kernel which depends on the state x of the mother just before the branching event and the number $\nu = k$ of offspring. Then, we restart the process for each of the new born branches at $F_1^{(k)}(x, \theta) \dots F_k^{(k)}(x, \theta)$ parametrized by a uniform random variable θ on $[0, 1]$. The new born branches evolve independently from each other.

This process is a branching Markov process, for which there has been a vast literature. We refer to Asmussen and Hering [?], Dawson [?] and Dawson *et al.* [?] for nonlocal branching processes similar to those considered here. Whereas the usual literature turns to limit theorems that consider superprocesses limits corresponding to high densities of small and rapid particles (see *e.g.* Dawson [?], Dynkin [?], Evans and Steinsaltz [?]), we stick here with the discrete continuous tree which we aim to characterize.

Let us also mention some results in the discrete time case. Markov chains indexed by a binary tree have been first studied in the symmetric independent case (see *e.g.* Athreya and Kang [?], Benjamini and Peres [?]), where for every x, $F^{(1)}(x, \theta)$ and $F^{(2)}(x, \theta)$ are i.i.d. A motivation for considering asymmetric branching comes from models for cell division. Indeed the binary tree can be used to describe a dividing cell genealogy in discrete time. The Markov chain indexed by this binary tree then indicates the evolution of some characteristic of the cell, such as its growth rate, its quantity of proteins or parasites... and depends on division events. Moreover experiments (Stewart *et al.* [?]) indicate that the transmission of this characteristic in the two daughter cells may be asymmetric. See Bercu *et al.* [?] or Guyon [?] for asymmetric models for cellular aging and Bansaye [?] for parasite infection. In Delmas and Marsalle [?] a generalization of these models where there might be 0, 1 or 2 daughters is studied. Indeed under stress conditions, cells may divide less or even die. The branching Markov chain, which in their case represents the cell's growth rate, is then restarted for each daughter cell at a value that depends on the mother's growth rate and on the total number of daughters.

We investigate the continuous time case and allow both asymmetry and random number of offspring. Let us give two simple examples of this model for parasite infection. In the first case, the cell divides in two daughter cells after an exponential time and a random fraction of parasites goes in one of the daughter cell, whereas the rest goes in the second one. In the second case, the cell divides in j daughter cells and the process X is equally shared between each of the j daughters: $\forall k \in \{1, \dots, j\}$, $F_k^{(j)}(x, \theta) = x/j$. Notice that another similar model has been investigated in Evans and Steinsaltz [?] where the evolution of damages in a population of dividing cells is studied, but with a superprocess point of view. The authors assume that the cell's death rate depends on the damage of the cell, which evolves as a diffusion between two fissions. When a division occurs, there is an unbalanced transmission of damaged material that leads to the consideration of nonlocal births. Further examples are developed in Section 5.

Our main purpose is to characterize the empirical distribution of this process. More precisely, if we denote by N_t the size of the living population V_t at time t, and if $(X_t^u)_{u\in V_t}$ denotes the values of the Markov process for the different individuals of V_t, we will focus on the following probability measure which describes the state of the population

$$\frac{\mathbf{1}_{\{N_t>0\}}}{N_t} \sum_{u\in V_t} \delta_{X_t^u}(dx), \quad t \in \mathbb{R}_+.$$

This is linked to the value of the process of an individual chosen uniformly at time t, say $U(t)$, as we can see from this simple identity:

$$\mathbb{E}\Big[\frac{\mathbf{1}_{\{N_t>0\}}}{N_t} \sum_{u\in V_t} f(X_t^u)\Big] = \mathbb{E}\big[\mathbf{1}_{\{N_t>0\}} f(X_t^{U(t)})\big].$$

We show that the distribution of the path leading from the ancestor to an uniformly chosen individual can be approximated by mean of an *auxiliary Markov process* Y with infinitesimal generator characterized by: $\forall f \in D(L)$,

$$Af(x) = Lf(x) + rm \sum_{k=1}^{+\infty} \frac{p_k}{m} \int_0^1 \sum_{j=1}^{k} \Big(f(F_j^{(k)}(x,\theta)) - f(x)\Big) d\theta \tag{1.1}$$

where we recall that r denotes the particle branching rate and where we introduce $m = \sum_{k=1}^{+\infty} kp_k$ the mean number of offspring. In this paper, we will be interested in the super-critical case $m > 1$, even if some remarks are made for the critical and sub-critical cases. The auxiliary process has the same generator as the Markov process running along the branches, plus jumps due to the branching. However, we can observe a bias phenomenon: the apparent jump rate rm is equal to the original rate r times the mean offspring number m and the apparent offspring distribution is the size-biased distribution $(kp_k/m, k \in \mathbb{N})$. For $m > 1$ for instance, this is heuristically explained by the fact that when one chooses an individual uniformly in the population at time t, an individual belonging to a lineage with more generations or with prolific ancestors is more likely to be chosen. Such biased phenomena have already been observed in the field of branching processes (see *e.g.* Chauvin *et al.* [?], Harris and Roberts [?]). Here, we allow nonlocal births, prove pathwise results and establish laws of large numbers when Y is ergodic. Our approach is entirely based on a probabilistic interpretation, *via* the auxiliary process.

In case Y is ergodic, we prove the laws of large numbers stated in Theorem 1.1 and 1.3, where W stands for the renormalized asymptotic size of the number of individuals at time t (*e.g.* Theorems 1 and 2 p. 111 of Athreya and Ney [?]):

$$W := \lim_{t\to+\infty} N_t/\mathbb{E}[N_t] \quad \text{a.s.} \quad \text{and} \quad \{W>0\} = \{\forall t\geq 0, N_t>0\} \quad \text{a.s.}$$

Theorem 1.1. *If the auxiliary process Y is ergodic with invariant measure π, we have for any real continuous bounded function f on E:*

$$\lim_{t\to\infty} \frac{\mathbf{1}_{\{N_t>0\}}}{N_t} \sum_{u\in V_t} f\big(X_t^u\big) = \mathbf{1}_{\{W>0\}} \int_E f(x)\pi(dx) \qquad \textit{in probability.} \tag{1.2}$$

This result in particular implies that for such function f,

$$\lim_{t\to+\infty} \mathbb{E}\big[f(X_t^{U(t)}) \,|\, N_t>0\big] = \int_E f(x)\pi(dx), \tag{1.3}$$

where $U(t)$ stands for a particle taken at random in the set V_t of living particles at time t.

Theorem 1.1 is a consequence of Theorem 4.2 (which gives similar results under weaker hypotheses) and of Remark 4.1. The convergence is proved using L^2 techniques.

Theorem 1.1 also provides a limit theorem for the empirical distribution of the tree indexed Markov process.

Corollary 1.2. *Under the assumption of Theorem 1.1,*

$$\lim_{t\to\infty} \frac{\mathbf{1}_{\{N_t>0\}}}{N_t} \sum_{u\in V_t} \delta_{X_t^u}(dx) = \mathbf{1}_{\{W>0\}}\,\pi(dx) \qquad \textit{in probability} \tag{1.4}$$

where the space $\mathcal{M}_F(E)$ of finite measures on E is embedded with the weak convergence topology.

We also give in Propositions 6.1 and 6.4 a result on the associated fluctuations. Notice that contrarily to the discrete case treated in [?], the fluctuation process is a Gaussian process with a finite variational part.

In addition, we generalize the result of Theorem 1.1 to ancestral paths of particles (Theorem 1.3):

Theorem 1.3. *Suppose that Y is ergodic with invariant measure π and that for any bounded measurable function f, $\lim_{t\to+\infty} \mathbb{E}_x[f(Y_t)] = \int_E f(x)\pi(dx)$, then for any real bounded measurable function φ on the Skorohod space $\mathbb{D}([0,T],E)$, we have*

$$\lim_{t\to\infty} \frac{\mathbf{1}_{\{N_t>0\}}}{N_t} \sum_{u\in V_t} \varphi(X_s^u,\, t-T\leq s<t) = \mathbb{E}_\pi\big[\varphi(Y_s,\, s<T)\big]\mathbf{1}_{\{W\neq 0\}} \quad \textit{in probability,}$$

where, for simplicity, X_s^u stands for the value of the tree indexed Markov process at time s for the ancestor of u living at this time.

Biases that are typical to all renewal problems have been known for long time in the literature (see *e.g.* Feller [?], Vol. 2 Chap. 1). Size biased trees are linked with the consideration of Palm measures, themselves related to the problem of building a population around the path of an individual picked uniformly at random from the population alive at a certain time t. In Chauvin *et al.* [?] and in Hardy and Harris [?], a spinal decomposition is obtained for continuous time branching processes. Their result states that along the chosen line of descent, which constitutes a bridge between the initial condition and the position of the particle chosen at time t, the birth times of the new branches form a homogeneous Poisson Point Process of intensity rm while the reproduction law that is seen along the branches is given by $(kp_k/m, k \in \mathbb{N})$. Other references for Palm and size-biased Galton-Watson decompositions in continuous time can be found in Gorostiza *et al.* [?], Geiger and Kauffmann [?], Geiger [?]. Notice that biases for an individual chosen uniformly in continuous tree had previously been observed by Samuels [?] and Biggins [?]. In the same vein, we refer to Nerman and Jagers [?] for consideration of the pedigree of an individual chosen randomly at time t and to Lyons *et al.* [?], Geiger [?] for spinal decomposition for size biased Galton-Watson processes in the discrete case.

Other motivating topics for this kind of results come from branching random walks (see *e.g.* Biggins [?], Rouault [?]) and homogeneous fragmentation (see Bertoin [?, ?]). We refer to the examples in Section 5 for more details.

The law of large numbers that we obtain is a continuous time version of the law of large numbers in Benjamini and Peres [?], Delmas and Marsalle [?], with possible asymmetric branching and random number of offspring. Similar laws of large numbers are obtained in Engländer and Turaev [?], Engländer and Winter [?] or Evans and Steinsaltz [?], with a superprocess renormalization, and in Georgii and Baake [?], with spectral techniques. In these works, the limit is characterized by mean of eigenfunctions of the generator A. Here, we stick with discrete continuous time branching processes and a statement of the results via the auxiliary process. This is interesting for statistical applications in which the population associated with the tree can not be considered as large, for instance.

In Section 2, we define our Markov process indexed by a continuous time Galton-Watson tree. We start with the description of the tree and then provide a measure-valued description of the process of interest. In Section 3, we build an *auxiliary process* Y and prove that its law is deeply related to the distribution of the lineage of an individual drawn uniformly in the population. In Section 4, we establish the laws of large numbers mentioned in Theorem 1.1 and 1.3. Several examples are then investigated in Section 5: splitting diffusions indexed by a Yule tree, a model for cellular aging generalizing [?] and an application to nonlocal branching random walks. Finally, a central limit theorem is considered for the case of splitting diffusions in Section 6.

2 Tree indexed Markov processes

We first give a description of the continuous time Galton-Watson trees and preliminary estimates in Section 2.1. Section 2.2 is devoted to the definition of tree indexed Markov processes.

2.1 Galton-Watson trees in continuous time

In a first step, we recall some definitions about discrete trees. In a second step, we introduce continuous time and finally, in a third step, we give the definition of the Galton-Watson tree in continuous time. For all this section, we refer mainly to [?, ?, ?].

Discrete trees. Let

$$\mathcal{U} = \bigcup_{m=0}^{+\infty} (\mathbb{N}^*)^m, \tag{2.1}$$

where $\mathbb{N}^* = \{1, 2, \dots\}$ and by convention $(\mathbb{N}^*)^0 = \{\emptyset\}$. For $u \in (\mathbb{N}^*)^m$, we define $|u| = m$ the generation of u. If $u = (u_1, \dots, u_n)$ and $v = (v_1, \dots, v_p)$ belong to $\mathcal{U}$, we write $uv = (u_1, \dots, u_n, v_1, \dots, v_p)$ for the concatenation of u and v. We identify both $\emptyset u$ and $u\emptyset$ with u. We also introduce the following order relation: $u \preceq v$ if there exists $w \in \mathcal{U}$ such that $v = uw$; if furthermore $w \neq \emptyset$, we write $u \prec v$. Finally, for u and v in $\mathcal{U}$ we define their most recent common ancestor (MRCA), denoted by $u \wedge v$, as the element $w \in \mathcal{U}$ of highest generation such that $w \preceq u$ and $w \preceq v$.

Definition 2.1. *A rooted ordered tree $\mathcal{T}$ is a subset of $\mathcal{U}$ such that:*

(i) $\emptyset \in \mathcal{T}$,

(ii) if $v \in \mathcal{T}$ then $u \preceq v$ implies $u \in \mathcal{T}$,

(iii) for every $u \in \mathcal{T}$, there exists a number $\nu_u \in \mathbb{N}$ such that if $\nu_u = 0$ then $v \succ u$ implies $v \notin \mathcal{T}$, otherwise $uj \in \mathcal{T}$ if and only if $1 \leq j \leq \nu_u$.

Notice that a rooted ordered tree $\mathcal{T}$ is completely defined by the sequence $(\nu_u, u \in \mathcal{U})$, which gives the number of children for every individual. To obtain a continuous time tree, we simply add the sequence of lifetimes.

Continuous time discrete trees. For a sequence $(l_u, u \in \mathcal{U})$ of nonnegative reals, let us define:

$$\forall u \in \mathcal{U},\ \alpha(u) = \sum_{v \prec u} l_v \quad \text{and} \quad \beta(u) = \sum_{v \preceq u} l_v = \alpha(u) + l_u, \tag{2.2}$$

with the convention $\alpha(\emptyset) = 0$. So to speak l_u stands for the lifetime of individual u while $\alpha(u)$ and $\beta(u)$ are its birth and death times. Let

$$\mathbb{U} = \mathcal{U} \times [0, +\infty). \tag{2.3}$$

Definition 2.2. *A continuous time rooted discrete tree (CT) is a subset $\mathbb{T}$ of $\mathbb{U}$ such that:*

(i) $(\emptyset, 0) \in \mathbb{T}$.

(ii) The projection of $\mathbb{T}$ on $\mathcal{U}$, $\mathcal{T}$, is a discrete rooted ordered tree,

(iii) There exists a sequence of nonnegative reals $(l_u, u \in \mathcal{U})$ such that for $u \in \mathcal{T}$, $(u, s) \in \mathbb{T}$ if and only if $\alpha(u) \leq s < \beta(u)$, where $\alpha(u)$ and $\beta(u)$ are defined by (2.2).

Let $\mathbb{T}$ be a CT. The set of individuals of $\mathbb{T}$ living at time t is denoted by V_t

$$V_t = \{u \in \mathcal{U} : (u, t) \in \mathbb{T}\} = \{u \in \mathcal{T} : \alpha(u) \leq t < \beta(u)\}. \tag{2.4}$$

The number of individuals alive at time t is $N_t = \mathrm{Card}(V_t)$. We denote by D_t the number of individuals which have died before time t:

$$D_t = \mathrm{Card}\{u \in \mathcal{T} : \beta(u) < t\}. \tag{2.5}$$

For $(u, s) \in \mathbb{T}$ and $t \leq s$, we introduce $u(t)$, the ancestor of u living at time t:

$$u(t) = v \quad \text{if} \quad (v \preceq u \quad \text{and} \quad (v, t) \in \mathbb{T}). \tag{2.6}$$

Eventually, for $(u, s) \in \mathbb{T}$, we define the shift of $\mathbb{T}$ at (u, s) by $\theta_{(u,s)}\mathbb{T} = \{(v, t) \in \mathbb{U} : (uv, s + t) \in \mathbb{T}\}$. Note that $\theta_{(u,s)}\mathbb{T}$ is still a CT.

Continuous time Galton-Watson trees.

Definition 2.3. *We say that a random CT is a continuous time Galton-Watson tree with offspring distribution $p = (p_k, k \in \mathbb{N})$ and exponential lifetime with mean $1/r$ if:*

(i) The sequence of the number of offspring, $(\nu_u, u \in \mathcal{U})$, is a sequence of independent random variables with common distribution p.

(ii) The sequence of lifetimes $(l_u, u \in \mathcal{U})$ is a sequence of independent exponential random variables with mean $1/r$.

(iii) The sequences $(\nu_u, u \in \mathcal{U})$ and $(l_u, u \in \mathcal{U})$ are independent.

We suppose that the offspring distribution p has finite second moment. We call

$$m = \sum_{k\geq 0} k\, p_k \quad \text{and} \quad \varsigma^2 = \sum_{k\geq 0} (k-m)^2\, p_k, \tag{2.7}$$

its expectation and variance. The offspring distribution is critical (resp. supercritical, resp. subcritical) if $m = 1$ (resp. $m > 1$, resp. $m < 1$). In this work, we mainly deal with the supercritical case.

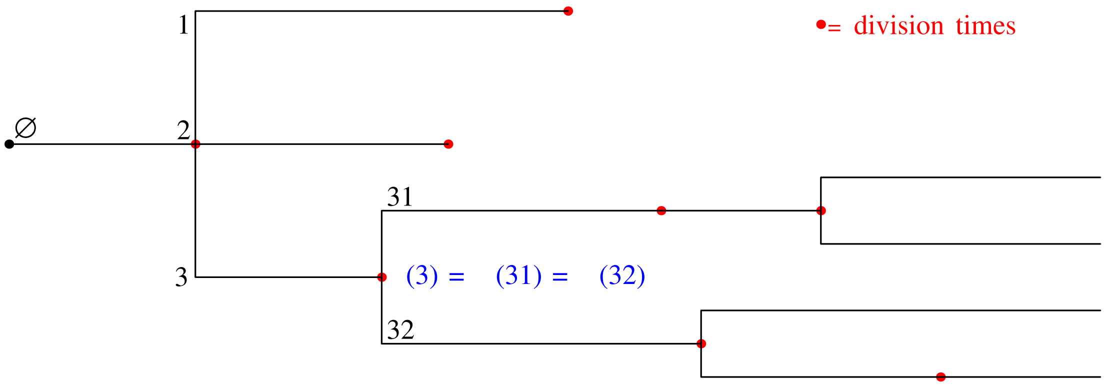

Figure 1: *Continuous time Galton-Watson tree.*

We end Section 2.1 with some estimates on N_t and D_t. To begin with, the following Lemma gives an equivalent for N_t.

Lemma 2.4. *For $t \in \mathbb{R}_+$, we have*

$$\mathbb{E}[N_t] = \mathrm{e}^{r(m-1)t}, \tag{2.8}$$

$$\mathbb{E}[N_t^2] = \begin{cases} \mathrm{e}^{r(m-1)t} + (\varsigma^2(m-1)^{-1} + m)(\mathrm{e}^{2r(m-1)t} - \mathrm{e}^{r(m-1)t}) & \textit{if } m \neq 1 \\ 1 + \varsigma^2 rt & \textit{if } m = 1. \end{cases} \tag{2.9}$$

If $m > 1$ there exists a nonnegative random variable W such that

$$\lim_{t\to+\infty} \frac{N_t}{\mathbb{E}[N_t]} = W \quad \textit{a.s and in } L^2, \tag{2.10}$$

and a.s. $\{W > 0\} = \{\forall t > 0,\ N_t > 0\}$.

Proof. The process $(N_t, t \geq 0)$ is a continuous time Markov branching process, so that the expectation and the variance of N_t are well-known, see [**?**] Chapter III Section 4. Almost sure convergence towards W is stated again in [**?**] Theorems 1 and 2 Chapter III Section 7. Finally, since the martingale $(N_t \,\mathrm{e}^{-r(m-1)t}, t \geq 0)$ is bounded in L^2, we obtain the L^2 convergence (*e.g.* Theorem 1.42 p. 11 of [**?**]). ■

We also give an equivalent for the asymptotic number of deaths D_t before t when $t \to +\infty$.

Lemma 2.5. *If $m > 1$, we have the following convergence a.s. and in L^2:*

$$\lim_{t\to+\infty} \frac{D_t}{\mathbb{E}[D_t]} = W, \tag{2.11}$$

with

$$\mathbb{E}[D_t] = (m-1)^{-1}(\mathrm{e}^{r(m-1)t} - 1) \tag{2.12}$$

and W defined by (2.10).

Proof. First remark that $(D_t, t \geq 0)$ is a counting process with intensity $rN_t\, dt$. We set $\Delta N_t = N_t - N_{t-}$ so that $dN_t = \Delta N_t \, dD_t$. To prove (2.11), it is sufficient to prove that $\mathrm{e}^{-r(m-1)t} I_t$ goes to 0 a.s. and in L^2, where $I_t = (m-1)D_t - N_t$. Since $I = (I_t, t \geq 0)$ satisfies the following stochastic equation driven by $(D_t, t \geq 0)$

$$dI_t = (m - 1 - \Delta N_t) dD_t, \tag{2.13}$$

we get that I is an L^2 martingale. We deduce that $d\langle I\rangle_t = \varsigma^2 r N_t \, dt$ and:

$$\mathbb{E}\big[I_t^2\big] = 1 + \mathbb{E}\big[\langle I\rangle_t\big] = 1 + \varsigma^2 r \int_0^t \mathrm{e}^{r(m-1)s} \, ds = 1 + \frac{\varsigma^2}{m-1}\big(\mathrm{e}^{r(m-1)t} - 1\big), \tag{2.14}$$

which implies the L^2 convergence of $\mathrm{e}^{-r(m-1)t} I_t$ to 0. Besides, $(\mathrm{e}^{-r(m-1)t} I_t,\, t \geq 0)$ is a supermartingale bounded in L^2 and hence the convergence also holds almost surely. ■

Example 1. Yule tree. The so-called Yule tree is a continuous time Galton-Watson tree with a deterministic offspring distribution: each individual of the population gives birth to 2 individuals that is $p_2 = 1$ (*i.e.* $p = \delta_2$, the Dirac mass at 2). The Yule tree is thus a binary tree, whose edges have independent exponential lengths with mean $1/r$. In that case, W is exponential with mean 1 (see *e.g.* [**?**] p. 112). We deduce from Lemma 2.4 that, for $t \in \mathbb{R}_+$,

$$\mathbb{E}[N_t] = \mathrm{e}^{rt} \quad \text{and} \quad \mathbb{E}[N_t^2] = 2\,\mathrm{e}^{2rt} - \mathrm{e}^{rt}. \tag{2.15}$$

2.2 Markov process indexed by the continuous time Galton-Watson tree

In this section, we define the Markov process $X_{\mathbb{T}} = (X_t^u, (u,t) \in \mathbb{T})$ indexed by the continuous time Galton-Watson tree $\mathbb{T}$ and with initial condition μ. Branching Markov processes have already been the object of an abundant literature (*e.g.* [**?**, **?**, **?**, **?**, **?**]). The process that we consider jumps at branching times (nonlocal branching property) but these jumps may be dependent.

Let $(E, \mathcal{E})$ be a Polish space. We denote by $\mathcal{P}(E)$ the set of probability measures on $(E, \mathcal{E})$.

Definition 2.6. *Let $X = (X_t, t \geq 0)$ be a càdlàg E-valued strong Markov process. Let $\tilde{F} = (F_j^{(k)}, 1 \leq j \leq k, k \in \mathbb{N}^*)$ be a family of measurable functions on $E \times [0,1]$. The continuous time branching Markov (CBM) process $X_{\mathbb{T}} = (X_t^u, (u,t) \in \mathbb{T})$ with offspring distribution p, exponential lifetimes with mean $1/m$, offspring position $\tilde{F}$, underlying motion X and starting distribution $\mu \in \mathcal{P}(E)$, is defined recursively as follows:*

(i) $\mathbb{T}$ *is a continuous time Galton-Watson tree with offspring distribution* p *and exponential lifetimes with mean* $1/m$*.*

(ii) Conditionally on $\mathbb{T}$*,* $X^{\emptyset} = (X^{\emptyset}_t, t \in [0, \beta(\emptyset)))$ *is distributed as* $(X_t, t \in [0, \beta(\emptyset)))$ *with* X_0 *distributed as* μ*.*

(iii) Conditionally on $\mathbb{T}$ *and* $X^{\emptyset}$*, the initial positions of the first generation offspring* $(X^u_{\alpha(u)}, 1 \leq u \leq \nu_{\emptyset})$ *are given by* $(F^{(\nu_{\emptyset})}_u(X^{\emptyset}_{\beta(\emptyset)-}, \Theta), 1 \leq u \leq \nu_{\emptyset})$ *where* Θ *is a uniform random variable on* $[0, 1]$*.*

(iv) Conditionally on $X^{\emptyset}$*,* $\nu_{\emptyset}$*,* $\beta_{\emptyset}$ *and* $(X^u_{\alpha(u)}, 1 \leq u \leq \nu_{\emptyset})$*, the tree-indexed Markov processes* $(X^{uv}_{\alpha(u)+t}, (v, t) \in \theta_{(u,\alpha(u))}\mathbb{T})$ *for* $1 \leq u \leq \nu_{\emptyset}$ *are independent and respectively distributed as* $X_{\mathbb{T}}$ *with starting distribution the Dirac mass at* $X^u_{\alpha(u)}$*.*

The law of $X_{\mathbb{T}}$ *is denoted by* $\mathbb{P}_\mu$*.*

We write $\mathbb{E}_\mu$ for the expectation with respect to $\mathbb{P}_\mu$, and when there is no ambiguity, we omit the subscript μ. If $\mu = \delta_x$, the Dirac mass at x, we write $\mathbb{P}_x$ and $\mathbb{E}_x$ for $\mathbb{P}_\mu$ and $\mathbb{E}_\mu$.

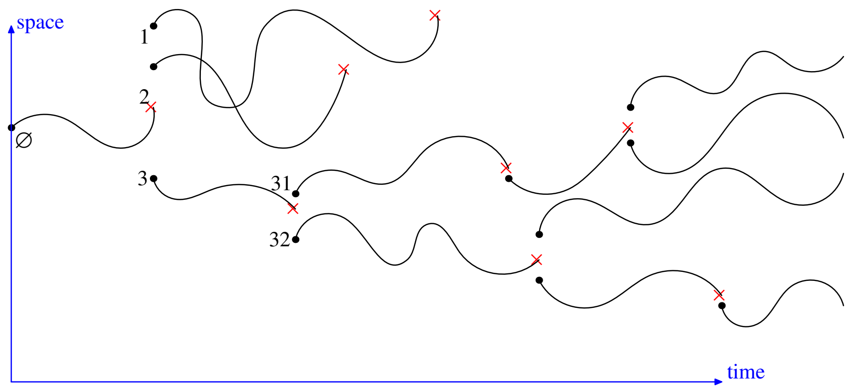

Figure 2: *Continuous time Markov process indexed by a Galton-Watson tree.*

2.3 Measure-valued description

Let $\mathcal{B}_b(E, \mathbb{R})$ be the set of real-valued measurable bounded functions on E and $\mathcal{M}_F(E)$ the set of finite measures on E embedded with the topology of weak convergence. For $\mu \in \mathcal{M}_F(E)$ and $f \in \mathcal{B}_b(E, \mathbb{R})$ we write $\langle \mu, f \rangle = \int_E f(x)\mu(dx)$.

We introduce the following measures to represent the population at time t:

$$\bar{Z}_t = \sum_{u \in V_t} \delta_{(u, X^u_t)}, \quad \text{and} \quad Z_t = \sum_{u \in V_t} \delta_{X^u_t}, \tag{2.16}$$

where V_t has been defined in (2.4). Note that $\langle Z_t, f \rangle = \sum_{u \in V_t} f(X^u_t)$. Since X is càdlàg, we get that the process $Z = (Z_t, t \geq 0)$ is a càdlàg measure-valued Markov process of $\mathbb{D}(\mathbb{R}_+, \mathcal{M}_F(E))$.

Following the works of Fournier and Méléard [?], we can describe the evolution of Z in terms of stochastic differential equations (SDE). Let $\rho(ds, du, dk, d\theta)$ be a Poisson point measure of intensity $r\, ds \otimes n(du) \otimes p(dk) \otimes d\theta$ where ds and $d\theta$ are Lebesgue measures on $\mathbb{R}_+$ and $[0, 1]$ respectively, $n(du)$

is the counting measure on $\mathcal{U}$ and $p(dk) = \sum_{i\in\mathbb{N}} p_i\delta_i(dk)$ is the offspring distribution. This measure ρ gives the information on the divisions. Let L be the infinitesimal generator of X. If $\mathcal{C}_b^{1,0}(\mathbb{R}_+ \times E, \mathbb{R})$ denotes the space of continuous bounded functions that are $\mathcal{C}^1$ in time with bounded derivatives, then for test functions $f : (t,x) \mapsto f_t(x)$ in $\mathcal{C}_b^{1,0}(\mathbb{R}_+ \times E, \mathbb{R})$ such that $\forall t \in \mathbb{R}_+, f_t \in D(L)$, we have

$$\begin{aligned}\langle Z_t, f_t\rangle =& f_0(X_0^{\emptyset}) + \int_0^t \int_{\mathcal{U}\times\mathbb{N}\times[0,1]} \mathbf{1}_{\{u\in V_{s-}\}} \left(\sum_{j=1}^k f_s(F_j^{(k)}(X_{s_-}^u,\theta)) - f_s(X_{s_-}^u)\right) \rho(ds,du,dk,d\theta)\\ &+ \int_0^t \int_{\mathbb{R}_+} (Lf(x) + \partial_s f(x))\, Z_s(dx)\, ds + W_t^f,\end{aligned} \tag{2.17}$$

where W_t^f is a martingale. Explicit expressions of this martingale and of the infinitesimal generator of $(Z_t, t \geq 0)$ can be obtained when L is precised.

Example 2. Splitted diffusions. The case when the Markov process X is a real diffusion ($E = \mathbb{R}$) is an interesting example. Let L be given by:

$$Lf(x) = b(x)f'(x) + \frac{\sigma^2(x)}{2} f''(x), \tag{2.18}$$

where we assume that b and σ are bounded and Lipschitz continuous. In this case, we can consider the following class of cylindrical functions from $\mathcal{M}_F(\mathbb{R})$ into $\mathbb{R}$ defined by $\phi_f(Z) = \phi(\langle Z, f\rangle)$ for $f \in \mathcal{C}_b^2(\mathbb{R},\mathbb{R})$ and $\phi \in \mathcal{C}_b^2(\mathbb{R})$ which is known to be convergence determining on $\mathcal{P}(\mathcal{M}_F(\mathbb{R}))$ (*e.g.* [**?**] Theorem 3.2.6). We can define the infinitesimal generator $\mathcal{L}$ of $(Z_t)_{t\geq 0}$ for these functions:

$$\mathcal{L}\phi_f(Z) = \mathcal{L}_1\phi_f(Z) + \mathcal{L}_2\phi_f(Z), \tag{2.19}$$

where $\mathcal{L}_1$ and $\mathcal{L}_2$ correspond to the branching and motion parts. Such decompositions were already used in Dawson [**?**] (Section 2.10) and in Roelly and Rouault [**?**] for instance. The generator $\mathcal{L}_1$ is defined by:

$$\mathcal{L}_1\phi_f(Z) =r \int_E \int_0^1 \sum_{k\in\mathbb{N}} \left(\phi\Big(\langle Z, f\rangle + \sum_{j=1}^k f(F_j^{(k)}(x,\theta)) - f(x)\Big) - \phi_f(Z)\right) p_k\, d\theta\, Z(dx), \tag{2.20}$$

with the convention that the sum over j is zero when $k = 0$. The generator $\mathcal{L}_2$ is given by:

$$\mathcal{L}_2\phi_f(Z) = \langle Z, Lf\rangle\phi'(\langle Z, f\rangle) + \langle Z, \sigma(x)f'^2(x)\rangle\phi''(\langle Z, f\rangle). \tag{2.21}$$

For a test function $f : (t,x) \mapsto f_t(x)$ in $\mathcal{C}_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}, \mathbb{R})$, the evolution of $(Z_t,\, t \geq 0)$ can then be described by the following SDE:

$$\begin{aligned}\langle Z_t, f_t\rangle =& f_0(X_0^{\emptyset}) + \int_0^t \int_{\mathcal{U}\times\mathbb{N}\times[0,1]} \mathbf{1}_{\{u\in V_{s-}\}} \left(\sum_{j=1}^k f_s(F_j^{(k)}(X_{s_-}^u,\theta)) - f_s(X_{s_-}^u)\right) \rho(ds,du,dk,d\theta)\\ &+ \int_0^t \int_{\mathbb{R}} (Lf_s(x) + \partial_s f_s(x))\, Z_s(dx)\, ds + \int_0^t \sum_{u\in V_s} \sqrt{2}\sigma(X_s^u)\partial_x f_s(X_s^u) dB_s^u.\end{aligned} \tag{2.22}$$

where $(B^u)_{u\in\mathcal{U}}$ a family of independent standard Brownian motions. In [**?**], such splitted diffusions are considered to describe a multi-level population. The cells, which correspond to the individuals in the present setting, undergo binary divisions, and contain a continuum of parasites that evolves as a Feller diffusion with drift $b(x) = (b - d)x$ and diffusion $\sigma(x) = 2\sigma^2 x$. At the branching time s for the individual u, each daughter inherits a random fraction of the value of the mother. The daughters $u1$ and $u2$ start respectively at $F_1^{(2)}(X^u_{s_-}, \theta) = G^{-1}(\theta)X^u_{s_-}$ and $F_2^{(2)}(X^u_{s_-}, \theta) = (1 - G^{-1}(\theta))X^u_{s_-}$, where G^{-1} is the generalized inverse of G, the cumulative distribution function of the random fraction. $\square$

3 The auxiliary Markov process and Many-To-One formulas

In this section, we are interested in the distribution of the path of an individual picked at random in the population at time t. By choosing uniformly among the individuals present at time t, we give a more important weight to branches where there have been more divisions and more children since the proportion of the corresponding offspring will be higher. Our pathwise approach generalizes [**?**] (discrete time) and [**?**] (continuous time Yule processes). As mentioned in the introduction, this size bias has already been observed by [**?**, **?**, **?**] for the tree structure when considering marginal distributions and by [**?**, **?**] for local branching Markov process.

In Section 3.1, we introduce an auxiliary Markov process which gives the distribution of an individual picked at random among all the individuals of all possible trees. This appears in (3.1) for an individual living at time t. We extend this relation to an individual picked at random in the whole tree (Section 3.2) and to a pair of individuals picked at random (Section 3.3).

3.1 Auxiliary process and Many-To-One formula at fixed time

We focus on the law of an individual picked at random and show that it is given by an auxiliary Markov process. This auxiliary Markov process $\widetilde{Y} = (Y, \Lambda)$ has two components. The component Y describes the motion on the space E. It behaves like X and has additional jumps which occur at rate rm. At these additional jump times, the new position is given by $F_J^{(H)}(., \Theta)$ where the "offspring" number H has the size biased distribution of the offspring distribution p, J is uniform on $\{1, \dots, H\}$ and Θ is uniform on $[0, 1]$. The component Λ records the additional jump times as well as the "offspring" numbers. For the definition of Λ, we shall consider the logarithm of the offspring number as this is the quantity that is involved in the Girsanov formulas.

By convention for a function f defined on an interval I, we set $f_J = (f(t), t \in J)$ for any $J \subset I$.

Definition 3.1. *Let $X_{\mathbb{T}}$ be as in definition 2.6 with starting distribution $\mu \in \mathcal{P}(E)$. The corresponding auxiliary process $\widetilde{Y} = (Y, \Lambda)$, with $Y = (Y_t, t \geq 0)$ and $\Lambda = (\Lambda_t, t \geq 0)$, is an $E \times \mathbb{R}$-valued càdlàg Markov process. The process (Y, Λ) and $\mathcal{I} = (I_k, k \in \mathbb{N}^*)$, a sequence of random variables, are defined as follows:*

(i) Λ is a compound Poisson process: $\Lambda_t = \sum_{k=1}^{S_t} \log(H_k)$, where $S = (S_t, t \geq 0)$ is a Poisson process with intensity rm, and $(H_k, k \in \mathbb{N}^*)$ are independent random variables independent of S and with common distribution the size biased distribution of p, $(hp_h/m,\ h \in \mathbb{N}^*)$.

(ii) Conditionally on Λ, $(I_k, k \in \mathbb{N}^*)$ are independent random variables and I_k is uniform on $\{1, \dots, H_k\}$.

(iii) Conditionally on $(\Lambda, \mathcal{I})$, the process $Y_{[0,\tau_1)}$, where $\tau_1 = \inf\{t \geq 0; S_t \neq S_0\}$, is distributed as $(X_t, t \in [0, \tau_1))$ with initial distribution μ and X independent of τ_1.

(iv) Conditionally on $(\Lambda, \mathcal{I}, Y_{[0,\tau_1)})$, Y_{τ_1} *is distributed as* $F_{I_1}^{(H_1)}(y, \Theta)$, *where* Θ *is an independent uniform random variable on* $[0,1]$.

(v) The distribution of $(Y_{\tau_1+t}, t \geq 0)$ *conditionally on* $(\Lambda, \mathcal{I}, Y_{[0,\tau_1]})$ *is equal to the distribution of* Y *conditionally on* $(\Lambda_{\tau_1+t} - \Lambda_{\tau_1}, t \geq 0)$ *and* $(I_{1+k}, k \in \mathbb{N}^*)$, *and started at* Y_{τ_1}.

We write $\mathbb{E}_\mu$ when we take the expection with respect to $(Y, \Lambda, \mathcal{I})$ and the starting measure is μ for the Y component. We also use the same convention as those described just after definition 2.6.

For $u \in \mathcal{U}$, we extend the definition of X_t^u for $t \in [0, \alpha(u))$ as follows: $X_t^u = X_t^{u(t)}$, where $u(t)$, defined by (2.6), denotes the ancestor of u living at time t. We define $\widetilde{X}_t^u = (X_t^u, \Lambda_t^u)$ where for $(u,s) \in \mathbb{T}$, for $t \leq s$

$$\Lambda_t^u = \sum_{v \prec u(t)} \log(\nu_v).$$

Notice that, for fixed u, the process $(\Lambda_t^u, t \in [0, \beta(u)))$ is a compound Poisson process with rate r for the underlying Poisson process and increments distributed as $\log(\nu)$ with ν distributed as p. It is finite if $u \in \mathcal{T}$. It allows to recover the birth time and family size of the ancestors of $u \in \mathcal{T}$.

The formula (3.1) in the next Proposition is similar to the so-called Many-to-One Theorem of Hardy and Harris [**?**] (Section 8.2) that enables expectation of sums over particles in the branching process to be calculated in terms of an expectation of an auxiliary process. Notice that in our setting an individual may have no offspring with positive probability (if $p_0 > 0$) which is not the case in [**?**].

Proposition 3.2 (Many-To-One formula at fixed time). *For* $t \geq 0$ *and for any nonnegative measurable function* $f \in \mathcal{B}(\mathbb{D}([0,t], E \times \mathbb{R}), \mathbb{R}_+)$ *and* $t \geq 0$, *we have:*

$$\frac{\mathbb{E}_\mu\big[\sum_{u \in V_t} f(\widetilde{X}_{[0,t]}^u)\big]}{\mathbb{E}[N_t]} = \mathbb{E}_\mu[f(\widetilde{Y}_{[0,t]})]. \tag{3.1}$$

Remark 3.3.

- Notice that we can not recover S from Λ unless there is no "offspring" number equal to 1, that is $p_1 = 0$. But this can always be achieved by changing the value of the jump rate r and adding the jumps related to $F_1^{(1)}$ to the process X. With a slight abuse, we will consider that the jumps of Λ are the jumps of S.
- For $m > 1$, a typical individual living at time t has prolific ancestors with shorter lives. For $m < 1$, a typical individual living at time t has still prolific ancestors but with longer lives.
- If births are local (*i.e.* for all $j \leq k$, $F_j^{(k)}(x, \theta) = x$), then Y is distributed as X.

Proof of Proposition 3.2. Let Λ be a compound Poisson process defined by (i). Using Girsanov theorem on compound Poisson process, we obtain for any nonnegative measurable function g that:

$$\mathbb{E}[g(\Lambda_{[0,t]})] = \mathbb{E}\left[g(\Lambda'_{[0,t]})\, \mathrm{e}^{-r(m-1)t + \Lambda'_t}\right], \tag{3.2}$$

where the process Λ' is a compound process with rate r for the underlying Poisson process and increments distributed as $\log(\nu)$ with ν distributed as p. Indeed, $g(\Lambda_{[0,t]})$ is a function of t, of the jump times $\tau_q = \inf\{t \geq 0; S_t = q\}$ and of jump sizes ν_q of Λ:

$$g(\Lambda_{[0,t]}) = \sum_{q=0}^{+\infty} G_q(t, \tau_1, \ldots, \tau_q, \nu_1, \ldots, \nu_q) \mathbf{1}_{\{\sum_{i=1}^q \tau_i \leq t < \sum_{i=1}^{q+1} \tau_i\}},$$

for some functions $(G_q, q \in \mathbb{N})$. We deduce that:

$$
\begin{aligned}
&\mathbb{E}[g(\Lambda_{[0,t]})] \\
=&\sum_{q=0}^{+\infty} \int_{\mathbb{R}_+^q} \sum_{h_1,\dots,h_q} (rm)^q \, \mathrm{e}^{-rmt} \, G(t_1,\dots,t_q,\log(h_1),\dots\log(h_q)) \prod_{i=1}^q \frac{p_{h_i} h_i}{m} \, dt_1 \dots dt_q \\
=&\sum_{q=0}^{+\infty} \int_{\mathbb{R}_+^q} \sum_{h_1,\dots,h_q} r^q \, \mathrm{e}^{-rmt} \, G(t_1,\dots,t_q,\log(h_1),\dots\log(h_q)) \, \mathrm{e}^{-r(m-1)t+\sum_{i=1}^q \log(h_i)} \prod_{i=1}^q p_{h_i} \, \mathrm{e}^{-rt} \; dt_1 \dots dt_q \\
=&\mathbb{E}\left[g(\Lambda'_{[0,t]}) \, \mathrm{e}^{-r(m-1)t+\Lambda'_t}\right].
\end{aligned}
$$

Recall that $(S_t, t \geq 0)$ is the underlying Poisson process of Λ. In particular, we have for $q \in \mathbb{N}$, $u \in \mathcal{U}$ such that $|u| = q$,

$$
\mathbb{E}[g(\Lambda_{[0,t]}) \mathbf{1}_{\{S_t=q\}}] = \mathbb{E}\left[g(\Lambda^u_{[0,t]}) \, \mathrm{e}^{-r(m-1)t+\Lambda^u_t} \, \mathbf{1}_{\{\Lambda^u_t > -\infty, \; \alpha(u) \leq t < \beta(u)\}}\right]. \tag{3.3}
$$

Let $q \in \mathbb{N}^*$. By construction, conditionally on $\{\Lambda_{[0,t]} = \lambda_{[0,t]}\}$, $\{S_t = q\}$, $\{(I_1,\dots,I_q) = u\}$, $Y_{[0,t]}$ is distributed as $X^u_{[0,t]}$ conditionally on $\{\Lambda^u_{[0,t]} = \lambda_{[0,t]}\}$. This holds also for $q = 0$ with the convention that $(I_1,\dots,I_q) = \emptyset$. Therefore, we have for any nonnegative measurable functions g and f,

$$
\begin{aligned}
\mathbb{E}_\mu[g(\Lambda_{[0,t]}) f(Y_{[0,t]})] &= \sum_{u\in\mathcal{U}} \sum_{q\in\mathbb{N}} \mathbf{1}_{\{|u|=q\}} \mathbb{E}_\mu[g(\Lambda_{[0,t]}) f(Y_{[0,t]}) \mathbf{1}_{\{(I_1,\dots,I_q)=u\}} \mathbf{1}_{\{S_t=q\}}] \\
&= \sum_{u\in\mathcal{U}} \sum_{q\in\mathbb{N}} \mathbf{1}_{\{|u|=q\}} \mathbb{E}[g(\Lambda_{[0,t]}) \mathbb{E}_\mu[f(X^u_{[0,t]}) | \Lambda^u_{[0,t]}]_{|\Lambda^u_{[0,t]}=\Lambda_{[0,t]}} \mathbf{1}_{\{(I_1,\dots,I_q)=u\}} \mathbf{1}_{\{S_t=q\}}] \\
&= \sum_{u\in\mathcal{U}} \sum_{q\in\mathbb{N}} \mathbf{1}_{\{|u|=q\}} \mathbb{E}[g(\Lambda_{[0,t]}) \mathbb{E}_\mu[f(X^u_{[0,t]}) | \Lambda^u_{[0,t]}]_{|\Lambda^u_{[0,t]}=\Lambda_{[0,t]}} \, \mathrm{e}^{-\Lambda_t} \, \mathbf{1}_{\{S_t=q\}}] \\
&= \sum_{u\in\mathcal{U}} \sum_{q\in\mathbb{N}} \mathbf{1}_{\{|u|=q\}} \mathbb{E}_\mu[g(\Lambda^u_{[0,t]}) \mathbb{E}_\mu[f(X^u_{[0,t]}) | \Lambda^u_{[0,t]}] \, \mathrm{e}^{-r(m-1)t} \, \mathbf{1}_{\{\Lambda^u_t > -\infty, \; \alpha(u) \leq t < \beta(u)\}}] \\
&= \sum_{u\in\mathcal{U}} \mathbb{E}_\mu[g(\Lambda^u_{[0,t]}) f(X^u_{[0,t]}) \, \mathrm{e}^{-r(m-1)t} \, \mathbf{1}_{\{u\in V_t\}}],
\end{aligned}
$$

where we used (3.3) for the fourth equation and $\{\Lambda^u_t > -\infty, \; \alpha(u) \leq t < \beta(u)\} = \{u \in V_t\}$ for the last. Then, we use (2.8) and a monotone class argument to conclude. ■

3.2 Many-to-Ones formulas over the whole tree

In this section, we generalize identity (3.1) on the link between the tree indexed process $X_{\mathbb{T}}$ and the auxiliary Markov process Y by considering sums over the whole tree.

Let us consider the space $\mathcal{D}$ of nonnegative measurable functions $f \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{D}(\mathbb{R}_+, E \times \mathbb{R}), \mathbb{R}_+)$ such that $f(t, y) = f(t, z)$ as soon as $y_{[0,t)} = z_{[0,t)}$. By convention, if y is defined at least on $[0, t)$, we will write $f(t, y_{[0,t)})$ for $f(t, z)$ where z is any function such that $z_{[0,t)} = y_{[0,t)}$.

Proposition 3.4 (Many-To-One formula over the whole tree)**.** *For all nonnegative measurable function f of $\mathcal{D}$, we have:*

$$
\mathbb{E}_\mu\left[\sum_{u\in\mathcal{T}} f(\beta(u), \widetilde{X}^u_{[0,\beta(u))})\right] = r \int_0^{+\infty} ds \;\, \mathrm{e}^{r(m-1)s} \, \mathbb{E}_\mu\left[f(s, \widetilde{Y}_{[0,s)})\right]. \tag{3.4}
$$

By convention for two functions f, g defined respectively on two intervals I_f, I_g, for $[a,b) \subset I_f$ and $[c,d) \subset I_g$, we define the concatenation $[f_{[a,b)}; g_{[c,d)}] = h_J$ where $J = [a, b + (d-c))$,

$$h(t) = \begin{cases} f(t) & \text{if } t \in [a,b) \\ g(c + (t-b)) & \text{if } t \in [b, d-c+b). \end{cases}$$

Proof. We first notice that if τ is an exponential random variable with mean $1/r$ ($r > 0$), then we have, for any nonnegative measurable function g,

$$\mathbb{E}\left[r \int_0^\tau g(t)dt\right] = \mathbb{E}[g(\tau)]. \tag{3.5}$$

Besides, we have

$$\mathbb{E}_\mu\left[\mathbf{1}_{\{u\in\mathcal{T}\}} f(\beta(u), \widetilde{X}^u_{[0,\beta(u))})\right] = \mathbb{E}_\mu\left[\mathbf{1}_{\{u\in\mathcal{T}\}} f(\beta(u), [\widetilde{X}^u_{[0,\alpha(u))}; \widetilde{X}_{[0,\beta(u)-\alpha(u))}])\right],$$

where conditionally on $\widetilde{X}^u_{[0,\alpha(u))}$, $\beta(u)$, $\{u \in \mathcal{T}\}$, $\widetilde{X} = (X, c)$ with X of distribution $\mathbb{P}_{X^u_{\alpha(u)}}$ and c the constant process equal to $\Lambda^u_{\alpha(u)}$. Notice that we have chosen $\widetilde{X}$ independent of $\beta(u)$. Thus, conditioning with respect to $[\widetilde{X}^u_{[0,\alpha(u))}; \widetilde{X}_{[0,+\infty)}]$, $\{u \in \mathcal{T}\}$ and using (3.5), we get

$$\mathbb{E}_\mu\left[\mathbf{1}_{\{u\in\mathcal{T}\}} f(\beta(u), \widetilde{X}^u_{[0,\beta(u))})\right] = r\mathbb{E}\left[\mathbf{1}_{\{u\in\mathcal{T}\}} \int_0^{\beta(u)-\alpha(u)} ds\ f(\alpha(u)+s, \widetilde{X}^u_{[0,\alpha(u)+s)})\right].$$

We deduce:

$$\begin{aligned}\mathbb{E}_\mu\left[\mathbf{1}_{\{u\in\mathcal{T}\}} f(\beta(u), \widetilde{X}^u_{[0,\beta(u))})\right] &= r\mathbb{E}\left[\mathbf{1}_{\{u\in\mathcal{T}\}} \int_{\alpha(u)}^{\beta(u)} ds\ f(s, \widetilde{X}^u_{[0,s)})\right] \\ &= r\int_0^{+\infty} ds\ \mathbb{E}\left[\mathbf{1}_{\{u\in V_s\}} f(s, \widetilde{X}^u_{[0,s)})\right],\end{aligned}$$

where we used (3.5) for the first equality. Using Proposition 3.2, we get

$$\begin{aligned}\mathbb{E}_\mu\left[\sum_{u\in\mathcal{T}} f(\beta(u), \widetilde{X}^u_{[0,\beta(u))})\right] &= r\int_0^{+\infty} ds\ \mathbb{E}_\mu\left[\sum_{u\in V_s} f(s, \widetilde{X}^u_{[0,s)})\right] \\ &= r\int_0^{+\infty} ds\ \mathrm{e}^{sr(m-1)}\,\mathbb{E}_\mu\left[f(s, \widetilde{Y}_{[0,s)})\right].\end{aligned}$$

■

The equality (3.4) means that adding the contributions over all the individuals in the Galton-Watson tree corresponds (at least for the first moment) to integrate the contribution of the auxiliary process over time with an exponential weight $\mathrm{e}^{r(m-1)t}$ which is the average number of living individuals at time t. Notice the weight is increasing if the Galton-Watson tree is supercritical and decreasing if it is subcritical. The left hand side of (3.4) corresponds heuristically to picking an individual uniformly among all the individuals of all the possible trees.

Remark 3.5. We shall give two alternative formulas for (3.4).

- We deduce from (3.4) that, for all nonnegative measurable function f,

$$\mathbb{E}_\mu\left[\sum_{u\in\mathcal{T}} f(\beta(u), \widetilde{X}^u_{[0,\beta(u))})\right] = \mathbb{E}_\mu\left[f(\tau, \widetilde{Y}_{[0,\tau)})\, \mathrm{e}^{rm\tau}\right], \tag{3.6}$$

  where τ is an independent exponential random variable of mean $1/r$. Thus, the right hand side of equation (3.4) can be read as the expectation of a functional of the process $\widetilde{Y}$ up to an independent exponential time τ of mean $1/r$, with a weight $\mathrm{e}^{rm\tau}$.

- Let $\tau_q = \inf\{t \geq 0; S_t = q\}$ the time of the q-th jump for the compound Poisson process Λ. Using (3.5), it is easy to check that, for any nonnegative measurable function g,

$$\frac{1}{m}\sum_{q\geq 1}\mathbb{E}_\mu[g(\widetilde{Y}_{[0,\tau_q)}, \tau_q)] = r\int_0^{+\infty} \mathbb{E}_\mu[g(\widetilde{Y}_{[0,s)}, s)]\, ds.$$

  Therefore, we deduce from (3.4) that, for all nonnegative measurable function f,

$$\mathbb{E}_\mu\left[\sum_{u\in\mathcal{T}} f(\beta(u), \widetilde{X}^u_{[0,\beta(u))})\right] = \frac{1}{m}\sum_{q\geq 1}\mathbb{E}_\mu\left[f(\tau_q, \widetilde{Y}_{[0,\tau_q)})\, \mathrm{e}^{r(m-1)\tau_q}\right]. \tag{3.7}$$

  This formula emphasizes that the jumps of the auxiliary process correspond to death times in the tree.

3.3 Identities for forks

In order to compute second moments, we shall need the distribution of two individuals picked at random in the whole population and which are not in the same lineage. As in the Many-To-One formula, it will involve the auxiliary process.

First, we define the following sets of forks:

$$\mathcal{FU} = \{(u,v)\in\mathcal{U}^2 : |u\wedge v| < \min(|u|,|v|)\} \quad \text{and} \quad \mathcal{FT} = \mathcal{FU}\cap\mathcal{T}^2. \tag{3.8}$$

Let $\widetilde{J}_2$ be the operator defined for all nonnegative measurable function f from $(E\times\mathbb{R})^2$ to $\mathbb{R}$ by:

$$\widetilde{J}_2 f(x,\lambda) = \int_0^1 \sum_{\substack{(a,b)\in(\mathbb{N}^*)^2\\ a\neq b}} \sum_{k\geq\max(a,b)} p_k\, f\Big(F_a^{(k)}(x,\theta), \lambda+\log(k), F_b^{(k)}(x,\theta), \lambda+\log(k)\Big)\, d\theta. \tag{3.9}$$

Informally, the functional $\widetilde{J}_2$ describes the starting positions of two siblings. Notice that we have

$$\widetilde{J}_2 f(x,\lambda) = m\int_0^1 \mathbb{E}\big[(H-1) f(F_I^{(H)}(x,\theta), \lambda+\log(H), F_K^{(H)}(x,\theta), \lambda+\log(H))\big]\, d\theta, \tag{3.10}$$

where H has the size-biased offspring distribution, and conditionally on H, (I,K) is distributed as a drawing without replacement among the integers $\{1,\ldots,H\}$.

For measurable real functions f and g on $E\times\mathbb{R}$, we denote by $f\otimes g$ the real measurable function on $(E\times\mathbb{R})^2$ defined by: $(f\otimes g)(\widetilde{x},\widetilde{y}) = f(\widetilde{x})g(\widetilde{y})$ for $\widetilde{x},\widetilde{y}\in E\times\mathbb{R}$.

Proposition 3.6 (Many-To-One formula for forks over the whole tree)**.** *For all nonnegative measurable functions $\varphi, \psi \in \mathcal{D}$, we have:*

$$\mathbb{E}_\mu\Big[\sum_{(u,v)\in\mathcal{FT}} \varphi(\beta(u), \widetilde{X}^u_{[0,\beta(u))})\psi(\beta(v), \widetilde{X}^v_{[0,\beta(v))})\Big]$$
$$= \mathbb{E}_\mu\Big[\mathrm{e}^{rm\tau}\, \widetilde{J}_2\Big(\mathbb{E}'_{\cdot}\big[\varphi\big(t+\tau',\, [\widetilde{y}_{[0,t)}; \widetilde{Y}'_{[0,\tau')}]\big)\,\mathrm{e}^{rm\tau'}\,\big]_{|t=\tau, \widetilde{y}=\widetilde{Y}}$$
$$\otimes\, \mathbb{E}'_{\cdot}\big[\psi\big(t+\tau',\, [\widetilde{y}_{[0,t)}; \widetilde{Y}'_{[0,\tau')}]\,\mathrm{e}^{rm\tau'}\,\big]_{|t=\tau, \widetilde{y}=\widetilde{Y}}\Big)(\widetilde{Y}_{\tau-})\Big], \quad (3.11)$$

where, under $\mathbb{E}_\mu$, τ is exponential with mean $1/r$ independent of $\widetilde{Y}$, and, under $\mathbb{E}'_{x,\lambda}$, $(\widetilde{Y}', \tau')$ is distributed as $((Y, \Lambda+\lambda), \tau)$ under $\mathbb{E}_x$.

Proof. Notice that $\{(u, v) \in \mathcal{FU}\}$ is equal to $\{\exists(w, \tilde{u}, \tilde{v}) \in \mathcal{U}^3, \exists(a, b) \in (\mathbb{N}^*)^2, a \neq b, u = wa\tilde{u}, v = wb\tilde{v}\}$. Let A be the l.h.s. of (3.11). We have:

$$A = \sum_{w\in\mathcal{U}} \sum_{\substack{a,b\in\mathbb{N}^*\\ a\neq b}} \sum_{\tilde{u},\tilde{v}\in\mathcal{U}} \mathbb{E}_\mu\Big[\varphi\big(\beta(w) + (\beta(wa\tilde{u}) - \beta(w)), [\widetilde{X}^w_{[0,\beta(w))}; \widetilde{X}^{wa\tilde{u}}_{[\beta(w),\beta(wa\tilde{u}))}]\big)\mathbf{1}_{\{wa\tilde{u}\in\mathcal{T}\}}$$
$$\times\, \psi\big(\beta(w) + (\beta(wb\tilde{v}) - \beta(w)),\, [\widetilde{X}^w_{[0,\beta(w))}; \widetilde{X}^{wb\tilde{v}}_{[\beta(w),\beta(wb\tilde{v}))}]\big)\mathbf{1}_{\{wb\tilde{v}\in\mathcal{T}\}}\Big].$$

Using the strong Markov property at time $\beta(w)$, the conditional independence between descendants and Proposition 3.4, we get:

$$A = \sum_{w\in\mathcal{U}} \sum_{\substack{a,b\in\mathbb{N}^*\\ a\neq b}} \mathbb{E}_\mu\Big[\mathbb{E}'_{\widetilde{X}^{wa}_{\alpha(wa)}}\Big[\varphi\big(t+\tau',\, [\widetilde{x}_{[0,t)};\, \widetilde{Y}_{[0,\tau')}]\big)\,\mathrm{e}^{rm\tau'}\,\Big]_{|t=\beta(w), \widetilde{x}=\widetilde{X}^w}\mathbf{1}_{\{wa\in\mathcal{T}\}}$$
$$\times\, \mathbb{E}'_{\widetilde{X}^{wb}_{\alpha(wb)}}\Big[\psi\big(t+\tau',\, [\widetilde{x}_{[0,t)};\, \widetilde{Y}_{[0,\tau')}]\big)\,\mathrm{e}^{rm\tau'}\,\Big]_{|t=\beta(w), \widetilde{x}=\widetilde{X}^w}\mathbf{1}_{\{wb\in\mathcal{T}\}}\Big], \quad (3.12)$$

where under $\mathbb{E}'_{x,\lambda}$, $(\widetilde{Y}', \tau')$ is distributed as $((Y, \Lambda+\lambda), \tau)$ under $\mathbb{E}_x$. As $\{wa, wb \in \mathcal{T}\} = \{w \in \mathcal{T}\} \cap \{\max\{a, b\} \leq \nu_w\}$ we have:

$$A = \sum_{w\in\mathcal{U}} \mathbb{E}_\mu\Big[\mathbf{1}_{\{w\in\mathcal{T}\}}\widetilde{J}_2\Big(\mathbb{E}'_{\cdot}\big[\varphi\big(t+\tau',\, [\widetilde{x}_{[0,t)}; \widetilde{Y}'_{[0,\tau')}]\big)\,\mathrm{e}^{rm\tau'}\,\big]_{|t=\beta(w), \widetilde{x}=\widetilde{X}^w}$$
$$\otimes\, \mathbb{E}'_{\cdot}\big[\psi\big(t+\tau',\, [\widetilde{x}_{[0,t)}; \widetilde{Y}'_{[0,\tau')}]\,\mathrm{e}^{rm\tau'}\,\big]_{|t=\beta(w), \widetilde{x}=\widetilde{X}^w}\Big)(\widetilde{X}^w_{\beta(w)_-})\Big], \quad (3.13)$$

with $\widetilde{J}_2$ defined by (3.9). The function under the expectation in (3.13) depends on $\beta(w)$ and $\widetilde{X}^w_{[0,\beta(w))}$. Equality (3.6) then gives the result. ∎

We shall give a version of Proposition 3.6, when the functions of the path depend only on the terminal value of the path. We shall define J_2 a simpler version of $\widetilde{J}_2$ (see definition (3.10)) acting only on the spatial motion: for all nonnegative measurable functions f from E^2 to $\mathbb{R}$,

$$J_2 f(x) = m \int_0^1 \mathbb{E}\big[(H-1) f(F^{(H)}_I(x, \theta), F^{(H)}_K(x, \theta))\big]\, d\theta, \quad (3.14)$$

where (H, I, K) are as in (3.10).

The following Corollary is a direct consequence of Proposition 3.6 and the fact that Y is càdlàg.

Corollary 3.7 (Many-To-One formula for forks over the whole tree)**.** *Let $(Q_t, t \geq 0)$ be the transition semi-group of Y. For all nonnegative measurable functions $f, g \in \mathcal{D}$, we have:*

$$\mathbb{E}_\mu\Big[\sum_{(u,v)\in\mathcal{FT}} f(\beta(u), X^u_{\beta(u)-}) g(\beta(v), X^v_{\beta(v)-})\Big] \\ = r^3 \int_{[0,\infty)^3} \mathrm{e}^{r(m-1)(s+t+t')}\, ds\, dt\, dt' \; \mu Q_s \left(J_2(Q_t f_{t+s} \otimes Q_{t'} g_{t'+s})\right), \quad (3.15)$$

where $f_t(x) = f(t, x)$ and $g_t(x) = g(t, x)$ for $t \geq 0$ and $x \in E$.

We can also derive a Many-To-One formula for forks at fixed time.

Proposition 3.8 (Many-To-One formula for forks at fixed time)**.** *Let $t \in \mathbb{R}_+$, and φ, ψ two nonnegative measurable functions on $\mathbb{D}([0, t], E)$. We have:*

$$\mathbb{E}_\mu\Big[\sum_{\substack{(u,v)\in V_t^2\\ u\neq v}} \varphi\big(\widetilde{X}^u_{[0,t]}\big)\, \psi\big(\widetilde{X}^v_{[0,t]}\big)\Big] = r\,\mathrm{e}^{2r(m-1)t} \int_0^t \mathrm{e}^{-r(m-1)a}\, da \\ \mathbb{E}_\mu\Big[\widetilde{J}_2\Big(\mathbb{E}'_{\cdot}\big[\varphi\big([\widetilde{y}_{[0,a)}; \widetilde{Y}'_{[0,t-a]}]\big)\big]_{|\widetilde{y}=\widetilde{Y}} \otimes \mathbb{E}'_{\cdot}\big[\psi\big([\widetilde{y}_{[0,a)}; \widetilde{Y}'_{[0,t-a]}]\big]_{|\widetilde{y}=\widetilde{Y}}\Big)(\widetilde{Y}_a)\Big], \quad (3.16)$$

where, under $\mathbb{E}'_{x,\lambda}$, $\widetilde{Y}'$ is distributed as $(Y, \Lambda + \lambda)$ under $\mathbb{E}_x$.

The l.h.s. of (3.16) heuristically corresponds to picking a pair of individuals uniformly from the population at time t for all possible trees. As in (3.4), we have in the r.h.s. of (3.16) an exponential weight $\mathrm{e}^{2r(m-1)t}$ which corresponds to the average number of pairs of individuals that can be picked at time t. The distribution of the paths associated with a random pair is described by the law of forks constituted of independent portions of the auxiliary process $\widetilde{Y}$ and splitted at a time $a \in [0, t]$. Notice that (3.16) indicates that the fork splits at an exponential random time with mean $1/r(m-1)$, conditioned to be less than t.

Proof of Proposition 3.8. The proof is similar to the proof of Proposition 3.6 except that we use Proposition 3.2 instead of Proposition 3.4 to obtain an analogue of (3.12). ■

4 Law of large numbers

In this Section, we are interested in averages over the population living at time t for large t. When the Galton-Watson tree is not supercritical we have almost sure extinction, and thus we assume here that $m > 1$.

4.1 Results and comments

Notice that $N_t = 0$ implies $Z_t = 0$ and by convention we set $Z_t/N_t = 0$ in this case. For $t \in \mathbb{R}_+$ and f a real function defined on E, we derive laws of large numbers for

$$\frac{\langle Z_t, f\rangle}{N_t} = \frac{\sum_{u\in V_t} f(X^u_t)}{N_t} \quad \text{and} \quad \frac{\langle Z_t, f\rangle}{\mathbb{E}[N_t]} = \frac{\sum_{u\in V_t} f(X^u_t)}{\mathbb{E}[N_t]}, \quad (4.1)$$

provided the auxiliary process introduced in the previous Section satisfies some ergodic conditions.

Let $(Q_t, t \geq 0)$ be the semigroup of the auxiliary process Y from definition 3.1:

$$\mathbb{E}_\mu[f(Y_t)] = \mu Q_t f \tag{4.2}$$

for all $\mu \in \mathcal{P}(E)$ and f nonnegative. Recall the operator J_2 defined in (3.14).

We shall consider the following ergodicity and integrability assumptions on f, a real measurable function defined on E, and $\mu \in \mathcal{P}(E)$.

(H1) There exists a nonnegative finite measurable function g such that $Q_t|f|(x) \leq g(x)$ for all $t \geq 0$ and $x \in E$.

(H2) There exists $\pi \in \mathcal{P}(E)$, such that $\langle \pi, |f| \rangle < +\infty$ and for all $x \in E$, $\lim_{t\to+\infty} Q_t f(x) = \langle \pi, f \rangle$.

(H3) There exists $\alpha < r(m-1)$ and $c_1 > 0$ such that $\mu Q_t f^2 \leq c_1 \,\mathrm{e}^{\alpha t}$ for every $t \geq 0$.

(H4) There exists $\alpha < r(m-1)$ and $c_2 > 0$ such that $\mu Q_t J_2(g \otimes g) \leq c_2 \,\mathrm{e}^{\alpha t}$ for every $t \geq 0$, with g defined in (H2).

Notice that in (H3-4), the constants α, c_1 and c_2 may depend on f and μ.

Remark 4.1. When the auxiliary process Y is ergodic (*i.e.* Y converges in distribution to $\pi \in \mathcal{P}(E)$), the class of continuous bounded functions satisfies (H1-4) with g constant and $\alpha = 0$. In some applications, one may have to consider polynomial growing functions. This is why we shall consider hypothesis (H1-4) instead of the ergodic property in Theorem 4.2 or in Proposition 4.3

The next Theorem states the law of large numbers: the asymptotic empirical measure is distributed as the stationary distribution π of Y.

Theorem 4.2. *For any $\mu \in \mathcal{P}(E)$ and f a real measurable function defined on E satisfying (H1-4), we have*

$$\lim_{t\to+\infty} \frac{\langle Z_t, f\rangle}{\mathbb{E}[N_t]} = \langle \pi, f\rangle W \quad \text{in } L^2(\mathbb{P}_\mu), \tag{4.3}$$

$$\lim_{t\to+\infty} \frac{\langle Z_t, f\rangle}{N_t} = \langle \pi, f\rangle \mathbf{1}_{\{W\neq 0\}} \quad \text{in } \mathbb{P}_\mu\text{-probability}, \tag{4.4}$$

with W defined by (2.10) and π defined in (H2).

For the proof which is postponed to Section 4.2, we use ideas developed in [**?**] in a discrete time setting. We give an intuition of the result. According to Proposition 3.2, an individual chosen at random at time t is heuristically distributed as Y_t, that is as π for large t thanks to the ergodic property of Y (see (H2)). Moreover two individuals chosen at random among the living individuals at time t have a MRCA who died early, which implies that they behave almost independently. Since Lemma 2.4 implies that the number of individuals alive at time t grows to infinity on $\{W \neq 0\}$, this yields the law of large numbers stated in Theorem 4.2.

Notice that Theorem 1.1 is a direct consequence of Theorem 4.2 and Remark 4.1.

We also present a law of large numbers when summing over the set of all individuals who died before time t. Recall that $D_t = \sum_{u\in\mathcal{T}} \mathbf{1}_{\{\beta(u)<t\}}$ denotes its cardinal.

Recall S in definition 3.1. Notice that $\mathbb{E}[S_t] = rmt$. We shall consider a slightly stronger hypothesis than (H3):

(H5) There exists $\alpha < r(m-1)$ and $c_3 > 0$ such that $\mathbb{E}_\mu[f^2(Y_t)S_t] \leq c_3 \,\mathrm{e}^{\alpha t}$ for every $t \geq 0$.

Proposition 4.3. *For any $\mu \in \mathcal{P}(E)$ and f a nonnegative measurable function defined on E satisfying (H1-5), we have*

$$\lim_{t\to+\infty} \frac{\sum_{u\in\mathcal{T}} f\big(X^u_{\beta(u)-}\big)\mathbf{1}_{\{\beta(u)<t\}}}{\mathbb{E}[D_t]} = \langle \pi, f\rangle\, W \quad \text{in } L^2(\mathbb{P}_\mu), \tag{4.5}$$

$$\lim_{t\to+\infty} \mathbf{1}_{\{N_t>0\}} \frac{\sum_{u\in\mathcal{T}} f\big(X^u_{\beta(u)-}\big)\mathbf{1}_{\{\beta(u)<t\}}}{D_t} = \langle \pi, f\rangle \mathbf{1}_{\{W\neq 0\}} \quad \text{in } \mathbb{P}_\mu\text{-probability}, \tag{4.6}$$

with W defined by (2.10) and π defined in (H2).

We can then extend these results to path dependent functions. In particular, the next theorem describes the asymptotic distribution of the motion and lineage of an individual taken at random in the tree. In order to avoid a set of complicated hypothesis we shall assume that Y is ergodic with limit distribution π and consider bounded functions.

Theorem 4.4. *We assume that there exists $\pi \in \mathcal{P}(E)$ such that for all $x \in E$, and all real-valued bounded measurable function f defined on E, $\lim_{t\to\infty} Q_t f(x) = \langle \pi, f\rangle$.*

Let $T > 0$. For any real bounded measurable function φ on $\mathbb{D}([0,T], E\times\mathbb{R}_+)$, we have

$$\lim_{t\to\infty} \frac{1}{\mathbb{E}[N_t]} \sum_{u\in V_t} \varphi(X^u_{[t-T,t]}, \Lambda^u_{[t-T,t]} - \Lambda^u_{t-T}) = \mathbb{E}_\pi\Big[\varphi(\widetilde{Y}_{[0,T]})\Big]\, W \quad \text{in } L^2(\mathbb{P}_\mu),$$

$$\lim_{t\to\infty} \frac{1}{N_t} \sum_{u\in V_t} \varphi(X^u_{[t-T,t]}, \Lambda^u_{[t-T,t]} - \Lambda^u_{t-T}) = \mathbb{E}_\pi\Big[\varphi(\widetilde{Y}_{[0,T]})\Big]\mathbf{1}_{\{W\neq 0\}} \quad \text{in } \mathbb{P}_\mu\text{-probability},$$

with W defined by (2.10).

Let J_1 be the transition kernel of Y at a jump of S, more precisely: for all nonnegative measurable function f from E to $\mathbb{R}$,

$$J_1 f(x) = m\int_0^1 \mathbb{E}\big[f(F_I^{(H)}(x,\theta))\big]d\theta, \tag{4.7}$$

where H has the size-biased offspring distribution, and conditionally on H, I is uniform on $\{1,\dots,H\}$.

Proposition 4.5. *We assume that there exists $\pi \in \mathcal{P}(E)$ such that for all $x \in E$, and all real-valued bounded measurable function f defined on E, $\lim_{t\to\infty} Q_t f(x) = \langle \pi, f\rangle$.*

Let φ be a real measurable function defined on E-valued paths. We set, for $x \in E$, $f(x) = \mathbb{E}_x[\varphi(Y_{[0,\tau_1)})]$, with τ_1 from definition 3.1. We have

$$\lim_{t\to+\infty} \frac{\sum_{u\in\mathcal{T}} \varphi\big(X^u_{[\alpha(u),\beta(u))}\big)\mathbf{1}_{\{\beta(u)<t\}}}{\mathbb{E}[D_t]} = \langle \pi, J_1 f\rangle\, W \quad \text{in } L^2(\mathbb{P}_\mu), \tag{4.8}$$

$$\lim_{t\to+\infty} \mathbf{1}_{\{N_t>0\}} \frac{\sum_{u\in\mathcal{T}} \varphi\big(X^u_{[\alpha(u),\beta(u))}\big)\mathbf{1}_{\{\beta(u)<t\}}}{D_t} = \langle \pi, J_1 f\rangle \mathbf{1}_{\{W\neq 0\}} \quad \text{in } \mathbb{P}_\mu\text{- probability},$$

with W defined by (2.10).

Remark 4.6. The hypothesis on Y in Theorem 4.4 and Proposition 4.5 is slightly stronger than the ergodic condition (*i.e.* Y converges in distribution to π), but it is fulfilled if Y converges to π for the distance in total variation (*i.e.* for all $x \in E$, $\lim_{t\to\infty}\sup_{A\in\mathcal{E}} |\mathbb{P}_x(Y_t\in A) - \pi(A)| = 0$). This property is very common for ergodic process.

4.2 Proofs

Proof of Theorem 4.2. We assume (H1-4). We shall first prove (4.3) for f such that $\langle \pi, f\rangle = 0$. We have

$$\mathbb{E}_\mu\Big[\frac{\langle Z_t, f\rangle^2}{\mathbb{E}[N_t]^2}\Big] = A_t + B_t$$

where

$$A_t = \mathbb{E}[N_t]^{-2}\mathbb{E}_\mu\Big[\sum_{u\in V_t} f^2(X_t^u)\Big] \quad \text{and} \quad B_t = \mathbb{E}[N_t]^{-2}\mathbb{E}_\mu\Big[\sum_{\substack{(u,v)\in V_t^2\\ u\neq v}} f(X_t^u)f(X_t^v)\Big].$$

Notice that

$$A_t = \mathrm{e}^{-r(m-1)t}\,\mathbb{E}_\mu\left[f^2(Y_t)\right] = \mathrm{e}^{-r(m-1)t}\,\mu Q_t f^2 \xrightarrow[t\to\infty]{} 0, \tag{4.9}$$

thanks to (2.8) and (3.1) for the first equality and (H3) for the convergence. We focus now on B_t. Notice that Proposition 3.8 and then (H1) and (H4) imply that

$$\mathbb{E}[N_t]^{-2}\mathbb{E}_\mu\Big[\sum_{\substack{(u,v)\in V_t^2\\ u\neq v}} |f(X_t^u)f(X_t^v)|\Big] = r\int_0^t \mu Q_s J_2\Big(Q_{t-s}|f|\otimes Q_{t-s}|f|\Big)\,\mathrm{e}^{-r(m-1)s}\,ds$$

is finite. We thus deduce that

$$B_t = r\int_0^t \mu Q_s J_2\Big(Q_{t-s}f\otimes Q_{t-s}f\Big)\,\mathrm{e}^{-r(m-1)s}\,ds. \tag{4.10}$$

Now, since $\langle \pi, f\rangle = 0$, we deduce from (H2) that for s fixed, and $y, z \in E$, $\lim_{t\to\infty}(Q_{t-s}f\otimes Q_{t-s}f)(y,z) = 0$. Thanks to (H1), there exists g such that $\mathbf{1}_{\{s\leq t\}}|(Q_{t-s}f\otimes Q_{t-s}f)| \leq (g\otimes g)$ and (H4) implies that $\int_0^\infty ds\;\mathrm{e}^{-r(m-1)s}\,\mu Q_s J_2(g\otimes g)$ is finite. Lebesgue Theorem entails that

$$\lim_{t\to\infty} B_t = \lim_{t\to\infty} r\int_0^t \mu Q_s J_2\big(Q_{t-s}f\otimes Q_{t-s}f\big)\,\mathrm{e}^{-r(m-1)s}\,ds = 0.$$

This ends the proof of (4.3) when $\langle \pi, f\rangle = 0$.

In the general case, we have

$$\frac{\langle Z_t, f\rangle}{\mathbb{E}[N_t]} - \langle \pi, f\rangle W = \frac{\langle Z_t, f - \langle \pi, f\rangle\rangle}{\mathbb{E}[N_t]} + \langle \pi, f\rangle\left(\frac{N_t}{\mathbb{E}[N_t]} - W\right). \tag{4.11}$$

Notice that if f and μ satisfy (H1-4) then so do $f - \langle \pi, f\rangle$ and μ. The first term of the sum in the r.h.s. of (4.11) converges to 0 in L^2 thanks to the first part of the proof. The second term converges to 0 in L^2 thanks to Lemma 2.4. Hence we get (4.3) if f and μ satisfy (H1-4).

We deduce (4.4) from (4.3) and (2.10). ■

Proof of Proposition 4.3. We assume (H1-5). We shall first prove (4.5) for f such that $\langle \pi, f\rangle = 0$.

We have:

$$\mathbb{E}[D_t]^{-2}\mathbb{E}_\mu\left[\left(\sum_{u\in\mathcal{T}} f\big(X^u_{\beta(u)-}\big)\mathbf{1}_{\{\beta(u)<t\}}\right)^2\right] = A_t + B_t + C_t,$$

where

$$A_t = \mathbb{E}[D_t]^{-2}\mathbb{E}_\mu\left[\sum_{u\in\mathcal{T}} f^2\big(X^u_{\beta(u)-}\big)\mathbf{1}_{\{\beta(u)<t\}}\right],$$

$$B_t = \mathbb{E}[D_t]^{-2}\mathbb{E}_\mu\left[\sum_{(u,v)\in\mathcal{FT}} f\big(X^u_{\beta(u)-}\big)f\big(X^v_{\beta(v)-}\big)\mathbf{1}_{\{\beta(u)<t,\beta(v)<t\}}\right],$$

$$C_t = 2\mathbb{E}[D_t]^{-2}\mathbb{E}_\mu\left[\sum_{u\prec v, v\in\mathcal{T}} f\big(X^u_{\beta(u)-}\big)f\big(X^v_{\beta(v)-}\big)\mathbf{1}_{\{\beta(v)<t\}}\right].$$

The terms A_t and B_t will be handled similarly as in the proof of Proposition 4.2. Notice that

$$A_t = r\mathbb{E}[D_t]^{-2}\int_0^t ds\, \mathrm{e}^{r(m-1)s}\,\mathbb{E}_\mu[f^2(Y_{s-})] = \frac{r(m-1)^2}{(\mathrm{e}^{r(m-1)t}-1)^2}\int_0^t ds\, \mathrm{e}^{sr(m-1)}\,\mu Q_s f^2 \xrightarrow[t\to\infty]{} 0, \tag{4.12}$$

thanks to (3.4) for the first equality, (2.12) for the second and (H3) for the convergence.

Notice that Corollary 3.7 and then (H1) and (H4) imply that

$$\begin{aligned}\mathbb{E}_\mu\Big[\sum_{(u,v)\in\mathcal{FT}} &|f\big(X^u_{\beta(u)-}\big)||f\big(X^v_{\beta(v)-}\big)|\mathbf{1}_{\{\beta(u)<t,\beta(v)<t\}}\Big]\\ &= r^3\int_{[0,+\infty)^3}\mu Q_s J_2\Big(Q_{s'}|f|\otimes Q_{s''}|f|\Big)\,\mathrm{e}^{r(m-1)(s+s'+s'')}\,\mathbf{1}_{\{s+s'<t,s+s''<t\}}\,ds ds' ds''\end{aligned}$$

is finite. We thus deduce that

$$\begin{aligned}B_t &= \frac{r^3(m-1)^2}{(\mathrm{e}^{r(m-1)t}-1)^2}\int_{[0,+\infty)^3}\mu Q_s J_2\Big(Q_{s'}f\otimes Q_{s''}f\Big)\,\mathrm{e}^{r(m-1)(s+s'+s'')}\,\mathbf{1}_{\{s+s'<t,s+s''<t\}}\,ds ds' ds''\\ &= \frac{r^3(m-1)^2\,\mathrm{e}^{2r(m-1)t}}{(\mathrm{e}^{r(m-1)t}-1)^2}\int_{[0,+\infty)^3}\mu Q_s J_2\Big(Q_{t-t'}f\otimes Q_{t-t''}f\Big)\,\mathrm{e}^{r(m-1)(s-t'-t'')}\,\mathbf{1}_{\{s<t'<t,s<t''<t\}}\,ds dt' dt''.\end{aligned}$$

Now, since $\langle \pi, f\rangle = 0$, we deduce from (H2) that for t', t'' fixed and $y, z \in E$, $\lim_{t\to\infty}(Q_{t-t'}f\otimes Q_{t-t''}f)(y,z) = 0$. Thanks to (H1), there exists g such that $|(Q_{t-t'}f\otimes Q_{t-t''}f)| \leq (g\otimes g)$. Then (H4) implies that $\int_{[0,+\infty)^3}\mu Q_s J_2(g\otimes g)\,\mathrm{e}^{r(m-1)(s-t'-t'')}\,\mathbf{1}_{\{s<t',s<t''\}}\,ds dt' dt''$ is finite. Lebesgue Theorem entails that

$$\lim_{t\to\infty} B_t = 0. \tag{4.13}$$

Let us now consider C_t. We have $C_t \leq C'_t + C''_t$ where

$$C'_t = \mathbb{E}[D_t]^{-2}\mathbb{E}_\mu\left[\sum_{u\prec v, v\in\mathcal{T}} f^2\big(X^v_{\beta(v)-}\big)\mathbf{1}_{\{\beta(v)<t\}}\right] \text{ and } C''_t = \mathbb{E}[D_t]^{-2}\mathbb{E}_\mu\left[\sum_{u\prec v, v\in\mathcal{T}} f^2\big(X^u_{\beta(u)-}\big)\mathbf{1}_{\{\beta(v)<t\}}\right].$$

We deduce from (3.4) that

$$\begin{aligned}C'_t &= \mathbb{E}[D_t]^{-2}\mathbb{E}_\mu\left[\sum_{v\in\mathcal{T}} |v| f^2\big(X^v_{\beta(v)-}\big)\mathbf{1}_{\{\beta(v)<t\}}\right]\\ &= \mathbb{E}[D_t]^{-2} r\int_0^t ds\ \mathrm{e}^{r(m-1)s}\,\mathbb{E}_\mu\left[S_{s-}f^2(Y_{s-})\right]\\ &= \frac{r(m-1)^2}{(\mathrm{e}^{r(m-1)t}-1)^2}\int_0^t ds\ \mathrm{e}^{r(m-1)s}\,\mathbb{E}_\mu\left[S_s f^2(Y_s)\right].\end{aligned}$$

We deduce from (H5) that

$$\lim_{t\to\infty} C'_t = 0. \tag{4.14}$$

Using the conditional expectation w.r.t. X^u, (2.12) and (3.4), we get

$$\begin{aligned}
C''_t &= \mathbb{E}[D_t]^{-2}\mathbb{E}_\mu\left[\sum_{u\in\mathcal{T}} f^2\big(X^u_{\beta(u)-}\big)\mathbf{1}_{\{\beta(u)<t\}} m\mathbb{E}[\sum_{v\in\mathcal{T}}\mathbf{1}_{\{\beta(v)<t'\}}]_{|t'=t-\beta(u)}\right]\\
&= \frac{m}{m-1}\mathbb{E}[D_t]^{-2}\mathbb{E}_\mu\left[\sum_{u\in\mathcal{T}} f^2\big(X^u_{\beta(u)-}\big)\mathbf{1}_{\{\beta(u)<t\}}(\mathrm{e}^{r(m-1)(t-\beta(u))}-1)\right]\\
&= \frac{m}{m-1}\mathbb{E}[D_t]^{-2}\int_0^t ds\ \mathrm{e}^{r(m-1)s}\,\mathbb{E}_\mu\left[f^2(Y_{s-})(\mathrm{e}^{r(m-1)(t-s)}-1)\right]\\
&\leq \frac{m(m-1)\,\mathrm{e}^{r(m-1)t}}{(\mathrm{e}^{r(m-1)t}-1)^2}\int_0^t ds\ \mu Q_s f^2.
\end{aligned}$$

We deduce from (H3) (or (H5)) that

$$\lim_{t\to\infty} C''_t = 0. \tag{4.15}$$

The proof of (4.5), when $\langle\pi, f\rangle = 0$, is then a consequence of (4.12), (4.13), (4.14) and (4.15).

In the general case, we have

$$\begin{aligned}
\mathbb{E}[D_t]^{-1}&\sum_{u\in\mathcal{T}} f\big(X^u_{\beta(u)-}\big)\mathbf{1}_{\{\beta(u)<t\}} - \langle\pi, f\rangle W\\
&= \mathbb{E}[D_t]^{-1}\sum_{u\in\mathcal{T}}\Big(f\big(X^u_{\beta(u)-}\big)-\langle\pi, f\rangle\Big)\mathbf{1}_{\{\beta(u)<t\}} + \langle\pi, f\rangle\left(\frac{D_t}{\mathbb{E}[D_t]}-W\right).
\end{aligned}\tag{4.16}$$

Notice that if f and μ satisfy (H1-5) then so do $f-\langle\pi, f\rangle$ and μ. The first term of the sum in the r.h.s. of (4.16) converges to 0 in L^2 thanks to the first part of the proof. The second term converges to 0 in L^2 thanks to Lemma 2.5. Hence we get (4.5) if f and μ satisfy (H1-5). The convergence in probability is thus obtained thanks to (4.5) and (2.10). ■

Proof of Theorem 4.4. The proof is similar to the proof of Theorem 4.2. Some arguments are shorter as we assume that φ is bounded.

We shall first consider the case $\mathbb{E}_\pi[\varphi(\widetilde{Y}_{[0,T]})] = 0$. We assume that $t > T$. We have

$$\mathbb{E}[N_t]^{-2}\mathbb{E}_\mu\Big[\Big(\sum_{u\in V_t}\varphi\big(X^u_{[t-T,t]}, \Lambda^u_{[t-T,t]}-\Lambda^u_{t-T}\big)\Big)^2\Big] = A_t + B'_t + B''_t,$$

where

$$\begin{aligned}
A_t &= \mathbb{E}[N_t]^{-2}\mathbb{E}_\mu\Big[\sum_{u\in V_t}\varphi^2\big(X^u_{[t-T,t]}, \Lambda^u_{[t-T,t]}-\Lambda^u_{t-T}\big)\Big],\\
B'_t &= \mathbb{E}[N_t]^{-2}\mathbb{E}_\mu\Big[\sum_{\substack{(u,v)\in V_t^2\\ u\neq v}}\varphi\big(X^u_{[t-T,t]}, \Lambda^u_{[t-T,t]}-\Lambda^u_{t-T}\big)\varphi\big(X^v_{[t-T,t]}, \Lambda^v_{[t-T,t]}-\Lambda^v_{t-T}\big)\mathbf{1}_{\{\beta(u\wedge v)\geq t-T\}}\Big],\\
B''_t &= \mathbb{E}[N_t]^{-2}\mathbb{E}_\mu\Big[\sum_{\substack{(u,v)\in V_t^2\\ u\neq v}}\varphi\big(X^u_{[t-T,t]}, \Lambda^u_{[t-T,t]}-\Lambda^u_{t-T}\big)\varphi\big(X^v_{[t-T,t]}, \Lambda^v_{[t-T,t]}-\Lambda^v_{t-T}\big)\mathbf{1}_{\{\beta(u\wedge v)<t-T\}}\Big].
\end{aligned}$$

We assume that φ is bounded by a constant, say c. We have $A_t \leq c^2 \mathbb{E}[N_t]^{-1}$ so that $\lim_{t\to\infty} A_t = 0$. We have, using Proposition 3.8,

$$|B'_t| \leq c^2 \mathbb{E}[N_t]^{-2} \mathbb{E}_\mu \Big[\sum_{\substack{(u,v)\in V_t^2 \\ u\neq v}} \mathbf{1}_{\{\beta(u\wedge v)\geq t-T\}} \Big] = c^2 r \int_0^t \mathrm{e}^{-r(m-1)a} \ \mathbf{1}_{\{a\geq t-T\}} da,$$

so that $\lim_{t\to\infty} B'_t = 0$.

We set $f(x) = \mathbb{E}_x[\varphi(\widetilde{Y}_{[0,T]})]$. Using Proposition 3.8 once more, we get

$$\begin{aligned} B''_t &= r \int_0^t \mathrm{e}^{-r(m-1)a} \, \mathbb{E}_\mu \Big[\widetilde{J}_2 \Big(\mathbb{E}'_{\cdot} \big[\varphi \big(Y'_{[t-a-T,t-a]}, \Lambda'_{[t-a-T,t-a]} - \Lambda'_{t-a-T} \big) \big] \\ &\qquad \otimes \mathbb{E}'_{\cdot} \big[\varphi \big(Y'_{[t-a-T,t-a]}, \Lambda'_{[t-a-T,t-a]} - \Lambda'_{t-a-T} \big) \big] \Big) (\widetilde{Y}_a) \Big] \mathbf{1}_{\{a<t-T\}} da \\ &= r \int_0^{t-T} \mathrm{e}^{-r(m-1)a} \ \mu Q_a J_2 \Big(Q_{t-a-T} f \otimes Q_{t-a-T} f \Big) da \end{aligned}$$

By hypothesis on Y, we have that, for fixed a, $\lim_{t\to\infty} Q_{t-a-T} f = \langle \pi, f \rangle = 0$. Using Lebesgue Theorem, we get $\lim_{t\to\infty} B''_t = 0$. This gives the result for the $L^2(\mathbb{P}_\mu)$ convergence when $\langle \pi, f\rangle = 0$. We conclude in the general case and for the convergence in probability as in the proof of Theorem 4.2. ■

Proof of Proposition 4.5. The proof is similar to the proof of Proposition 4.3. Some arguments are shorter as we assume that φ is bounded.

We shall first prove (4.8) for φ such that $\langle \pi, J_1 f\rangle = 0$. We have:

$$\mathbb{E}[D_t]^{-2} \mathbb{E}_\mu \left[\left(\sum_{u\in\mathcal{T}} \varphi\big(X^u_{[\alpha(u),\beta(u))}\big) \mathbf{1}_{\{\beta(u)<t\}} \right)^2 \right] = A_t + B_t + C_t,$$

where

$$\begin{aligned} A_t &= \mathbb{E}[D_t]^{-2} \mathbb{E}_\mu \left[\sum_{u\in\mathcal{T}} \varphi\big(X^u_{[\alpha(u),\beta(u))}\big)^2 \mathbf{1}_{\{\beta(u)<t\}} \right], \\ B_t &= \mathbb{E}[D_t]^{-2} \mathbb{E}_\mu \left[\sum_{(u,v)\in\mathcal{FT}} \varphi\big(X^u_{[\alpha(u),\beta(u))}\big) \varphi\big(X^v_{[\alpha(v),\beta(v))}\big) \mathbf{1}_{\{\beta(u)<t, \beta(v)<t\}} \right], \\ C_t &= 2\mathbb{E}[D_t]^{-2} \mathbb{E}_\mu \left[\sum_{u\prec v, v\in\mathcal{T}} \varphi\big(X^u_{[\alpha(u),\beta(u))}\big) \varphi\big(X^v_{[\alpha(v),\beta(v))}\big) \mathbf{1}_{\{\beta(v)<t\}} \right]. \end{aligned}$$

We assume that φ is bounded by a constant, say c. We have $A_t \leq c^2/\mathbb{E}[D_t]$ so that $\lim_{t\to\infty} A_t = 0$. Thanks to Corollary 3.7, we have

$$\begin{aligned} |C_t| &\leq 2c^2 \mathbb{E}[D_t]^{-2} \mathbb{E}_\mu [\sum_{v\in\mathcal{T}} |v| \mathbf{1}_{\{\beta(v)<t\}}] \\ &= 2c^2 \mathbb{E}[D_t]^{-2} \int_0^t ds \ \mathrm{e}^{sr(m-1)} \, \mathbb{E}_\mu[S_s] \\ &= 2c^2 \mathbb{E}[D_t]^{-2} \int_0^t ds \ srm \, \mathrm{e}^{sr(m-1)}. \end{aligned}$$

This implies that $\lim_{t\to\infty} C_t = 0$.

We set $h_t(x) = \mathbb{E}_x[\varphi(X_{[0,\tau)})\mathbf{1}_{\{\tau<t\}}]$, where τ is an exponential random variable with mean 1, independent of X.

Using the conditional expectation w.r.t. $X^{u'}$, where u' is the ancestor of u, and $X^{v'}$, where v' is the ancestor of v, we have, according to $u'=v'$ or $u'\neq v'$,

$$B_t = B'_t + B''_t,$$

where

$$B'_t = \mathbb{E}[D_t]^{-2}\mathbb{E}_\mu\left[\sum_{u'\in\mathcal{T}} J_2(h_{t-\beta(u')}\otimes h_{t-\beta(u')})(X^{u'}_{\beta(u')-})\mathbf{1}_{\{\beta(u')<t\}}\right],$$

$$B''_t = \mathbb{E}[D_t]^{-2}\mathbb{E}_\mu\left[\sum_{(u',v')\in\mathcal{FT}} J_1(h_{t-\beta(u')})(X^{u'}_{\beta(u')-})J_1(h_{t-\beta(v')})(X^{v'}_{\beta(v')-})\mathbf{1}_{\{\beta(u')<t,\beta(v')<t\}}\right].$$

Using the definition of J_2, (3.14), we get $|B'_t|\leq c^2\mathbb{E}[D_t]^{-1}(\varsigma^2+m^2-m)$ and thus $\lim_{t\to\infty} B'_t=0$.

We deduce from Corollary 3.7, that

$$\begin{aligned}
B''_t &= \frac{r^3(m-1)^2}{(\mathrm{e}^{r(m-1)t}-1)^2}\int_{[0,+\infty)^3} ds ds' ds'' \\
&\qquad \mu Q_s J_2\Big(Q_{s'}J_1 h_{t-s-s'}\otimes Q_{s''}J_1 h_{t-s-s''}\Big)\,\mathrm{e}^{r(m-1)(s+s'+s'')}\,\mathbf{1}_{\{s+s'<t,s+s''<t\}} \\
&= \frac{r^3(m-1)^2\,\mathrm{e}^{2r(m-1)t}}{(\mathrm{e}^{r(m-1)t}-1)^2}\int_{[0,+\infty)^3} ds dv' dv'' \\
&\qquad \mu Q_s J_2\Big(Q_{t-s-v'}J_1 h_{v'}\otimes Q_{t-s-v''}J_1 h_{v''}\Big)\,\mathrm{e}^{-r(m-1)(s+v'+v'')}\,\mathbf{1}_{\{v'<t-s,v''<t-s\}}.
\end{aligned}$$

By hypothesis on Y, we have that, for fixed s and v, $\lim_{t\to\infty} Q_{t-s-v}J_1h_v = \langle\pi, J_1h_v\rangle$. Using Lebesgue Theorem, we get

$$\lim_{t\to\infty} B''_t = r^3(m-1)^2\int_{[0,+\infty)^3} ds dv' dv'' \langle\pi, J_1 h_{v'}\rangle\langle\pi, J_1 h_{v''}\rangle\,\mathrm{e}^{-r(m-1)(s+v'+v'')}$$

Notice that $h_t(x) = \frac{1}{m}\,\mathbb{E}_x[\varphi(Y_{[0,\tau_1)})\,\mathrm{e}^{r(m-1)\tau_1}\,\mathbf{1}_{\{\tau_1<t\}}]$ so that

$$r(m-1)\int_0^{+\infty} dt\; h_t(x)\,\mathrm{e}^{-r(m-1)t} = \frac{1}{m}\,\mathbb{E}_x[\varphi(Y_{[0,\tau_1)})].$$

Recall $f(x)=\mathbb{E}_x[\varphi(Y_{[0,\tau_1)})]$. We get $\lim_{t\to\infty} B''_t = \frac{1}{(m-1)m^2}\langle\pi,J_1f\rangle^2=0$. Therefore, we get that

$$\lim_{t\to\infty} A_t+B_t+C_t=0,$$

which gives the result for the $L^2(\mathbb{P}_\mu)$ convergence when $\langle\pi,J_1f\rangle=0$. We conclude in the general case and for the convergence in probability as in the proof of Proposition 4.3. ■

5 Examples

We now investigate several examples. In Section 5.1, splitted diffusions are considered as scholar examples. In subsection 5.2, we give a biological application to "cellular aging" when cells divide in continuous time, which is one of the motivation of this work. In Section 5.3, we give a central limit theorem for nonlocal branching Lévy processes.

5.1 Real splitted real diffusions

A first example consists in binary branching: the continuous tree $\mathbb{T}$ is a Yule tree. For the Markov process X, we consider a real diffusion with generator:

$$Lf(x) = b(x)f'(x) + \frac{\sigma^2(x)}{2}f''(x). \tag{5.1}$$

We assume that b and σ are such that there exists a unique strong solution to the corresponding SDE, see for instance [?] Theorem 3.2 p.182.

When a branching occurs, each daughter inherits a random fraction of the value of the mother:

$$F^{(1)}(x,\theta) = G^{-1}(\theta)x, \quad F^{(2)}(x,\theta) = (1-G^{-1}(\theta))x,$$

where G is the cumulative distribution function of the random fraction in $[0,1]$ associated with the branching event. We assume the distribution of the random fraction is symmetric: $G(x) = 1-G(1-x)$.

The infinitesimal generator of Y is characterized for $f \in \mathcal{C}_b^2(\mathbb{R},\mathbb{R})$ by:

$$\begin{aligned}
Af(x) =& b(x)f'(x) + \sigma(x)f''(x) + 2r\int_0^1 \left(\frac{1}{2}\big(f(G^{-1}(\theta)x) - f(x)\big) + \frac{1}{2}\big(f((1-G^{-1}(\theta))x) - f(x)\big)\right) d\theta \\
=& b(x)f'(x) + \sigma(x)f''(x) + 2r\int_0^1 \big(f(qx) - f(x)\big)G(dq).
\end{aligned} \tag{5.2}$$

Particular choices for the functions b and σ are the following ones:

(i) If $b(x) = 0$ and $\sigma(x) = \sigma$, we obtain the splitted Brownian process.

(ii) If $b(x) = -\beta(x-\alpha)$ and $\sigma(x) = \sigma$, we obtain the splitted Ornstein-Uhlenbeck process.

(iii) If $b(x) = 1$ and $\sigma(x) = 0$, the deterministic process X can represent the linear growth of some biological content of the cell (nutriments, proteins, parasites...) which is shared randomly in the two daughter cells when the cell divides. More precisely here, each daughter inherits random fraction of this biological content.

Let us note that if $b(x) = \beta x$ and $\sigma(x)^2 = \sigma^2 x$, we obtain the splitted Feller diffusion. But in this case, almost surely, the auxiliary process either becomes extinct or goes to infinity as $t \to \infty$. The assumption (H2) is not satisfied. This natural model for parasite infection is studied in [?].

The following results give the asymptotic limit of the splitted diffusion under some condition which is satisfied by the examples (i-iii). For this we use results due to Meyn and Tweedie [?, ?].

Proposition 5.1. *Assume that Y is Feller and irreducible (see [?] p. 520) and that there exists $K \in \mathbb{R}_+$, such that for every $|x| > K$, $b(x)/x < r'$ with $r' < r$. Then, the auxiliary process Y of generator A is ergodic with stationary probability π. Furthermore $\sum_{u\in V_t} \delta_{X_t^u}(dx)/N_t$ converges weakly to π as $t \to \infty$ and this convergence holds in probability.*

Proof. Once we check that Y is ergodic, then Corollary 1.2 and the fact that W defined by (2.10) is a.s. positive readily imply the weak convergence of the Proposition. To prove the ergodicity of Y, we use Theorems 4.1 of [?] and 6.1 of [?]. Since Y is Feller and irreducible, the process Y admits a unique invariant probability measure π and is exponentially ergodic provided the condition (CD3) in [?] is satisfied, namely, if there exists a positive measurable function $V : x \mapsto V(x)$ such that $\lim_{x\to\pm\infty} V(x) = +\infty$ and for which:

$$\exists c > 0,\, d \in \mathbb{R},\, \forall x \in \mathbb{R},\, AV(x) \le -cV(x) + d. \tag{5.3}$$

For $V(x) = |x|$ regularized on an ε-neighborhood of 0 ($0 < \varepsilon < 1$), we have:

$$\forall |x| > \varepsilon,\, AV(x) = \text{sign}(x)b(x) + 2r|x| \int_0^1 (q-1)G(dq) = \text{sign}(x)b(x) - r|x|, \tag{5.4}$$

as the distribution of G is symmetric. By assumption, there exists $\eta > 0$ and $K > \varepsilon$, such that (5.4) implies:

$$\forall x \in \mathbb{R},\, AV(x) \le -\eta V(x) + \big(\sup_{|x|\le K} |b(x)| + rK\big)\mathbf{1}_{\{|x|\le K\}}. \tag{5.5}$$

This implies (5.3) and finishes the proof ; the geometric ergodicity expresses here as:

$$\exists \beta > 0,\, B < +\infty,\, \forall t \in \mathbb{R}_+,\, \forall x \in \mathbb{R}, \quad \sup_{g \,/\, |g(u)| \le 1+|u|} \big|Q_t g(x) - \langle \pi, g\rangle\big| \le B(1+|x|)\,\mathrm{e}^{-\beta t}. \tag{5.6}$$

■

Remark 5.2. The examples (i-iii) satisfy the assumptions of Proposition 5.1. If b and σ are bounded Lipschitz functions, X is Feller (*e.g.* Theorem 6.3.4 p. 152 of [?]), and thus Y is also Feller. The Feller property also holds for Ornstein-Uhlenbeck processes. The irreducibility property is well known for diffusions as (i) and (ii) and trivial for (iii).

Remark 5.3. If there exists $K > 0$ in Prop. 5.1 such that for every $|x| > K$, $2b(x)/x + 6\sigma(x)/x^2 < r'$ with $r' < r \int_0^1 (1-q^4)^2 G(dq)$, then we can use similar arguments as in the proof of Proposition 5.1. We get that the auxiliary process Y is geometrically ergodic with

$$\exists \beta > 0,\, B < +\infty,\, \forall t \in \mathbb{R}_+,\, \forall x \in \mathbb{R}, \quad \sup_{g \,/\, |g(u)| \le 1+|u|^4} \big|Q_t g(x) - \langle \pi, g\rangle\big| \le B(1+|x|^4)\,\mathrm{e}^{-\beta t} \tag{5.7}$$

instead of (5.6). This result will be used for the proof of the central limit theorem.

5.2 Cellular-aging process

We now present a generalization to the continuous time of Guyon [?] and Delmas and Marsalle [?] about cellular aging. When a rode shaped cell divides, it produces a new end per progeny cell. So each new cell has a pole (or end) which is new and an other one which was created one or more generations ago. This number of generations is the age of the cell. Since each cell has a new pole and an older one, at the next division one of the two daughters will inherit the new pole and the other one will inherit the older pole. Experiments indicate that the first one has a larger growth rate than the second one (see Stewart *et al.* [?] for details), which indicates aging.

To detect this aging effect, [?, ?] used discrete time Markov models by looking at cells of a given generation. To take into account the asynchrony of cell divisions, it may be useful to consider continuous time genealogical tree.

We consider the following model. Cells are characterized by a type $\eta \in \{0,1\}$ (type 0 corresponds to a cell of age 1 and type 1 to cell of greater age) and a quantity ζ (growth rate, quantity of damage in the cell) that evolves according to a Markov process depending on the type of the cell. Cells may die, which leads us to the following model. At rate r, each cell is replaced by one cell of type 0 (resp. 1) with probability $p_0 \geq 0$ (resp. $p_1 \geq 0$), to two cells of type 0 and 1 with probability $p_{0,1} \geq 0$, or to no cell with probability $1 - p_0 - p_1 - p_{0,1} \geq 0$. The way the quantity ζ is given to a daughter depends on its type and on the fact that it has or not a sister.

This can be stated in the framework of Sections 2.2 and 3. For the sake of simplicity, we shall assume that ζ evolves as a real diffusion between two branching times.

Let L^0 and L^1 be two diffusion generators: for $f \in \mathcal{C}^2(\mathbb{R} \times \{0,1\}, \mathbb{R})$:

$$L^\eta f(\zeta,\eta) = b(\zeta,\eta)\partial_\zeta f(\zeta,\eta) + \sigma(\zeta,\eta)\partial^2_{\zeta,\zeta} f(\zeta,\eta), \quad \eta \in \{0,1\}. \tag{5.8}$$

We assume there exists a unique strong solution to the corresponding two SDE, see for instance [?] Theorem 3.2 p.182. We consider the underlying process $X = ((\zeta_t,\eta_t), t \geq 0)$ with generator

$$Lf(\zeta,\eta) = \mathbf{1}_{\{\eta=0\}} L^0 f(\zeta,0) + \mathbf{1}_{\{\eta=1\}} L^1 f(\zeta,1).$$

Notice the process $(\eta_t, t \geq 0)$ is constant between two branching times. The offspring distribution is

$$p(dk) = (1 - p_0 - p_1 - p_{0,1})\delta_0(dk) + (p_0 + p_1)\delta_1(dk) + p_{0,1}\delta_2(dk). \tag{5.9}$$

The offspring position is given by:

$$\begin{aligned} F_1^{(1)}((\zeta,\eta),\theta) =& (g_0(\zeta,\eta,\theta_1),0)\mathbf{1}_{\{\theta_2 \leq p_0/(p_0+p_1)\}} + (g_1(\zeta,\eta,\theta_1),1)\mathbf{1}_{\{\theta_2 > p_0/(p_0+p_1)\}} \\ F_{i+1}^{(2)}((\zeta,\eta),\theta) =& (g_i^{(2)}((\zeta,\eta,\theta),i) \quad \text{for } i \in \{0,1\}, \end{aligned} \tag{5.10}$$

for some functions g_0, g_1, g_2^0, g_2^1 and (θ_1,θ_2) a function of θ such that if θ is uniform on $[0,1]$, then θ_1 and θ_2 are independent and uniform on $[0,1]$. The division is asymmetric if $g_0^{(2)} \neq g_1^{(2)}$. One important issue is, using the law of large number (Section 4) and fluctuation results, to test if the division is asymmetric, which means aging, or not. Let us mention that a natural question would be to give the test in a more general model in which the division rate depends on the state of the cell and of the quantity of interest ζ (which is realistic if for example ζ describe the quantity of damage of the cell).

Let us consider a test function $f : (t,\zeta,\eta) \mapsto f_t(\zeta,\eta)$ in $\mathcal{C}_b^{1,2}(\mathbb{R}_+ \times (\mathbb{R} \times \{0,1\}), \mathbb{R})$, and let $(B^u)_{u\in\mathcal{U}}$ be a family of independent standard Brownian motions. The SDE describing the evolution of the population of cells then becomes with the notations of (5.8), (5.9) and (5.10):

$$\begin{aligned} \langle Z_t, f_t\rangle = \langle Z_0, f_0\rangle &+ \int_0^t \int_{\mathcal{U}\times\{0,1,2\}\times[0,1]} \mathbf{1}_{\{u\in V_{s_-}\}} \left(\sum_{j=1}^k f_s(F_j^{(k)}((\zeta^u_{s_-},\eta^u_{s_-}),\theta)) - f_s(\zeta^u_{s_-}) \right) \rho(ds,du,dk,d\theta) \\ &+ \int_0^t \int_{\mathbb{R}\times\{0,1\}} (L^\eta f(\zeta,\eta) + \partial_s f_s(\zeta,\eta))\, Z_s(d\zeta,d\eta)\, ds \qquad (5.11) \\ &+ \int_0^t \sum_{u\in V_s} \sqrt{2\sigma(\zeta^u_s,\eta^u_s)}\partial_\zeta f(\zeta^u_s,\eta^u_s) dB^u_s. \qquad (5.12) \end{aligned}$$

If Y is ergodic, then $\mathbf{1}_{\{N_t>0\}} \sum_{u\in V_t} \delta_{X_t^u}(dx)/N_t$ converges to a deterministic non degenerated measure on $\mathbb{R}_+ \times \{0,1\}$. Given a particular choice for the parameters g_0, g_1, g_2^0, g_2^1, L^0, L^1, p_0, p_1 and $p_{0,1}$ of the model, one can use arguments similar to the ones used in Proposition 5.1 and Remark 5.2 to prove the ergodicity of Y.

5.3 Branching Lévy process

We consider particles moving independently on $\mathbb{R}$ following a Lévy process X and reproducing with constant rate r. Each child jumps from the location of the mother when the branching occurs. We are interested in the rescaled population location at large time.

The generator of the underlying process X is given by:

$$Lf(x) = bf'(x) + \frac{\sigma^2}{2} f''(x) + \int_{\mathbb{R}\setminus\{0\}} \big(f(x+y) - f(x) - yf'(x)\mathbf{1}_{\{|y|<1\}}\big) h(dy) \tag{5.13}$$

with $b \in \mathbb{R}$, $\sigma \in \mathbb{R}_+$ and h a measure on $\mathbb{R}\setminus\{0\}$ such that $\int_{\mathbb{R}\setminus\{0\}} y^2 \; h(dy) < +\infty$. The particles reproduce at rate r in a random number of offspring distributed as $p = (p_k, k \in \mathbb{N})$, such that $\sum_{k\geq 1} kp_k > 1$ (supercritical case). The offspring position is defined as follows:

$$F_j^{(k)}(x,\theta) = x + \Delta_j^k(\theta), \quad j \in \{1,\dots,k\}, \tag{5.14}$$

where we recall that x is the location just before branching time and k is the number of offspring. We assume the following second moment condition: $\sum_{k\in\mathbb{N}} p_k \sum_{j=1}^k \mathbb{E}[\Delta_j^k(\Theta)^2] < \infty$, where Θ is uniform on $[0,1]$.

Proposition 5.4. *We have the following weak convergence in $\mathcal{M}_F(\mathbb{R})$:*

$$\lim_{t\to+\infty} \frac{1}{N_t} \sum_{u\in V_t} \delta_{\frac{X_t^u - \beta t}{\sqrt{t}}}(dx) = \pi_\Sigma(dx)\mathbf{1}_{\{W>0\}}, \qquad \textit{in probability} \tag{5.15}$$

where π_Σ is the centered Gaussian probability measure with variance Σ and

$$\beta = b + \int_{\mathbb{R}\setminus\{0\}} y\mathbf{1}_{\{|y|\geq 1\}} \; h(dy) + r\sum_{k=1}^{+\infty} p_k \sum_{j=1}^{k} \mathbb{E}[\Delta_j^k(\Theta)], \tag{5.16}$$

$$\Sigma = \sigma^2 + \int_{\mathbb{R}\setminus\{0\}} y^2 \; h(dy) + r\sum_{k=1}^{+\infty} p_k \sum_{j=1}^{k} \mathbb{E}[\Delta_j^k(\Theta)^2]. \tag{5.17}$$

Proof. The auxiliary process Y is a Lévy process with generator:

$$\begin{aligned} Af(x) = {} & bf'(x) + \frac{\sigma^2}{2} f''(x) + \int_{\mathbb{R}} \big(f(x+y) - f(x) - yf'(x)\mathbf{1}_{\{|y|<1\}}\big) \; h(dy) \\ & + rm \sum_{k=1}^{+\infty} \frac{kp_k}{m} \int_0^1 \sum_{j=1}^{k} \frac{1}{k} \big(f(x+\Delta_j^k(\theta)) - f(x)\big) d\theta. \end{aligned}$$

In particular, we have for all $x \in \mathbb{R}$:

$$\mathbb{E}_x[Y_t] = x + t\Big(b + \int_{\mathbb{R}\setminus\{0\}} y\mathbf{1}_{\{|y|\geq 1\}}\ h(dy) + r\sum_{k=1}^{+\infty} p_k \sum_{j=1}^{k} \mathbb{E}[\Delta_j^k(\Theta)]\Big) = x + \beta t \tag{5.18}$$

$$\mathbb{E}_x[Y_t^2] - \mathbb{E}_x[Y_t]^2 = t\Big(\sigma^2 + \int_{\mathbb{R}\setminus\{0\}} y^2\ h(dy) + r\sum_{k=1}^{+\infty} p_k \sum_{j=1}^{k} \mathbb{E}[\Delta_j^k(\Theta)^2]\Big) = \Sigma t. \tag{5.19}$$

Then, we deduce from the central limit theorem for Lévy processes or directly from Lévy Khintchine formula, that $((Y_t - \beta t)/\sqrt{t}, t \geq 0)$ converges in distribution to π_Σ. This implies that for any fixed s, $((Y_{t-s} - \beta t)/\sqrt{t}, t \geq 0)$ converges in distribution to π_Σ.

Let φ be a continuous bounded real function and define

$$f_t(x) := \varphi\Big((x - \beta t)/\sqrt{t}\Big) \qquad \text{for } t \geq 0,\ x \in \mathbb{R}.$$

Let $(Q_t, t \geq 0)$ be the transition semi-group of Y. We get that for any fixed s and $x \in \mathbb{R}$,

$$\lim_{t\to+\infty} Q_{t-s} f_t(x) = \langle \pi_\Sigma, \varphi\rangle. \tag{5.20}$$

It is then very easy to adapt the proof of Theorem 4.2 with f replaced by $f_t - \langle \pi_\Sigma, \varphi\rangle$: (4.9) holds since f_t is uniformly bounded; (4.10) holds using similar arguments with (5.20) instead of (H2) and f_t uniformly bounded instead of (H1) and (H4) arguments. Similar arguments as in the end of the proof of Theorem 4.2 imply that for any continuous bounded real function φ, the following convergence in probability holds:

$$\lim_{t\to+\infty} \frac{1}{N_t} \sum_{u\in V_t} \varphi\Big(\frac{X_t^u - \beta t}{\sqrt{t}}\Big) = \lim_{t\to+\infty} \frac{1}{N_t} \langle Z_t, f_t\rangle = \langle \pi_\Sigma, \varphi\rangle\ \mathbf{1}_{\{W>0\}}.$$

This gives (5.15). ■

6 Central limit theorem

6.1 Fluctuation process

In order to study the fluctuations associated to the LLNs, Theorem 4.2, we shall use the martingale associated to Z_t, see (2.17). We focus on the simple case of splitted diffusions developed in Section 5.1. Our main result for this section is stated as Proposition 6.4.

In the sequel, C denotes a constant that may change from line to line. We keep notations from Section 5.1. We assume that b and σ are such that there exists a unique strong solution to the corresponding SDE, see for instance [?] Theorem 3.2 p.182.

We consider the following sequence of fluctuation processes indexed by $T > 0$. For $f \in \mathcal{B}_b(\mathbb{R}_+, \mathbb{R})$,

$$\langle \eta_t^T, f\rangle = \sqrt{\mathbb{E}[N_{t+T}]}\left(\frac{\langle Z_{t+T}, f\rangle}{\mathbb{E}[N_{t+T}]} - \frac{\langle Z_T, Q_t f\rangle}{\mathbb{E}[N_T]}\right) \tag{6.1}$$

where we recall that $N_t = \mathrm{Card}(V_t) = \langle Z_t, 1\rangle$ and Q_t has been defined in (4.2). The family Q_t is the transition semigroup of the auxiliary process Y, which is given by:

$$Y_t = X_0 + \int_0^t b(Y_s)ds + \int_0^t \sigma(Y_s)dB_s - \int_0^t (1-q)Y_{s_-}\rho(ds,dq) \tag{6.2}$$

where X_0 is an initial condition with distribution μ, where $(B_t)_{t\in\mathbb{R}_+}$ is a standard real Brownian motion and where $\rho(ds,dq)$ is a Poisson point measure with intensity $2r\,ds\otimes\widetilde{G}(dq)$ with $\widetilde{G}$ such that $\int_{[0,1]} f(q)\widetilde{G}(dq) = \int_{[0,1]}(f(q)/2 + f(1-q)/2)G(dq)$. As in Section 5.1, we will assume in the sequel that G is symmetric. In this case, $\widetilde{G}(dq) = G(dq)$.
The idea in (6.1) is to compare the independent trees that have grown from the particles of Z_T between times T and $t+T$, with the positions of independent auxiliary processes at time t and started at the positions Z_T. We recall that L is the generator defined in (5.1), and let J be the operator defined on the space of locally integrable functions by

$$Jf(x) = -\frac{3r}{2}f(x) + r\int_0^1 \Big(f(qx) + f((1-q)x)\Big)G(dq) = -\frac{3r}{2}f(x) + 2r\int_0^1 f(qx)G(dq). \tag{6.3}$$

Proposition 6.1. *The fluctuation process (6.1) satisfies the following evolution equation:*

$$\langle \eta_t^T, f\rangle = \int_0^t \int_{\mathbb{R}} \Big(Lf(x) + Jf(x)\Big)\eta_s^T(dx)\,ds + M_t^T(f), \tag{6.4}$$

where $M_t^T(f)$ is a square integrable martingale with quadratic variation:

$$\langle M^T(f)\rangle_t = \int_0^t ds \int_{\mathbb{R}} \frac{Z_{s+T}(dx)}{\mathbb{E}[N_{s+T}]}\left[r\int_0^1 \Big(f(qx) + f((1-q)x) - f(x)\Big)^2 G(dq) + 2\sigma^2(x)f'(x)^2\right]. \tag{6.5}$$

The proof of this proposition is given in Section 6.3. In the following, we are interested in the behavior of the fluctuation process when $T \to +\infty$. The processes η^T take their values in the space $\mathcal{M}_S(\mathbb{R})$ of signed measures. Since this space endowed with the topology of weak convergence is not metrizable, we follow the approach of Métivier [?] and Méléard [?] (see also [?, ?]) and embed $\mathcal{M}_S(\mathbb{R})$ in weighted distribution spaces. This is described in the sequel. We then prove the convergence of the fluctuation processes to a distribution-valued diffusion driven by a Gaussian white noise (Proposition 6.4).

6.2 Convergence of the fluctuation process: the Central Limit Theorem

Let us introduce the Sobolev spaces that we will use (see *e.g.* Adams [?]). We follow in this the steps of [?, ?]. To obtain estimates of our fluctuation processes, the following additional regularities for b and σ are required, as well as assumptions on our auxiliary process.

Assumption 6.2. *We assume that:*
(i) b and σ are in $\mathcal{C}^8(\mathbb{R},\mathbb{R})$ with bounded derivatives.
(ii) There exists $K > 0$ such that for every $|x| > K$, $2b(x)/x + 6\sigma(x)/x^2 < r'$ with $r' < r\int_0^1 (1-q^4)^2 G(dq)$.
(iii) Y is ergodic with stationary measure π such that $\langle \pi, |x|^8\rangle < +\infty$.
(iv) for every initial condition μ such that $\langle \mu, |x|^8\rangle < +\infty$, $\sup_{t\in\mathbb{R}_+}\mathbb{E}_\mu[Y_t^8] < +\infty$.

Remark 6.3. (i) Notice that under Assumption 6.2 (i), there exist $\bar{b}$ and $\bar{\sigma} > 0$ s.t. forall $x \in \mathbb{R}$, we have $|b(x)| \leq \bar{b}(1 + |x|)$ and $|\sigma(x)| \leq \bar{\sigma}(1 + |x|)$.
(ii) Conditions for the ergodicity of Y have been provided in Proposition 5.1 and Remarks 5.2 and 5.3. Under Assumption 6.2 (ii), Remark 5.3 applies and we have geometrical ergodicity with (5.7).
(iii) The moment hypothesis of Assumption 6.2 (iv) is fulfilled for the examples (i-iii) of Section 5.1 provided the initial condition satisfies $\langle \mu, |x|^8 \rangle < +\infty$. This can be seen by using Itô's formula (*e.g.* [?], Th. 5.1 p. 67) and Gronwall's Lemma. Moreover, for every $p \in \{1, \dots, 8\}$, $\mathbb{E}_\mu\big[|Y_t|^p\big] < +\infty$.
(iv) Assumptions 6.2 (iii) and (iv) imply: $\forall p \in \{1, \dots, 7\}$, $\int_{\mathbb{R}} |x|^p \pi(dx) < +\infty$ and $\lim_{t \to +\infty} \mathbb{E}_\mu[|Y|^p] = \int_{\mathbb{R}} |x|^p \pi(dx)$. This is a consequence of the equi-integrability of $(|Y_t|^p)_{t \geq 0}$ for $p \in \{1, \dots, 7\}$.

For $j \in \mathbb{N}$ and $\alpha \in \mathbb{R}_+$, we denote by $W^{j,\alpha}$ the closure of $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ with respect to the norm:

$$\|g\|_{W^{j,\alpha}} := \left(\sum_{k \leq j} \int_{\mathbb{R}} \frac{|g^{(k)}(x)|^2}{1 + |x|^{2\alpha}} dx \right)^{1/2}, \tag{6.6}$$

where $g^{(k)}$ is the k^{th} derivative of g. The space $W^{j,\alpha}$ endowed with the norm $\|.\|_{W^{j,\alpha}}$ defines a Hilbert space. We denote by $W^{-j,\alpha}$ the dual space. Let $C^{j,\alpha}$ be the space of functions g with j continuous derivatives and such that

$$\forall k \leq j, \lim_{|x| \to +\infty} \frac{|g^{(k)}(x)|}{1 + |x|^\alpha} = 0.$$

When endowed with the norm:

$$\|g\|_{C^{j,\alpha}} := \sum_{k \leq j} \sup_{x \in \mathbb{R}} \frac{|g^{(k)}(x)|}{1 + |x|^\alpha}, \tag{6.7}$$

these spaces are Banach spaces, and their dual spaces are denoted by $C^{-j,\alpha}$.

In the sequel, we will use the following embeddings (see [?, ?]):

$$\begin{aligned} &C^{7,0} \hookrightarrow W^{7,1} \hookrightarrow_{H.S.} W^{5,2} \hookrightarrow C^{4,2} \hookrightarrow W^{4,3} \hookrightarrow C^{3,3} \hookrightarrow W^{3,4} \hookrightarrow C^{2,4} \\ &C^{-2,4} \hookrightarrow W^{-3,4} \hookrightarrow C^{-3,3} \hookrightarrow W^{-4,3} \hookrightarrow C^{-4,2} \hookrightarrow W^{-5,2} \hookrightarrow_{H.S.} W^{-7,1} \hookrightarrow C^{-7,0}, \end{aligned} \tag{6.8}$$

where $H.S.$ means that the corresponding embedding is Hilbert-Schmidt (see [?] p.173). Let us explain briefly why we use these embeddings. Following the preliminary estimates of [?] (Proposition 3.4), it is possible to choose $W^{-3,4}$ as a reference space for our study. We control the norm of the martingale part in $W^{-4,3}$ using the embeddings $W^{4,3} \hookrightarrow C^{3,3} \hookrightarrow W^{3,4}$. We obtain uniform estimate for the norm of η^T_t in $C^{-4,2}$. The spaces $W^{-5,2}$ and $W^{-7,1}$ are used to apply the tightness criterion in [?] (see our Lemma 6.8). The space $C^{-7,0}$ is used for proving uniqueness of the accumulation point of the family $(\eta^T)_{T \geq 0}$.

Proposition 6.4. *Let $\Upsilon > 0$. The sequence $(\eta^T)_{T \in \mathbb{R}_+}$ converges in $\mathbb{D}([0, \Upsilon], C^{-7,0})$ when $T \to +\infty$ to the unique solution in $C([0, \Upsilon], C^{-7,0})$ of the following evolution equation:*

$$\langle \eta_t, f \rangle = \int_0^t \int_{\mathbb{R}} \Big(Lf(x) + Jf(x) \Big) \eta_s(dx)\, ds + \sqrt{W} \mathcal{W}_t(f), \tag{6.9}$$

where $\mathcal{W}(f)$ is a Gaussian martingale independent of W and which bracket is $V(f) \times t$ with:

$$V(f) = \int_{\mathbb{R}} \left(r \int_0^1 \big(f(qx) + f((1-q)x) - f(x) \big)^2 G(dq) + 2\sigma^2(x) f'(x)^2 \right) \pi(dx). \tag{6.10}$$

Notice that unlike the discrete case treated in [?], our fluctuation process has here a finite variational part.

6.3 Proofs

We begin by establishing the evolution equation for η^T that are announced in Proposition 6.1.

Proof of Proposition 6.1. From Lemma 2.4 and applying (2.22) with $f_t(x) = \mathrm{e}^{-rt/2} f(x)$, we obtain:

$$\begin{aligned}\langle Z_{t+T}, f\rangle \,\mathrm{e}^{-r(t+T)/2} =& \langle Z_T, f\rangle \,\mathrm{e}^{-rT/2} + M_t^T(f) \\ &+ \int_0^t \int_{\mathbb{R}} \Big(Lf(x) + Jf(x)\Big)\, \mathrm{e}^{-r(s+T)/2}\, Z_{s+T}(dx)\, ds,\end{aligned} \tag{6.11}$$

where $M_t^T(f)$ is a square integrable martingale with quadratic variation:

$$\langle M^T(f)\rangle_t = \int_T^{t+T} ds \int_{\mathbb{R}} \mathrm{e}^{-rs}\, Z_s(dx) \left[r \int_0^1 \Big(f(qx) + f((1-q)x) - f(x)\Big)^2 G(dq) + 2\sigma^2(x) f'(x)^2 \right],$$

which is the bracket announced in (6.5). Computing in the same way $\langle Z_t, f\rangle \,\mathrm{e}^{-rt/2}$ and taking the expectation gives, with (4.2) and Proposition 3.2:

$$Q_t f(x)\, \mathrm{e}^{rt/2} = f(x) + \int_0^t Q_s\Big(Lf + Jf\Big)(x)\, \mathrm{e}^{rs/2}\, ds.$$

Integrating with respect to Z_T and multiplying by $\mathrm{e}^{-rT/2}$ imply:

$$\langle Z_T, Q_t f\rangle \,\mathrm{e}^{-r(T-t)/2} = \langle Z_T, f\rangle \,\mathrm{e}^{-rT/2} + \int_0^t \mathrm{e}^{-r(T-s)/2}\, ds\, \langle Z_T,\, Q_s(Lf + Jf)\rangle. \tag{6.12}$$

We deduce the announced result from (6.1), (6.11) and (6.12). ■

We now prove that our fluctuation process η^T can be viewed as a process with values in $W^{-3,4}$, by following the preliminary estimates of [?] (Proposition 3.4). This space $W^{-3,4}$ is then chosen as reference space and in all the spaces appearing in the second line of (6.8) that contain $W^{-3,4}$, the norm of η_t^T is finite and well defined.

Lemma 6.5. *Let $\Upsilon > 0$. There exists a finite constant C that does not depend on T nor on Υ such that*

$$\sup_{t\in[0,\Upsilon]} \mathbb{E}_\mu\Big[\|\eta_t^T\|^2_{W^{-3,4}}\Big] \le C\, \mathrm{e}^{r(\Upsilon+T)}. \tag{6.13}$$

Proof. Let $(\varphi_p)_{p\in\mathbb{N}^*}$ be a complete orthonormal basis of $W^{3,4}$ that are $\mathcal{C}^\infty$ with compact support. We have by Riesz representation theorem and Parseval's identity:

$$\begin{aligned}\mathrm{e}^{-r(t+T)}\, \|\eta_t^T\|^2_{W^{-4,3}} =& \mathrm{e}^{-r(t+T)} \sum_{p\ge 1} \langle \eta_t^T, \varphi_p\rangle^2 \\ =& \mathrm{e}^{-r(t+T)}\, \mathbb{E}[N_{t+T}] \sum_{p\ge1} \Big(\frac{\langle Z_{t+T}, \varphi_p\rangle}{\mathbb{E}[N_{t+T}]} - \frac{\langle Z_T, Q_t\varphi_p\rangle}{\mathbb{E}[N_T]} \Big)^2 \\ \le& 2 \sum_{p\ge1} \Big(\frac{\langle Z_{t+T}, \varphi_p\rangle^2}{\mathbb{E}[N_{t+T}]^2} + \frac{\langle Z_T, Q_t\varphi_p\rangle^2}{\mathbb{E}[N_T]^2} \Big).\end{aligned} \tag{6.14}$$

Under the Assumption 6.2 (iii) and thanks to Remark 4.1 and Example 1, we use the same proof as in Theorem 4.2, especially (4.9) and (4.10):

$$\begin{aligned}
0 < \mathbb{E}_\mu\Big[\frac{\langle Z_{t+T},\varphi_p\rangle^2}{\mathbb{E}[N_{t+T}]^2}+\frac{\langle Z_T,Q_t\varphi_p\rangle^2}{\mathbb{E}[N_T]^2}\Big]&\\
=&\,\mathrm{e}^{-r(t+T)}\,\mu Q_{t+T}\varphi_p^2+r\int_0^{t+T}\mu Q_s J_2\Big(Q_{t+T-s}\varphi_p\otimes Q_{t+T-s}\varphi_p\Big)\,\mathrm{e}^{-rs}\,ds\\
&+\mathrm{e}^{-rT}\,\mu Q_T\big(Q_t\varphi_p\big)^2+r\int_0^{T}\mu Q_s J_2\Big(Q_{T-s}Q_t\varphi_p\otimes Q_{T-s}Q_t\varphi_p\Big)\,\mathrm{e}^{-rs}\,ds\\
\leq&\,2\,\mathrm{e}^{-rT}\,\mu Q_{t+T}\varphi_p^2+4r\int_0^{t+T}\int_0^1\int_{\mathbb{R}}\varphi_p^2(qx)\,\mathrm{e}^{-rs}\;\mu Q_{t+T}(dx)\;G(dq)\;ds \qquad (6.15)
\end{aligned}$$

since by (3.14), Cauchy-Schwarz' inequality and symmetry of G:

$$\begin{aligned}
J_2\Big(Q_{t+T-s}|\varphi_p|\otimes Q_{t+T-s}|\varphi_p|\Big)(x)=&\,2\int_0^1\Big(Q_{t+T-s}|\varphi_p|(qx)Q_{t+T-s}|\varphi_p|((1-q)x)\Big)G(dq)\\
\leq&\,2\int_0^1 Q_{t+T-s}\varphi_p^2(qx)G(dq).
\end{aligned}$$

We deduce from (6.14) and (6.15) that:

$$\begin{aligned}
\mathrm{e}^{-r(t+T)}\,\mathbb{E}_\mu\big[\|\eta_t^T\|^2_{W^{-4,3}}\big]\leq&\,4\,\mathrm{e}^{-rT}\int_{\mathbb{R}}\sum_{p\geq 1}\varphi_p^2(x)\;\mu Q_{t+T}(dx)\\
&+8r\int_0^{t+T}\int_0^1\int_{\mathbb{R}}\sum_{p\geq 1}\varphi_p^2(qx)\,\mathrm{e}^{-rs}\;\mu Q_{t+T}(dx)\,G(dq)\,ds. \qquad (6.16)
\end{aligned}$$

Let us consider the linear forms $D_{x,F}(g)=g(Fx)$ for $F\in[0,1]$, $x\in\mathbb{R}$ and $g\in W^{3,4}\hookrightarrow C^{2,4}$:

$$|D_{x,F}(g)|=|g(Fx)|\leq(1+|x|^4)\|g\|_{C^{2,4}}\leq C(1+|x|^4)\|g\|_{W^{3,4}}$$

Using Riesz representation theorem and Parseval's identity, we get:

$$\sum_{p\geq 1}D_{x,F}(\varphi_p)^2=\|D_{x,F}\|^2_{W^{-3,4}}\leq C(1+|x|^4). \qquad (6.17)$$

We deduce from (4.2), (6.16) and Assumption 6.2 (iv) that:

$$\mathrm{e}^{-r(t+T)}\,\mathbb{E}_\mu\big[\|\eta_t^T\|^2_{W^{-4,3}}\big]\leq C\mathbb{E}_\mu\Big[1+|Y_{t+T}|^4\Big]\Big(\mathrm{e}^{-rT}+\frac{1-\mathrm{e}^{-r(t+T)}}{r}\Big)\leq C, \qquad (6.18)$$

where the constant C is finite and does not depend on Υ nor T. This completes the proof. ∎

We now turn to the proof of the central limit theorem stated in Proposition 6.4. To achieve this aim, we first prove the next Lemma on moment estimates.

Lemma 6.6. *We assume Assumption 6.2 and let $\Upsilon\in\mathbb{R}_+$.*
(i) We have:

$$\sup_{T\in\mathbb{R}_+}\sup_{t\leq\Upsilon}\mathbb{E}_\mu\big[\|\eta_t^T\|^2_{C^{-4,2}}\big]<+\infty. \qquad (6.19)$$

(ii) Let us denote by M^T_t the operator that associates $M^T_t(f)$ to f.

$$\sup_{T\in\mathbb{R}_+}\sup_{t\leq\Upsilon}\mathbb{E}_\mu\big[\|M^T_t\|^2_{W^{-4,3}}\big]<+\infty. \tag{6.20}$$

Proof. Let us first deal with (6.20). We consider the following linear forms: $D_{x,\sigma}(g)=\sigma(x)g'(x)$ and $D_{x,q}(g)=g(qx)+g((1-q)x)-g(x)$. Notice that for $g\in W^{4,3}\hookrightarrow C^{3,3}$, $x\in\mathbb{R}$ and $q\in[0,1]$,

$$\begin{aligned}|D_{x,\sigma}(g)|&=|\sigma(x)g'(x)|\leq\bar{\sigma}(1+|x|)|g'(x)|\leq C(1+|x|^4)\|g\|_{C^{3,3}}\leq C(1+|x|^4)\|g\|_{W^{4,3}},\\|D_{x,q}(g)|&=|g(qx)+g((1-q)x)-g(x)|\leq 3(1+|x|^3)\|g\|_{C^{3,3}}\leq C(1+|x|^3)\|g\|_{W^{4,3}},\end{aligned} \tag{6.21}$$

where C does not dependent on x nor on q. This implies that $D_{x,\sigma}$ and $D_{x,q}$ are continuous from $W^{4,3}$ into $\mathbb{R}$, and their norms in $W^{-4,3}$ are upper bounded by $C(1+|x|^4)$ and $C(1+|x|^3)$ respectively. Let us consider $(\varphi_p)_{p\in\mathbb{N}^*}$ a complete orthonormal basis of $W^{4,3}$ that are $\mathcal{C}^\infty$ with compact support. Using Riesz representation Theorem and Parseval's identity, we get

$$\sum_{p\geq1}D_{x,\sigma}(\varphi_p)^2=\|D_{x,\sigma}\|^2_{W^{-4,3}}\leq C(1+|x|^8)\quad\text{and}\quad\sum_{p\geq1}D_{x,q}(\varphi_p)^2=\|D_{x,q}\|^2_{W^{-4,3}}\leq C(1+|x|^6). \tag{6.22}$$

We have

$$\begin{aligned}\mathbb{E}_\mu\big[\sup_{t\leq\Upsilon}\|M^T_t\|^2_{W^{-4,3}}\big]&\leq\mathbb{E}_\mu\Big[\sum_{p\geq1}\sup_{t\leq\Upsilon}M^T_t(\varphi_p)^2\Big]\\&\leq 4\sum_{p\geq1}\mathbb{E}_\mu\Big[\langle M^T(\varphi_p)\rangle_\Upsilon\Big]\\&=4\int_T^{T+\Upsilon}ds\;\mathbb{E}_\mu\Big[\int_{\mathbb{R}}\frac{Z_s(dx)}{\mathbb{E}[N_s]}\Big(r\int_0^1\sum_{p\geq1}D_{x,q}(\varphi_p)^2\;G(dq)+2\sum_{p\geq1}D_{x,\sigma}(\varphi_p)^2\Big)\Big]\\&\leq C\int_T^{T+\Upsilon}ds\;\mathbb{E}_\mu\Big[\int_{\mathbb{R}}\frac{Z_s(dx)}{\mathbb{E}[N_s]}(1+|x|^8)\Big]\\&=C\int_T^{T+\Upsilon}ds\;\mathbb{E}_\mu\Big[(1+|Y_s|^8)\Big],\end{aligned} \tag{6.23}$$

where the first inequality comes from [?] Lemma 6.52, the second is Doob's inequality, the third line is a consequence of (6.5), the fourth inequality comes from the bounds (6.22) and the last equality comes from (3.1). The proof is then finished since by Assumption 6.2 (iv), $\sup_{t\geq0}\mathbb{E}_\mu[Y_t^8]<\infty$.

Let us now consider the proof of (6.19). Recall J defined by (6.3). It is clear that J is a bounded operator from $C^{4,2}$ into itself:

$$\|J\varphi\|_{C^{4,2}}\leq C\|\varphi\|_{C^{4,2}}, \tag{6.24}$$

where C does not depend on $\varphi\in C^{4,2}$.

Let us denote by $U(t)$ the semi-group of the diffusion with generator L given by (5.1). Proposition 3.9 in [?] and Assumptions 6.2 yield that for $\varphi\in C^{4,2}$ and $\psi\in C^{3,3}$

$$\sup_{t\leq\Upsilon}\|U(t)(\varphi)\|_{C^{4,2}}\leq C\|\varphi\|_{C^{4,2}}\text{ and }\sup_{t\leq\Upsilon}\|U(t)(\psi)\|_{C^{3,3}}\leq C\|\psi\|_{C^{3,3}} \tag{6.25}$$

where C does not depend on φ nor on ψ.

Let us consider the test function $\psi_t : (s,x) \mapsto U(t-s)\varphi(x)$ with $\varphi \in C^{4,2}$. Using Itô's formula:

$$\langle \eta^T_t, \varphi\rangle = \int_0^t \langle \eta^T_s, JU(t-s)\varphi\rangle ds + \int_0^t \langle dM^T_s, U(t-s)\varphi\rangle,$$

that is $\eta^T_t = \displaystyle\int_0^t U(t-s)^* J^* \eta^T_s \; ds + \int_0^t U(t-s)^* \; dM^T_s$, where $U(t-s)^*$ and J^* stand for the adjoint operators of $U(t-s)$ and J. We deduce that for $t \leq \Upsilon$:

$$\mathbb{E}_\mu\big[\|\eta^T_t\|^2_{C^{-4,2}}\big] \leq 2\Upsilon \int_0^t \mathbb{E}_\mu\Big[\|U(t-s)^* J^* \eta^T_s\|^2_{C^{-4,2}}\Big] ds + 2\mathbb{E}_\mu\Big[\| \int_0^t U(t-s)^* dM^T_s\|^2_{C^{-4,2}}\Big]. \quad (6.26)$$

Thanks to (6.24) and (6.25), we have for $s \leq t \leq \Upsilon$,

$$\mathbb{E}_\mu\Big[\|U(t-s)^* J^* \eta^T_s\|^2_{C^{-4,2}}\Big] ds \leq C\mathbb{E}_\mu\Big[\|\eta^T_s\|^2_{C^{-4,2}}\Big] ds. \quad (6.27)$$

The second term of the r.h.s. of (6.26) is upper bounded by considering the norm in $W^{-4,3}$. To prove that

$$\sup_{T\in\mathbb{R}_+} \sup_{t\leq\Upsilon} \mathbb{E}_\mu\Big[\| \int_0^t U(t-s)^* dM^T_s\|^2_{W^{-4,3}}\Big] < +\infty, \quad (6.28)$$

we use similar arguments as those used for the proof of (6.20) and (6.25). In the proof below, we replace the linear forms $D_{x,\sigma}$ and $D_{x,q}$ by $\bar{D}_{x,t-s,\sigma}$ and $\bar{D}_{x,t-s,q}$ with $\bar{D}_{x,t-s,\sigma}(\varphi) = D_{x,\sigma}(U(t-s)\varphi)$ and $\bar{D}_{x,t-s,q}(\varphi) = D_{x,q}(U(t-s)\varphi)$. Notice that by (6.21) for $g \in W^{4,3} \hookrightarrow C^{3,3}$, $x \in \mathbb{R}$ and $q \in [0,1]$,

$$\begin{aligned}
|\bar{D}_{x,t,\sigma}(g)| &= |D_{x,\sigma}(U(t)g)| \leq C(1+|x|^4)\|U(t)g\|_{C^{3,3}} \leq C(1+|x|^4)\|g\|_{C^{3,3}} \leq C(1+|x|^4)\|g\|_{W^{4,3}},\\
|\bar{D}_{x,t,q}(g)| &= |D_{x,q}(U(t)g)| \leq C(1+|x|^3)\|U(t)g\|_{C^{3,3}} \leq C(1+|x|^3)\|g\|_{W^{4,3}},
\end{aligned}$$

where C does not dependent on x. Using again Riesz representation Theorem and Parseval's identity, we get

$$\sum_{p\geq 1} \bar{D}_{x,t,\sigma}(\varphi_p)^2 = \|\bar{D}_x\|^2_{W^{-4,3}} \leq C(1+|x|^8) \quad \text{and} \quad \sum_{p\geq 1} \bar{D}_{x,t,q}(\varphi_p)^2 = \|\bar{D}_{x,q}\|^2_{W^{-4,3}} \leq C(1+|x|^6), \quad (6.29)$$

where C does not dependent on x nor on q. We have with the same arguments as in (6.23):

$$\begin{aligned}
\mathbb{E}_\mu[\sup_{t\leq\Upsilon} \| \int_0^t U(t-s)^* dM^T_s\|^2_{W^{-4,3}}] &\leq \mathbb{E}_\mu\Big[\sum_{p\geq 1} \sup_{t\leq\Upsilon} \int_0^t (U(t-s)\varphi_p)^2 dM^T_s\Big]\\
&\leq 4\sum_{p\geq 1} \mathbb{E}_\mu\Big[\int_0^\Upsilon (U(t-s)\varphi_p)^2 d\langle M^T\rangle_s\Big]\\
&= 4\int_T^{T+\Upsilon} ds\; \mathbb{E}_\mu\Big[\int_{\mathbb{R}} \left(r\int_0^1 \sum_{p\geq 1} \bar{D}_{x,t-s,q}(\varphi_p)^2 \; G(dq) + 2\sum_{p\geq 1} \bar{D}_{x,t-s,\sigma}(\varphi_p)^2 \right) \frac{Z_s(dx)}{\mathbb{E}[N_s]}\Big]\\
&\leq C\int_T^{T+\Upsilon} ds\; \mathbb{E}_\mu\Big[\int_{\mathbb{R}} \frac{Z_s(dx)}{\mathbb{E}[N_s]}(1+|x|^8)\Big]\\
&= C\int_T^{T+\Upsilon} ds\; \mathbb{E}_\mu\Big[(1+|Y_s|^8)\Big].
\end{aligned}$$

The proof is then done as $\sup_{t\geq 0}\mathbb{E}_\mu[Y_t^8]<\infty$ by Assumption 6.2 (iv).

Thus we get from (6.26), (6.27) and (6.28):

$$\mathbb{E}_\mu\big[\|\eta_t^T\|^2_{C^{-4,2}}\big]\leq C\left(1+\int_0^t\mathbb{E}_\mu\Big[\|\eta_s^T\|^2_{C^{-4,2}}\Big]ds\right).$$

We use Gronwall's Lemma and the fact that $\mathbb{E}_\mu\big[\|\eta_t^T\|^2_{C^{-4,2}}\big]$ is locally bounded (see Lemma 6.5) to conclude. ∎

We now prove the tightness of the fluctuation process.

Proposition 6.7. *Let $\Upsilon>0$. The sequence $(\eta^T)_{T\in\mathbb{R}_+}$ is tight in $\mathbb{D}([0,\Upsilon],W^{-7,1})$.*

We use a tightness criterion from [?], which we recall (see Lemma C p.217 in [?]).

Lemma 6.8. *(see Lemma C p.217 in [?])*
A sequence $(\Theta^T)_{T\in\mathbb{R}_+}$ of Hilbert H-valued càdlàg processes is tight in $\mathbb{D}([0,\Upsilon],H)$ if the following conditions are satisfied:
(i) There exists a Hilbert space H_0 such that $H_0\hookrightarrow_{H.S.}H$ and $\forall t\leq\Upsilon$, $\sup_{T\in\mathbb{R}_+}\mathbb{E}\big[\|\Theta_t^T\|^2_{H_0}\big]<+\infty$,
(ii) (Aldous condition) For every $\varepsilon>0$, there exists $\delta>0$ and $T_0\in\mathbb{R}_+$ such that for every sequence of stopping time $\tau_T\leq\Upsilon$,

$$\sup_{T>T_0}\sup_{\theta<\delta}\mathbb{P}\big(\|\Theta^T_{\tau_T+\theta}-\Theta^T_{\tau_T}\|_H>\varepsilon\big)<\varepsilon.$$

Proof of Prop. 6.7. We shall use Lemma 6.8 with $H_0=W^{-5,2}$ and $H=W^{-7,1}$. Condition (i) is a direct consequence of the uniform estimates obtained in (6.19) and of the fact that $\|\eta_t^T\|^2_{W^{-5,2}}\leq C\|\eta_t^T\|^2_{C^{-4,2}}$.

Let us now turn to condition (ii). By the Rebolledo criterion (see *e.g.* [?]), it is sufficient to show the Aldous condition for the the finite variation part and for the trace of the martingale part of (6.4). Let $(\varphi_p)_{p\geq1}$ be a complete orthonormal system of $W^{7,1}\hookrightarrow C^{6,1}$. We recall that the trace of the martingale part is defined as $\mathrm{tr}_{W^{-7,1}}\langle\!\langle M^T\rangle\!\rangle_t=\sum_{p\geq1}\langle M^T(\varphi_p)\rangle_t$ (see *e.g.* Joffe and Métivier [?]). Let $\varepsilon>0$ and $\tau_T\leq\Upsilon$ be a sequence of stopping times. For $T_0>0$ and $\delta>0$, following the steps of (6.23), we get:

$$\begin{aligned}&\sup_{T>T_0}\sup_{\theta<\delta}\mathbb{P}\Big(\big|\mathrm{tr}_{W^{-7,1}}\langle\!\langle M^T\rangle\!\rangle_{\tau_T+\theta}-\mathrm{tr}_{W^{-7,1}}\langle\!\langle M^T\rangle\!\rangle_{\tau_T}\big|>\varepsilon\Big)\\&\leq\sup_{T>T_0}\sup_{\theta<\delta}\frac{1}{\varepsilon}\mathbb{E}\Big[\int_{\tau_T}^{\tau_T+\theta}\langle\frac{Z_{s+T}}{\mathbb{E}[N_{s+T}]},r\int_0^1\sum_{p\geq1}D_{x,q}(\varphi_p)^2G(dq)+2\sum_{p\geq1}D_{x,\sigma}(\varphi_p)^2\rangle ds\Big]\end{aligned}\tag{6.30}$$

Using the embedding $W^{7,1}\hookrightarrow C^{6,1}$ and computations similar to (6.21), we obtain that:

$$\sum_{p\geq1}D_{x,q}(\varphi_p)^2=\|D_{x,q}\|^2_{W^{-7,1}}\leq C(1+|x|^2)\text{ and }\sum_{p\geq1}D_{x,\sigma}(\varphi_p)^2=\|D_{x,\sigma}\|^2_{W^{-7,1}}\leq C(1+|x|^4).$$

Thus (6.30) gives:

$$
\begin{aligned}
\sup_{T>T_0}\sup_{\theta<\delta}\mathbb{P}_\mu\Big(\big|\mathrm{tr}_{W^{-7,1}}\langle\!\langle M^T\rangle\!\rangle_{\tau_T+\theta}-\mathrm{tr}_{W^{-7,1}}\langle\!\langle M^T\rangle\!\rangle_{\tau_T}\big|>\varepsilon\Big)&\\
\leq&\frac{C}{\varepsilon}\sup_{T>T_0}\mathbb{E}_\mu\Big[\int_{\tau_T}^{\tau_T+\delta}\langle\frac{Z_{s+T}}{\mathbb{E}[N_{s+T}]},1+|x|^4\rangle\Big]ds\\
\leq&\frac{C}{\varepsilon}\sup_{T>T_0}\mathbb{E}_\mu\Big[\int_0^{\delta}\langle Z_{s+\tau_T+T},1+|x|^4\rangle\,\mathrm{e}^{-r(s+\tau_T+T)}\Big]ds\\
\leq&\frac{C}{\varepsilon}\sup_{T>T_0}\int_0^{\delta}\mathbb{E}_\mu\Big[\mathbb{E}_{Z_{\tau_T}}\Big[\langle Z_{s+T},1+|x|^4\rangle\,\mathrm{e}^{-r(s+T)}\Big]\Big]ds
\end{aligned}
\tag{6.31}
$$

by using the strong Markov property of $(Z_t)_{t\geq 0}$. Now, using the branching property:

$$
\begin{aligned}
\mathbb{E}_{Z_{\tau_T}}\Big[\langle Z_{s+T},1+|x|^4\rangle\,\mathrm{e}^{-r(s+T)}\Big]=&\int_{\mathbb{R}}\mathbb{E}_y\Big[\langle Z_{s+T},1+|x|^4\rangle\,\mathrm{e}^{-r(s+T)}\Big]Z_{\tau_T}(dy)\\
=&\int_{\mathbb{R}}\mathbb{E}_y\big[1+|Y_{s+T}|^4\big]Z_{\tau_T}(dy)\\
\leq&\int_{\mathbb{R}}\Big(\langle\pi,1+|x|^4\rangle+B(1+|y|^4)\,\mathrm{e}^{-\beta(s+T)}\Big)Z_{\tau_T}(dy).
\end{aligned}
\tag{6.32}
$$

for some β and $B>0$ given by (5.7) (see Remark 6.3 (ii)). Since we have a Yule tree, $\mathbb{E}[N_{\tau_T}]\leq\mathbb{E}[N_\Upsilon]=\exp(r\Upsilon)$. Moreover, using (2.22) where the integrand in the second term of the r.h.s. is negative for our choice $f(x)=|x|^4$ and noticing that Z_s is a positive measure, we obtain with localizing arguments that for any $t\in\mathbb{R}_+$:

$$
\mathbb{E}_\mu\big[\langle Z_{t\wedge\tau_T},1+|x|^4\rangle\big]\leq\langle\mu,1+|x|^4\rangle+\int_0^t(8\bar{b}+24\bar{\sigma})\mathbb{E}_\mu[\langle Z_{s\wedge\tau_T},1+|x|^4\rangle]ds.
$$

We deduce from Gronwall's lemma that:

$$
\mathbb{E}_\mu\big[\langle Z_{\tau_T},1+|x|^4\rangle\big]\leq\langle\mu,1+|x|^4\rangle\,\mathrm{e}^{(8\bar{b}+24\bar{\sigma})\Upsilon}.\tag{6.33}
$$

Then (6.31), (6.32) and (6.33) imply that:

$$
\sup_{T>T_0}\sup_{\theta<\delta}\mathbb{P}_\mu\Big(\big|\mathrm{tr}_{W^{-7,1}}\langle\!\langle M^T\rangle\!\rangle_{\tau_T+\theta}-\mathrm{tr}_{W^{-7,1}}\langle\!\langle M^T\rangle\!\rangle_{\tau_T}\big|>\varepsilon\Big)\leq\frac{C\delta}{\varepsilon}\big(\,\mathrm{e}^{r\Upsilon}+\mathrm{e}^{(8\bar{b}+24\bar{\sigma})\Upsilon}\,\big)\tag{6.34}
$$

which finishes the proof of the Aldous inequality for the trace of the martingale.

Remark 6.9. Notice that this also shows that $(M^T)_{T\geq 0}$ is tight in $W^{-7,1}$.

For the finite variation part:

$$\begin{aligned}
&\sup_{T>T_0}\sup_{\theta<\delta}\mathbb{P}\Big(\|\int_0^{\tau_T+\theta}(L+J)^*\eta_s^T ds-\int_0^{\tau_T}(L+J)^*\eta_s^T ds\|_{W^{-7,1}}>\varepsilon\Big)\\
&\qquad\le \sup_{T>T_0}\sup_{\theta<\delta}\frac{1}{\varepsilon^2}\mathbb{E}\Big[\|\int_{\tau_T}^{\tau_T+\theta}(L+J)^*\eta_s^T ds\|^2_{W^{-7,1}}\Big]\\
&\qquad\le \sup_{T>T_0}\sup_{\theta<\delta}\frac{\theta}{\varepsilon^2}\mathbb{E}\Big[\int_{\tau_T}^{\tau_T+\theta}\|(L+J)^*\eta_s^T\|^2_{W^{-7,1}}ds\Big]\\
&\qquad\le \sup_{T>T_0}\frac{C\delta}{\varepsilon^2}\int_0^{\Upsilon+\delta}\mathbb{E}\Big[\|\eta_t^T\|^2_{C^{-4,2}}\Big]dt\\
&\qquad\le \frac{C\delta(\Upsilon+\delta)}{\varepsilon^2}\sup_{T>T_0}\sup_{t\le\Upsilon}\mathbb{E}\Big[\|\eta_t^T\|^2_{C^{-4,2}}\Big].
\end{aligned}\tag{6.35}$$

We use Cauchy-Schwarz' inequality for the second inequality. The third inequality is obtained by noticing that under the Assumption 6.2 and for $\varphi\in W^{7,1}$:

$$\|L\varphi\|_{C^{4,2}}\le C\|\varphi\|_{C^{6,1}}\le C\|\varphi\|_{W^{7,1}}\tag{6.36}$$

as $W^{7,1}\hookrightarrow C^{6,1}$. We can make the r.h.s. of (6.35) as small as we wish thanks to (6.19), and this ends the proof of the tightness. ■

Then, we identify the limit by showing that the limiting values solve an equation for which uniqueness holds. This will prove the central limit theorem.

Proof of Proposition 6.4. First of all, by Remark 6.9, the sequence of martingales $(M^T)_{T\ge0}$ is tight in $W^{-7,1}$ and thus also in $C^{-7,0}$ by (6.8). Let us prove that in the latter space, it is moreover C-tight in the sense of Jacod and Shiryaev [?] p.315. Using the Proposition 3.26 (iii) of this reference, it remains to prove the convergence of $\sup_{t\le\Upsilon}\|\Delta M_t^T\|_{C^{-7,0}}$ to 0 where $\Delta M_t^T=M_t^T-M_{t_-}^T$. Since the finite variation part of (6.4) is continuous, $\Delta M_t^T=\Delta\eta_t^T$, and since in (6.1) $t\mapsto\langle Z_t,Q_tf\rangle$ is continuous, we have for $f\in C^{7,0}$:

$$\sup_{t\le\Upsilon}|\Delta M_t^T(f)|=\sup_{t\le\Upsilon}\mathrm{e}^{-\frac{r(t+T)}{2}}\,|f\big(q(\omega,t+T)X_{t+T}^{u(\omega,t+T)}\big)+f\big((1-q(\omega,t+T))X_{t+T}^{u(\omega,t+T)}\big)-f\big(X_{t+T}^{u(\omega,t+T)}\big)|\tag{6.37}$$

where $u(\omega,t+T)\in V_{t+T}$ is the label of the particle that undergoes division at $t+T$, and where $q(\omega,t+T)$ is the fraction which appears in the splitting. By convention, if there is no splitting at $t+T$, the term in the supremum of the r.h.s. of (6.37) is 0. Thus $\sup_{t\le\Upsilon}|\Delta M_t^T(f)|\le 3\,\mathrm{e}^{-rT/2}\,\|f\|_\infty\le 3\,\mathrm{e}^{-rT/2}\,\|f\|_{C^{7,0}}$. This proves that:

$$\sup_{t\le\Upsilon}\|\Delta M_t^T\|_{C^{-7,0}}\le 3\,\mathrm{e}^{-rT/2},\tag{6.38}$$

which converges a.s. to 0 when $T\to+\infty$. This finishes the proof of the C-tightness of $(M^T)_{T\ge0}$ in $C^{-7,0}$. The inequality (6.38) also ensures that the sequence $\sup_{t\le\Upsilon}\|\Delta M_t^T\|_{W^{-7,1}}$ is uniformly integrable. From the LLN of Proposition 4.2, the integrand of (6.5) converges to $W\times V(f)$ which does not depend on s any more. Since W is $\cap_{\varepsilon>0}\sigma(\eta_\varepsilon)$-measurable, it follows that W and $\mathcal{W}$ are independent. Thus, using Theorem 3.12 p. 432 in [?], we obtain that $(M^T)_{T\ge0}$ converges in distribution in

$\mathbb{D}([0,\Upsilon], C^{-7,0})$ to a Gaussian process $\mathcal{W}$ with the announced quadratic variation.

By Proposition 6.7, the sequence $(\eta^T)_{T\geq 0}$ is tight in $W^{-7,1}$ and hence also in $C^{-7,0}$ by (6.8). Let η be an accumulation point in $\mathbb{D}([0,\Upsilon], C^{-7,0})$. Because of (6.4) and (6.38), η is almost surely a continuous process. Let us call again by $(\eta^T)_{T\geq 0}$, with an abuse of notation, the subsequence that converges in law to η. Since η is continuous, we get from (6.4) that it solves (6.9). Using Gronwall's inequality, we obtain that this equation admits in $\mathcal{C}([0,\Upsilon], C^{-7,0})$ a unique solution for a given Gaussian white noise W which is in $C^{-7,0}$. This achieves the proof. ■

Acknowledgements: This work has been financed by the ANR MAEV (ANR-08-BLAN-0190) and Viroscopy (ANR-08-SYSC-016-03).