

DETECTION OF CELLULAR AGING IN A GALTON-WATSON PROCESS

JEAN-FRANÇOIS DELMAS AND LAURENCE MARSALLE

ABSTRACT. We consider the bifurcating Markov chain model introduced by Guyon to detect cellular aging from cell lineage. To take into account the possibility for a cell to die, we use an underlying Galton-Watson process to describe the evolution of the cell lineage. We give in this more general framework a weak law of large number, an invariance principle and thus fluctuation results for the average over one generation or up to one generation. We also prove the fluctuations over each generation are independent. Then we present the natural modifications of the tests given by Guyon in cellular aging detection within the particular case of the auto-regressive model.

1. INTRODUCTION

This work is motivated by experiments done by biologists on *Escherichia coli*, see Stewart and al. [16]. *E. coli* is a rod-shaped single celled organism which reproduces by dividing in the middle. It produces a new end per progeny cell. We shall call this new end the new pole whereas the other end will be called the old pole. The age of a cell is given by the age of its old pole (i.e. the number of generations in the past of the cell before the old pole was produced). Notice that at each generation a cell gives birth to 2 cells which have a new pole and one of the two cells has an old pole of age 1 (which corresponds to the new pole of its mother), while the other has an old pole with age larger than one (which corresponds to the old pole of its mother). The former is called the new pole daughter and the latter the old pole daughter. Experimental data, see [16], suggest strongly that the growth rate of the new pole daughter is significantly larger than the growth rate of the old pole daughter. For asymmetric aging see also Ackermann [2] for an other case of asymmetric division, and Lindner and al. [12] or Ackermann and al. [1] on asymmetric damage repartition.

Guyon [9] studied a mathematical model, called bifurcation Markov chain (BMC), of an asymmetric Markov chain on a regular binary tree. This model allows to represent an asymmetric repartition for example of the growth rate of a cell between new pole and old pole daughters. Using this model, Guyon provides tests to detect a difference of the growth rate between new pole and old pole on a single experimental data set, whereas in [16] averages over many experimental data sets have to be done to detect this difference. In the BMC model, cells are assumed to never die (a death corresponds to no more division). Indeed few death appear in normal nutriment saturated conditions. However, under stress condition, dead cells can represent a significant part of the population. It is therefore natural to take this random effect into account by using a Galton-Watson (GW) process. Our purpose is to study a model of bifurcating Markov chains on a Galton Watson tree instead of a regular tree. Notice that inferences on symmetric bifurcating processes on regular trees have been studied, see the survey of Hwang, Basawa and Yea [11] and the seminal work of Cowan and Staudte [7]. We also learned of a recent independent work on inferences for asymmetric auto-regressive models by

Date: July 4, 2008.

2000 Mathematics Subject Classification. 60F05, 60J80, 92D25, 62M05.

Key words and phrases. Aging, Galton-Watson process, bifurcating Markov process, stable convergence.

Bercu, De Saporta and Gégout-Petit [6]. Other models on cell lineage with differentiation have been investigated, see for example Bansaye [5, 4] on parasite infection and Evans and Steinsaltz [8] on asymptotic models relying on super-Brownian motion.

1.1. The statistical model. In order to study the behavior of the growth rate of cells in [16], we set some notations: we index the genealogical tree by the regular binary tree $\mathbb{T} = \{\emptyset\} \cup \bigcup_{k \in \mathbb{N}^*} \{0, 1\}^k$; \emptyset is the label of the ancestor and if i denotes a cell, let $i0$ denote the new pole progeny cell, and $i1$ the old pole progeny cell. The growth rate of cell i is X_i . When the mother gives birth to two cells among which a unique one divides, we consider that the cell which doesn't divide doesn't grow. We work with the following model (see Section 1.2 for a more general model of BMC on GW tree):

- With probability $p_{1,0}$, i gives birth to two cells $i0$ and $i1$ which will both divide. The growth rates of the daughters X_{i0} and X_{i1} are then linked to the mother's one X_i through the following auto-regressive equations

$$(1) \quad \begin{cases} X_{i0} &= \alpha_0 X_i + \beta_0 + \varepsilon_{i0} \\ X_{i1} &= \alpha_1 X_i + \beta_1 + \varepsilon_{i1}, \end{cases}$$

where $\alpha_0, \alpha_1 \in (-1, 1)$, $\beta_0, \beta_1 \in \mathbb{R}$ and $((\varepsilon_{i0}, \varepsilon_{i1}), i \in \mathbb{T})$ is a sequence of independent centered bi-variate Gaussian random variables, with covariance matrix

$$\sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad \sigma^2 > 0, \quad \rho \in (-1, 1).$$

- With probability p_0 , only the new pole $i0$ divides. Its growth rate X_{i0} is linked to its mother's one X_i through the relation

$$(2) \quad X_{i0} = \alpha'_0 X_i + \beta'_0 + \varepsilon'_{i0},$$

where $\alpha'_0 \in (-1, 1)$, $\beta'_0 \in \mathbb{R}$ and $(\varepsilon'_{i0}, i \in \mathbb{T})$ is a sequence of independent centered Gaussian random variables with variance $\sigma_0^2 > 0$.

- With probability p_1 , only the old pole $i1$ divides. Its growth rate X_{i1} is linked to its mother's one through the relation

$$(3) \quad X_{i1} = \alpha'_1 X_i + \beta'_1 + \varepsilon'_{i1},$$

where $\alpha'_1 \in (-1, 1)$, $\beta'_1 \in \mathbb{R}$ and $(\varepsilon'_{i1}, i \in \mathbb{T})$ is a sequence of independent centered Gaussian random variables with variance $\sigma_1^2 > 0$.

- The sequences $((\varepsilon_{i0}, \varepsilon_{i1}), i \in \mathbb{T})$, $(\varepsilon'_{i0}, i \in \mathbb{T})$ and $(\varepsilon'_{i1}, i \in \mathbb{T})$ are independent.

In Section 6, we first compute the maximum likelihood estimator (MLE) of the parameter

$$(4) \quad \theta = (\alpha_0, \beta_0, \alpha_1, \beta_1, \alpha'_0, \beta'_0, \alpha'_1, \beta'_1, p_{1,0}, p_0, p_1)$$

and of $\kappa = (\sigma, \rho, \sigma_0, \sigma_1)$. Then, we prove that they are strongly consistent and that the MLE of θ is asymptotically normal, see Proposition 6.3 and Remark 6.4. Notice that the MLE of $(p_{1,0}, p_0, p_1)$, which is computed only on the underlying GW tree, was already known, see for example [13]. Eventually, we explicit a test for aging detection, for instance the null hypothesis $\{(\alpha_0, \beta_0) = (\alpha_1, \beta_1)\}$ against its alternative $\{(\alpha_0, \beta_0) \neq (\alpha_1, \beta_1)\}$, see Proposition 6.7. It appears that, for those hypothesis, using the test statistic from [9] with incomplete data due to death cells instead of the test statistic from Proposition 6.7 is not conservative, see Remark 6.8.

To prove those results, we shall consider a more general framework of BMC which is described in Section 1.2. An important tool is the auxiliary Markov chain which is defined in

Section 1.3. Eventually easy to read version of our main general results are given in Section 1.4.

1.2. The mathematical model of bifurcating Markov chain (BMC). We first introduce some notations related to the regular binary tree. Let $\mathbb{G}_0 = \{\emptyset\}$, $\mathbb{G}_k = \{0, 1\}^k$ for $k \in \mathbb{N}^*$, $\mathbb{T}_r = \bigcup_{0 \leq k \leq r} \mathbb{G}_k$. The new (resp. old) pole daughter of a cell $i \in \mathbb{T}$ is denoted by $i0$ (resp. $i1$) and 0 (resp. 1) if $i = \emptyset$ is the initial cell or root of the tree. The set \mathbb{G}_k corresponds to all possible cells in the k -th generation. We denote by $|i|$ the generation of i ($|i| = k$ if and only if $i \in \mathbb{G}_k$).

For a cell $i \in \mathbb{T}$, let X_i denote a quantity of interest (for example its growth rate). We assume that the quantity of interest of the daughters of a cell i , conditionally on the generations previous to i , depends only on X_i . This property is stated using the formalism of BMC. More precisely, let (E, \mathcal{E}) be a measurable space, P a probability kernel on $E \times \mathcal{E}^2$ with values in $[0, 1]$: $P(\cdot, A)$ is measurable for all $A \in \mathcal{E}^2$ and $P(x, \cdot)$ is a probability measure on (E^2, \mathcal{E}^2) , and for any measurable real-valued bounded function g defined on E^3 we set

$$Pg(x) = \int_{E^2} g(x, y, z) P(x, dy, dz).$$

Definition 1.1. We say a stochastic process indexed by \mathbb{T} , $X = (X_i, i \in \mathbb{T})$, is a bifurcating Markov chain on a measurable space (E, \mathcal{E}) with initial distribution ν and probability kernel P , a P -BMC in short, if:

- X_{\emptyset} is distributed as ν .
- For any measurable real-valued bounded functions $(g_i, i \in \mathbb{T})$ defined on E^3 , we have for all $k \geq 0$,

$$\mathbb{E} \left[\prod_{i \in \mathbb{G}_k} g_i(X_i, X_{i0}, X_{i1}) | \sigma(X_j; j \in \mathbb{T}_k) \right] = \prod_{i \in \mathbb{G}_k} Pg_i(X_i).$$

We consider a metric measurable space (S, \mathcal{S}) and add a cemetery point to S , ∂ . Let $\bar{S} = S \cup \{\partial\}$ and $\bar{\mathcal{S}}$ be the σ -field generated by \mathcal{S} and $\{\partial\}$. (In the biological framework of the previous Section, S corresponds to the state space of the quantities of interest and ∂ is the default value for dead cells.) Let P^* be a probability kernel defined on $\bar{S} \times \bar{\mathcal{S}}^2$ such that

$$(5) \quad P^*(\partial, \{(\partial, \partial)\}) = 1.$$

Notice that this condition means that ∂ is an absorbing state. (In the biological framework of the previous Section, condition (5) states that no dead cell can give birth to a living cell.)

Definition 1.2. Let $X = (X_i, i \in \mathbb{T})$ be a P^* -BMC on $(\bar{S}, \bar{\mathcal{S}})$, with P^* satisfying (5). We call $(X_i, i \in \mathbb{T}^*)$, with $\mathbb{T}^* = \{i \in \mathbb{T}, X_i \neq \partial\}$, a bifurcating Markov chain on a Galton-Watson tree. The P^* -BMC is said spatially homogeneous if $p_{1,0} = P^*(x, S \times S)$, $p_0 = P^*(x, S \times \{\partial\})$ and $p_1 = P^*(x, \{\partial\} \times S)$ do not depend on $x \in S$. A spatially homogeneous P^* -BMC is said super-critical if $m > 1$ where $m = 2p_{1,0} + p_1 + p_0$.

Notice that condition (5) and the spatially homogeneity property implies that \mathbb{T}^* is a GW tree. This justify the name of BMC on a Galton-Watson tree. The GW tree is super-critical if and only if $m > 1$. From now on, we shall only consider super-critical spatially homogeneous P^* -BMC on a Galton-Watson tree. (In the biological framework of the previous Section, \mathbb{T}^* denotes the sub-tree of living cells and the notations $p_{1,0}$, p_0 and p_1 are consistent since, for instance, $P^*(x, S \times S)$ represents the probability that a living cell with growing rate x gives birth to two living cells.)

We now consider the Galton-Watson sub-tree \mathbb{T}^* . For any subset $J \subset \mathbb{T}$, let

$$(6) \quad J^* = J \cap \mathbb{T}^* = \{j \in J; X_j \neq \partial\}$$

be the subset of J of living cells and $|J|$ be the cardinal of J . The process $Z = (Z_k, k \in \mathbb{N})$, where $Z_k = |\mathbb{G}_k^*|$, is a GW process with reproduction generating function

$$\psi(z) = (1 - p_0 - p_1 - p_{1,0}) + (p_0 + p_1)z + p_{1,0}z^2.$$

Notice the average number of daughters alive is m . We have, for $k \geq 0$,

$$(7) \quad \mathbb{E}[|\mathbb{G}_k^*|] = m^k \quad \text{and} \quad \mathbb{E}[|\mathbb{T}_r^*|] = \sum_{q=0}^r \mathbb{E}[|\mathbb{G}_q^*|] = \sum_{q=0}^r m^q = \frac{m^{r+1} - 1}{m - 1}.$$

Let us recall some well-known facts on super-critical GW, see e.g. [10] or [3]. The extinction probability of the GW process Z is $\eta = \mathbb{P}(|\mathbb{T}^*| < \infty) = 1 - \frac{m-1}{p_{1,0}}$. There exists a random variable W s.t.

$$(8) \quad W = \lim_{q \rightarrow \infty} m^{-q} |\mathbb{G}_q^*| \quad \text{a.s. and in } L^2,$$

$\mathbb{P}(W = 0) = \eta$ and whose Laplace transform, $\varphi(\lambda) = \mathbb{E}[e^{-\lambda W}]$, satisfies $\varphi(\lambda) = \psi(\varphi(\lambda/m))$ for $\lambda \geq 0$. Notice the distribution of W is completely characterized by this functional equation and $\mathbb{E}[W] = 1$.

For $i \in \mathbb{T}$, we set $\Delta_i = (X_i, X_{i_0}, X_{i_1})$, the mother-daughters quantities of interest. For a finite subset $J \subset \mathbb{T}$, we set

$$(9) \quad M_J(f) = \begin{cases} \sum_{i \in J} f(X_i) & \text{for } f \in \mathcal{B}(\bar{S}), \\ \sum_{i \in J} f(\Delta_i) & \text{for } f \in \mathcal{B}(\bar{S}^3), \end{cases}$$

with the convention that $M_\emptyset(f) = 0$, and the following two averages of f over J

$$(10) \quad \bar{M}_J(f) = \frac{1}{|J|} M_J(f) \quad \text{if } |J| > 0 \quad \text{and} \quad \tilde{M}_J(f) = \frac{1}{\mathbb{E}[|J|]} M_J(f) \quad \text{if } \mathbb{E}[|J|] > 0.$$

We shall study the asymptotic limit of the averages of a function f for the BMC over the n -th generation, $\bar{M}_{\mathbb{G}_n^*}(f)$ and $\tilde{M}_{\mathbb{G}_n^*}(f)$, or over all the generations up to the n -th, $\bar{M}_{\mathbb{T}_n^*}(f)$ and $\tilde{M}_{\mathbb{T}_n^*}(f)$, as n goes to infinity. Notice the no death case studied in [9] corresponds to $p_{1,0} = 1$ that is $m = 2$.

1.3. The auxiliary Markov chain. We define the sub-probability kernel on $S \times \mathcal{S}^2$ $P(\cdot, \cdot) = P^*(\cdot, \cdot \cap \mathcal{S}^2)$ and two sub-probability kernels on $S \times \mathcal{S}$:

$$P_0^* = P^*(\cdot, (\cdot \cap S) \times \bar{S}) \quad \text{and} \quad P_1^* = P(\cdot, \bar{S} \times (\cdot \cap S)).$$

Notice that P_0^* (resp. P_1^*) is the restriction of the first (resp. second) marginal of P^* to S . From spatial homogeneity, we have for all $x \in S$, $P(x, \mathcal{S}^2) = p_{1,0}$ and, for $\delta \in \{0, 1\}$,

$$P_\delta^*(x, \{\partial\}) = 0 \quad \text{and} \quad P_\delta^*(x, S) = p_\delta + p_{1,0}.$$

We introduce an auxiliary Markov chain (see Guyon [9] for the case $m = 2$). Let $Y = (Y_n, n \in \mathbb{N})$ be a Markov chain on S with Y_0 distributed as X_\emptyset and transition kernel

$$Q = \frac{1}{m}(P_0^* + P_1^*).$$

The distribution of Y_n corresponds intuitively to the distribution of X_I conditionally on $I \in \mathbb{T}^*$, where I is chosen at random in \mathbb{G}_n , see Lemma 2.1 for a precise statement. We shall write \mathbb{E}_x when $X_\emptyset = x$ (i.e. ν is the Dirac mass at $x \in S$).

Last, we need some more notations : if (E, \mathcal{E}) is a metric measurable space, then $\mathcal{B}_b(E)$ (resp. $\mathcal{B}_+(E)$) denotes the set of bounded (resp. non-negative) real-valued measurable functions defined on E . The set $\mathcal{C}_b(E)$ (resp. $\mathcal{C}_+(E)$) denotes the set of bounded (resp. non-negative) real-valued continuous functions defined on E . For a finite measure λ on (E, \mathcal{E}) and $f \in \mathcal{B}_b(E) \cup \mathcal{B}_+(E)$ we shall write $\langle \lambda, f \rangle$ for $\int f(x) d\lambda(x)$.

We consider the following hypothesis (H) :

The Markov chain Y is ergodic, that is there exists a probability measure μ on (S, \mathcal{S}) s.t., for all $f \in \mathcal{C}_b(S)$ and all $x \in S$, $\lim_{k \rightarrow \infty} \mathbb{E}_x[f(Y_k)] = \langle \mu, f \rangle$.

Notice that under (H) , the probability measure μ is the unique stationary distribution of Y and $(Y_n, n \in \mathbb{N})$ converges in distribution to μ .

1.4. The main results. We can now state our principal results on the weak law of large numbers and fluctuations for the averages over a generation or up to a generation. Those results are a particular case of the more general statements given in Theorem 3.4 and Theorem 5.2, using Remark 2.2.

Theorem 1.3. *Let $(X_i, i \in \mathbb{T}^*)$ be a super-critical spatially homogeneous P^* -BMC on a GW tree and W be defined by (8). We assume that (H) holds and that $x \mapsto P^*g(x) \in \mathcal{C}_b(\bar{S})$ for all $g \in \mathcal{C}_b(\bar{S}^3)$. Let $f \in \mathcal{C}_b(\bar{S}^3)$.*

- **Weak law of large numbers.** *We have the following convergence in probability:*

$$\mathbf{1}_{\{|\mathbb{G}_r^*| > 0\}} \frac{1}{|\mathbb{G}_r^*|} \sum_{i \in \mathbb{G}_r^*} f(\Delta_i) \xrightarrow[r \rightarrow \infty]{\mathbb{P}} \langle \mu, P^*f \rangle \mathbf{1}_{\{W \neq 0\}},$$

$$\mathbf{1}_{\{|\mathbb{T}_r^*| > 0\}} \frac{1}{|\mathbb{T}_r^*|} \sum_{i \in \mathbb{T}_r^*} f(\Delta_i) \xrightarrow[r \rightarrow \infty]{\mathbb{P}} \langle \mu, P^*f \rangle \mathbf{1}_{\{W \neq 0\}}.$$

- **Fluctuations.** *We have the following convergence in distribution:*

$$\mathbf{1}_{\{|\mathbb{G}_r^*| > 0\}} \frac{1}{\sqrt{|\mathbb{T}_r^*|}} \sum_{i \in \mathbb{T}_r^*} \left(f(\Delta_i) - P^*f(X_i) \right) \xrightarrow[r \rightarrow \infty]{(d)} \mathbf{1}_{\{W \neq 0\}} \sigma G,$$

where $\sigma^2 = \langle \mu, P^*(f^2) - (P^*f)^2 \rangle$ and G is a Gaussian random variable with mean zero, variance 1, and independent of W .

One can get the strong law of large numbers under stronger hypothesis on Y (such as geometric ergodicity) using similar arguments as in [9]. We also can prove that the fluctuations over each generation are asymptotically independent.

Theorem 1.4. *Let $(X_i, i \in \mathbb{T}^*)$ be a super-critical spatially homogeneous P^* -BMC on a GW tree. We assume that (H) holds and that $x \mapsto P^*g(x) \in \mathcal{C}_b(\bar{S})$ for all $g \in \mathcal{C}_b(\bar{S}^3)$. Let $d \geq 1$, and for $\ell \in \{1, \dots, d\}$, $f_\ell \in \mathcal{C}_b(\bar{S}^3)$ and $\sigma_\ell^2 = \langle \mu, P^*(f_\ell^2) - (P^*f_\ell)^2 \rangle$. We set for $f \in \mathcal{C}_b(\bar{S}^3)$*

$$N_n(f) = \frac{1}{\sqrt{|\mathbb{G}_n^*|}} \sum_{i \in \mathbb{G}_n^*} \left(f(\Delta_i) - P^*f(X_i) \right).$$

Then we have the following convergence in distribution:

$$(N_n(f_1), \dots, N_{n-d+1}(f_d)) \mathbf{1}_{\{|\mathbb{G}_n^*| > 0\}} \xrightarrow[n \rightarrow \infty]{(d)} \mathbf{1}_{\{W \neq 0\}} (\sigma_1 G_1, \dots, \sigma_d G_d),$$

where G_1, \dots, G_d are independent Gaussian random variables with mean zero and variance 1 and are independent of W given by (8).

Even if the results on fluctuations in Theorem 1.3 are not complete, see the Remark 1.5 below, they are still sufficient to study the statistical model we gave in Section 1.1 for the detection of cellular aging from cell lineage when death of cells can occur.

Remark 1.5. Let $V = (V_r, r \geq 0)$ be a Markov chain on a finite state space. We assume V is irreducible, with transition matrix R and unique invariant probability μ . Then it is well known, see [14], that $\frac{1}{r} \sum_{i=1}^r h(V_i)$ converges a.s. to $\langle \mu, h \rangle$ and that, to prove the fluctuations result, one solves the Poisson equation $H - RH = h - \langle \mu, h \rangle$, writes

$$(11) \quad \frac{1}{\sqrt{r}} \sum_{i=1}^r (h(V_i) - \langle \mu, h \rangle) = \frac{1}{\sqrt{r}} \sum_{i=1}^r (H(V_i) - RH(V_{i-1})) + \frac{1}{\sqrt{r}} RH(V_0) - \frac{1}{\sqrt{r}} RH(V_r),$$

and then uses martingale theory (we use similar techniques to prove the fluctuations in Theorem 1.3) to obtain the asymptotic normality of $\frac{1}{\sqrt{r}} \sum_{i=1}^r H(V_i) - RH(V_{i-1})$. It then only remains to say that $\frac{1}{\sqrt{r}} RH(V_0)$ and $\frac{1}{\sqrt{r}} RH(V_r)$ converge to 0 to conclude.

Assume that hypothesis of Theorem 1.3 hold and that $x \mapsto P^*(x, A)$ is continuous for all $A \in \mathcal{B}(\bar{S}^2)$. Let $h \in \mathcal{C}_b(\bar{S})$. Theorem 1.3 implies that $\mathbf{1}_{\{|\mathbb{G}_r^*| > 0\}} \frac{1}{|\mathbb{T}_r^*|} \sum_{i \in \mathbb{T}_r^*} h(X_i)$ converges in probability to $\langle \mu, h \rangle \mathbf{1}_{\{W \neq 0\}}$. To get the fluctuations, that is the limit of

$$\mathbf{1}_{\{|\mathbb{G}_r^*| > 0\}} \frac{1}{\sqrt{|\mathbb{T}_r^*|}} \sum_{i \in \mathbb{T}_r^*} (h(X_i) - \langle \mu, h \rangle)$$

as r goes to infinity, using martingale theory, one can think of using the same kind of approach in order to use the result on fluctuations of Theorem 1.3. But then notice that what will correspond to the boundary term in (11) at time r , $\frac{1}{\sqrt{r}} RH(V_r)$, will now be a boundary term over the last generation \mathbb{G}_r^* whose cardinal is of the same order as $|\mathbb{T}_r^*|$. Thus the order of the boundary term is not negligible, which unable us to conclude.

The fluctuations for $\sum_{i \in \mathbb{T}_r^*} h(X_i)$ are still an open question.

1.5. Organization of the paper. We quickly study the auxiliary chain in Section 2. We state the first result on the weak law of large number in Section 3. Section 4 is devoted to some preparatory results in order to apply results on fluctuations for martingale. Our main result, Theorem 5.2, is stated and proved in Section 5. The biological model of Section 1.1 is analyzed in Section 6.

2. PRELIMINARY RESULT AND NOTATIONS

Recall the Markov chain Y defined in Section 1.3.

Lemma 2.1. *We have, for $f \in \mathcal{B}_b(S) \cup \mathcal{B}_+(S)$,*

$$(12) \quad \mathbb{E}[f(Y_n)] = m^{-n} \sum_{i \in \mathbb{G}_n} \mathbb{E}[f(X_i) \mathbf{1}_{\{i \in \mathbb{T}^*\}}] = \frac{\sum_{i \in \mathbb{G}_n} \mathbb{E}[f(X_i) \mathbf{1}_{\{i \in \mathbb{T}^*\}}]}{\sum_{i \in \mathbb{G}_n} \mathbb{P}(i \in \mathbb{T}^*)} = \mathbb{E}[f(X_I) | I \in \mathbb{T}^*],$$

where I is a uniform random variable on \mathbb{G}_n independent of X .

Proof. We consider the first equality. Recall that Y_0 has distribution ν . For $i = i_1 \dots i_n \in \mathbb{G}_n$, we have, thanks to (5) and the definition of P^* ,

$$\mathbb{E}[f(X_i)\mathbf{1}_{\{i \in \mathbb{T}^*\}}] = \mathbb{E}[f(X_i)\mathbf{1}_{\{X_i \neq \partial\}}] = \langle \nu, (P_{i_1}^* \dots P_{i_n}^*) f \rangle,$$

so that

$$\begin{aligned} \sum_{i \in \mathbb{G}_n} \mathbb{E}[f(X_i)\mathbf{1}_{\{i \in \mathbb{T}^*\}}] &= \sum_{i_1, \dots, i_n \in \{0,1\}} \langle \nu, (P_{i_1}^* \dots P_{i_n}^*) f \rangle \\ &= \langle \nu, (P_0^* + P_1^*)^n f \rangle = m^n \langle \nu, Q^n f \rangle = m^n \mathbb{E}[f(Y_n)]. \end{aligned}$$

This gives the first equality. Then take $f = 1$ in the previous equality to get $m^n = \sum_{i \in \mathbb{G}_n} \mathbb{P}(i \in \mathbb{T}^*)$ and the second equality of (12). The last equality of (12) is obvious. \square

We recall that ν denotes the distribution of X_\emptyset . Any function f defined on S is extended to \bar{S} by setting $f(\partial) = 0$. Let F be a vector subspace of $\mathcal{B}(S)$ s.t.

- (i) F contains the constants;
- (ii) $F^2 := \{f^2; f \in F\} \subset F$;
- (iii) – $F \otimes F \subset L^1(P(x, \cdot))$ for all $x \in S$ and $P(f_0 \otimes f_1) \in F$ for all $f_0, f_1 \in F$;
– For $\delta \in \{0, 1\}$, $F \subset L^1(P_\delta^*(x, \cdot))$ for all $x \in S$ and $P_\delta^*(f) \in F$ for all $f \in F$;
- (iv) There exists a probability measure μ on (S, \mathcal{S}) s.t. $F \subset L^1(\mu)$ and $\lim_{n \rightarrow \infty} \mathbb{E}_x[f(Y_n)] = \langle \mu, f \rangle$ for all $x \in S$ and $f \in F$;
- (v) For all $f \in F$, there exists $g \in F$ s.t. for all $r \in \mathbb{N}$, $|Q^r f| \leq g$;
- (vi) $F \subset L^1(\nu)$.

By convention a function defined on \bar{S} is said to belong to F if its restriction to S belongs to F .

Remark 2.2. Notice that if (H) is satisfied and if $x \mapsto P^*g(x)$ is continuous on S for all $g \in \mathcal{C}_b(\bar{S}^3)$ then the set $\mathcal{C}_b(S)$ fulfills (i) – (vi).

3. WEAK LAW OF LARGE NUMBERS

We give the first result of this section. Recall notations (6) and (9).

Theorem 3.1. *Let $(X_i, i \in \mathbb{T}^*)$ be a super-critical spatially homogeneous P^* -BMC on a GW tree. Let F satisfy (i)-(vi) and $f \in F$. The sequence $(\tilde{M}_{\mathbb{G}_q^*}(f), q \in \mathbb{N})$ converges to $\langle \mu, f \rangle W$ in L^2 , where W is defined by (8). We also have that the sequence $(\bar{M}_{\mathbb{G}_q^*}(f)\mathbf{1}_{\{|\mathbb{G}_q^*| > 0\}}, q \in \mathbb{N})$ converges to $\langle \mu, f \rangle \mathbf{1}_{\{W \neq 0\}}$ in probability.*

Proof. We first assume that $\langle \mu, f \rangle = 0$. We have,

$$\| \sum_{i \in \mathbb{G}_q^*} f(X_i) \|_{L^2}^2 = \mathbb{E} \left[\left(\sum_{i \in \mathbb{G}_q} f(X_i) \mathbf{1}_{\{i \in \mathbb{T}^*\}} \right)^2 \right] = \sum_{i \in \mathbb{G}_q} \mathbb{E}[f^2(X_i) \mathbf{1}_{\{i \in \mathbb{T}^*\}}] + B_q = m^q \mathbb{E}[f^2(Y_q)] + B_q,$$

with $B_q = \sum_{(i,j) \in \mathbb{G}_q^2, i \neq j} \mathbb{E}[f(X_i)f(X_j)\mathbf{1}_{\{(i,j) \in \mathbb{T}^{*2}\}}]$, where we used (12) for the last equality.

Since the sum in B_q concerns all pairs of distinct elements of \mathbb{G}_q , we have that $i \wedge j$, the most recent common ancestor of i and j , does not belong to \mathbb{G}_q . We shall compute B_q by

decomposing according to the generation of $k = i \wedge j$: $B_q = \sum_{r=0}^{q-1} \sum_{k \in \mathbb{G}_r} C_k$ with

$$C_k = \sum_{(i,j) \in \mathbb{G}_q^2, i \wedge j = k} \mathbb{E}[f(X_i)f(X_j)\mathbf{1}_{\{(i,j) \in \mathbb{T}^{*2}\}}].$$

If $|k| = q - 1$, using the Markov property of X and of the GW process at generation $q - 1$, we get

$$\begin{aligned} C_k &= \sum_{(i,j) \in \mathbb{G}_1^2, i \wedge j = \emptyset} \mathbb{E}[\mathbb{E}_{X_k}[f(X_i)f(X_j)\mathbf{1}_{\{(i,j) \in \mathbb{T}^{*2}\}}]\mathbf{1}_{\{k \in \mathbb{T}^*\}}] \\ &= 2\mathbb{E}[P(f \otimes f)(X_k)\mathbf{1}_{\{k \in \mathbb{T}^*\}}]. \end{aligned}$$

If $|k| < q - 1$, we have, with $r = |k|$,

$$\begin{aligned} C_k &= 2 \sum_{(i,j) \in \mathbb{G}_{q-r-1}^2} \mathbb{E}[\mathbb{E}_{X_{k_0}}[f(X_i)\mathbf{1}_{\{i \in \mathbb{T}^*\}}]\mathbb{E}_{X_{k_1}}[f(X_j)\mathbf{1}_{\{j \in \mathbb{T}^*\}}]\mathbf{1}_{\{k_0 \in \mathbb{T}^*, k_1 \in \mathbb{T}^*\}}] \\ &= 2\mathbb{E}[\sum_{i \in \mathbb{G}_{q-r-1}} \mathbb{E}_{X_{k_0}}[f(X_i)\mathbf{1}_{\{i \in \mathbb{T}^*\}}] \sum_{j \in \mathbb{G}_{q-r-1}} \mathbb{E}_{X_{k_1}}[f(X_j)\mathbf{1}_{\{j \in \mathbb{T}^*\}}]\mathbf{1}_{\{k_0 \in \mathbb{T}^*, k_1 \in \mathbb{T}^*\}}] \\ &= 2m^{2(q-r-1)}\mathbb{E}[\mathbb{E}_{X_{k_0}}[f(Y_{q-r-1})]\mathbb{E}_{X_{k_1}}[f(Y_{q-r-1})]\mathbf{1}_{\{k_0 \in \mathbb{T}^*, k_1 \in \mathbb{T}^*\}}] \\ &= 2m^{2(q-r-1)}\mathbb{E}[P(Q^{q-r-1}f \otimes Q^{q-r-1}f)(X_k)\mathbf{1}_{\{k \in \mathbb{T}^*\}}], \end{aligned}$$

where we used the Markov property of X and of the GW process at generation $r + 1$ for the first equality, (12) for the third equality and the Markov property at generation r for the last equality.

In particular, we get that $C_k = 2m^{2(q-r-1)}\mathbb{E}[P(Q^{q-r-1}f \otimes Q^{q-r-1}f)(X_k)\mathbf{1}_{\{k \in \mathbb{T}^*\}}]$ for all k s.t. $|k| \leq q - 1$. Using (12), we deduce that

$$\begin{aligned} B_q &= 2 \sum_{r=0}^{q-1} m^{2(q-r-1)} \sum_{k \in \mathbb{G}_r} \mathbb{E}[P(Q^{q-r-1}f \otimes Q^{q-r-1}f)(X_k)\mathbf{1}_{\{k \in \mathbb{T}^*\}}] \\ &= 2 \sum_{r=0}^{q-1} m^{2q-r-2} \langle \nu, Q^r P(Q^{q-r-1}f \otimes Q^{q-r-1}f) \rangle. \end{aligned}$$

Therefore, we get

$$\begin{aligned} \|\tilde{M}_{\mathbb{G}_q^*}(f)\|_{L^2}^2 &= m^{-2q} \left\| \sum_{i \in \mathbb{G}_q^*} f(X_i) \right\|_{L^2}^2 \\ (13) \quad &= m^{-q} \mathbb{E}[f^2(Y_q)] + 2m^{-2} \sum_{r=0}^{q-1} m^{-r} \langle \nu, Q^r P(Q^{q-r-1}f \otimes Q^{q-r-1}f) \rangle. \end{aligned}$$

As $f \in F$, properties (ii), (iv), (v) and (vi) imply that $\lim_{q \rightarrow \infty} m^{-q} \mathbb{E}[f^2(Y_q)] = 0$. Properties (iii), (iv) and (v) with $\langle \mu, f \rangle = 0$ implies that $P(Q^{q-r-1}f \otimes Q^{q-r-1}f)$ converges to 0 as q goes to infinity (with r fixed) and is bounded uniformly in $q > r$ by a function of F . Thus, properties (v) and (vi) imply that $\langle \nu, Q^r P(Q^{q-r-1}f \otimes Q^{q-r-1}f) \rangle$ converges to 0 as q goes to infinity (with r fixed) and is bounded uniformly in $q > r$ by a finite constant, say K . For

any $\varepsilon > 0$, we can choose r_0 s.t. $\sum_{r>r_0} m^{-r} K \leq \varepsilon$ and $q_0 > r_0$ s.t. for $q \geq q_0$ and $r \leq r_0$, we have $|\langle \nu, Q^r P(Q^{q-r-1} f \otimes Q^{q-r-1} f) \rangle| \leq \varepsilon/r_0$. We then get that for all $q \geq q_0$

$$\sum_{r=0}^{q-1} m^{-r} |\langle \nu, Q^r P(Q^{q-r-1} f \otimes Q^{q-r-1} f) \rangle| \leq \sum_{r=0}^{r_0} r_0^{-1} \varepsilon + \sum_{r=r_0+1}^{q-1} m^{-r} K \leq 2\varepsilon.$$

This gives that $\lim_{q \rightarrow \infty} \sum_{r=0}^{q-1} m^{-r} \langle \nu, Q^r P(Q^{q-r-1} f \otimes Q^{q-r-1} f) \rangle = 0$. Eventually, we get from (13)

that if $\langle \mu, f \rangle = 0$, then $\lim_{q \rightarrow \infty} \|\tilde{M}_{\mathbb{G}_q^*}(f)\|_{L^2} = 0$.

For any function $f \in F$, we have, with $g = f - \langle \mu, f \rangle$,

$$\tilde{M}_{\mathbb{G}_q^*}(f) = \tilde{M}_{\mathbb{G}_q^*}(g) + \langle \mu, f \rangle m^{-q} |\mathbb{G}_q^*|.$$

As $g \in F$ and $\langle \mu, g \rangle = 0$, the previous computations yield that $\lim_{q \rightarrow \infty} \|\tilde{M}_{\mathbb{G}_q^*}(g)\|_{L^2} = 0$. As $(m^{-q} |\mathbb{G}_q^*|, q \geq 1)$ converges in L^2 (and a.s.) to W , we get that $\tilde{M}_{\mathbb{G}_q^*}(f)$ converges to $\langle \mu, f \rangle W$ in L^2 .

Then use that $m^{-q} |\mathbb{G}_q^*|$ converges a.s. to W to get the second part of the Theorem. \square

We now prove a similar result for the average over the r -th first generations. We set $t^r = \mathbb{E}[|\mathbb{T}_r^*|]$, see (7). We first state an elementary Lemma whose proof is left to the reader.

Lemma 3.2. *Let $(v_r, r \in \mathbb{N})$ be a sequence of non negative real numbers converging to $a \in \mathbb{R}_+$, and m a real such that $m > 1$. Let*

$$w_r = \sum_{q=0}^r m^{q-r-1} v_q.$$

Then the sequence $(w_r, r \in \mathbb{N})$ converges to $a/(m-1)$.

Recall notations (6) and (9).

Theorem 3.3. *Let $(X_i, i \in \mathbb{T}^*)$ be a super-critical spatially homogeneous P^* -BMC on a GW tree. Let F satisfy (i)-(vi) and $f \in F$. The sequence $(\tilde{M}_{\mathbb{T}_r^*}(f), r \in \mathbb{N})$ converges to $\langle \mu, f \rangle W$ in L^2 , where W is defined by (8). We also have that the sequence $(\tilde{M}_{\mathbb{T}_r^*}(f) \mathbf{1}_{\{|\mathbb{T}_r^*| > 0\}}, r \in \mathbb{N})$ converges to $\langle \mu, f \rangle \mathbf{1}_{\{W \neq 0\}}$ in probability.*

Proof. We have

$$\begin{aligned} \left\| \frac{1}{t_r} \sum_{i \in \mathbb{T}_r^*} f(X_i) - \langle \mu, f \rangle W \right\|_{L^2} &= \left\| \sum_{q=0}^r \frac{m^q}{t_r} \left(\tilde{M}_{\mathbb{G}_q^*}(f) - \langle \mu, f \rangle W \right) \right\|_{L^2} \\ &\leq \sum_{q=0}^r \frac{m^q}{t_r} \left\| \tilde{M}_{\mathbb{G}_q^*}(f) - \langle \mu, f \rangle W \right\|_{L^2} \\ &= \frac{m-1}{1-m^{-r-1}} \sum_{q=0}^r m^{q-r-1} \left\| \tilde{M}_{\mathbb{G}_q^*}(f) - \langle \mu, f \rangle W \right\|_{L^2}. \end{aligned}$$

The first part of the Theorem follows from Theorem 3.1 and Lemma 3.2.

Use that $m^{-q} |\mathbb{G}_q^*|$ converges a.s. to W to deduce that $t_r^{-1} |\mathbb{T}_r^*|$ converges a.s. to W , and thus get the second part of the Theorem. \square

We end this section, with an extension of the results to functions defined on the mother-daughters quantities of interest $\Delta_i = (X_i, X_{i0}, X_{i1}) \in \mathcal{S}^3$. Recall notations (6) and (9).

Theorem 3.4. *Let $(X_i, i \in \mathbb{T}^*)$ be a super-critical spatially homogeneous P^* -BMC on a GW tree. Let F satisfy (i)-(vi) and $f \in \mathcal{B}(\mathcal{S}^3)$. We assume that P^*f and $P^*(f^2)$ exist and belong to F . Then the sequences $(\tilde{M}_{\mathbb{G}_q^*}(f), q \in \mathbb{N})$ and $(\tilde{M}_{\mathbb{T}_r^*}(f), r \in \mathbb{N})$ converge to $\langle \mu, P^*f \rangle W$ in L^2 , where W is defined by (8); and the sequences $(\bar{M}_{\mathbb{G}_q^*}(f)\mathbf{1}_{\{|\mathbb{G}_q^*|>0\}}, q \in \mathbb{N})$ and $(\bar{M}_{\mathbb{T}_r^*}(f)\mathbf{1}_{\{|\mathbb{T}_r^*|>0\}}, r \in \mathbb{N})$ converge to $\langle \mu, P^*f \rangle \mathbf{1}_{\{W \neq 0\}}$ in probability.*

Proof. Recall that $M_{\mathbb{G}_q^*}(f) = \sum_{i \in \mathbb{G}_q^*} f(\Delta_i)$. The Markov property for BMC gives

$$\|M_{\mathbb{G}_q^*}(f)\|_{L^2}^2 = \|M_{\mathbb{G}_q^*}(P^*f)\|_{L^2}^2 + \mathbb{E}[M_{\mathbb{G}_q^*}(P^*(f^2) - (P^*f)^2)].$$

Since $(m^{-q}M_{\mathbb{G}_q^*}(P^*(f^2) - (P^*f)^2), q \in \mathbb{N})$ converges to $\langle \mu, P^*(f^2) - (P^*f)^2 \rangle$ in L^2 and thus in L^1 , we have that $m^{-2q}\mathbb{E}[M_{\mathbb{G}_q^*}(P^*(f^2) - (P^*f)^2)]$ converges to 0 as q goes to infinity. Then, we deduce the convergence of $(\tilde{M}_{\mathbb{G}_q^*}(f), q \in \mathbb{N})$ and $(\bar{M}_{\mathbb{G}_q^*}(f)\mathbf{1}_{\{|\mathbb{G}_q^*|>0\}}, q \in \mathbb{N})$ from Theorem 3.1.

The proof for the convergence of $(\tilde{M}_{\mathbb{T}_r^*}(f), r \in \mathbb{N})$ and $(\bar{M}_{\mathbb{T}_r^*}(f)\mathbf{1}_{\{|\mathbb{T}_r^*|>0\}}, r \in \mathbb{N})$ mimics then the proof of Theorem 3.3. \square

4. TECHNICAL RESULTS ABOUT THE WEAK LAW OF LARGE NUMBERS

The technical Propositions of this section deal with the average of a function f when going through \mathbb{T}^* via timescales $(\tau_n(t), t \in [0, 1])$ preserving the genealogical order. Roughly speaking, these timescales allow to visit the sub-tree \mathbb{T}^* . In order to define $(\tau_n(t), t \in [0, 1])$ we need to define I_n^* , set of the n “first” cells of \mathbb{T}^* . Let $(X_i, i \in \mathbb{T}^*)$ be a super-critical spatially homogeneous P^* -BMC on a GW tree and \mathcal{G} be the σ -field generated by $(X_i, i \in \mathbb{T}^*)$.

- We consider random variables $(\Pi_q^*, q \in \mathbb{N}^*)$ which are conditionally on \mathcal{G} , independent and s.t. Π_q^* is distributed as a uniform random permutation on \mathbb{G}_q^* . In particular, given $|\mathbb{G}_q^*| = k$, $(\Pi_q^*(1), \dots, \Pi_q^*(k))$ can be viewed as a random drawing of all the elements of \mathbb{G}_q^* , without replacement.
- For each integer $n \in \mathbb{N}^*$, we define the random variable $\rho_n = \inf\{k; n \leq |\mathbb{T}_k^*|\}$, with the convention $\inf \emptyset = \infty$. Loosely speaking, ρ_n is the number of the generation to which belongs the n -th element of \mathbb{T}^* . Notice that $\rho_1 = 0$.
- Let $\tilde{\Pi}$ be the application from \mathbb{N}^* to $\mathbb{T}^* \cup \{\partial_{\mathbb{T}}\}$, where $\partial_{\mathbb{T}}$ is a cemetery point added to \mathbb{T}^* , given by $\tilde{\Pi}(1) = \emptyset$ and for $k \geq 2$:

$$\tilde{\Pi}(k) = \begin{cases} \Pi_{\rho_k}^*(k - |\mathbb{T}_{\rho_k-1}^*|) & \text{if } \rho_k < +\infty \\ \partial_{\mathbb{T}} & \text{if } \rho_k = +\infty. \end{cases}$$

Notice that $\tilde{\Pi}$ defines a random order on \mathbb{T}^* which preserves the genealogical order: if $k \leq n$ then $|\tilde{\Pi}(k)| \leq |\tilde{\Pi}(n)|$, with the convention $|\partial_{\mathbb{T}}| = \infty$. We thus define the set of the n “first” elements of \mathbb{T}^* (when $|\mathbb{T}^*| \geq n$):

$$(14) \quad I_n^* = \{\tilde{\Pi}(k), 1 \leq k \leq n \wedge |\mathbb{T}^*|\}.$$

We can now introduce the timescales: for $n \geq 1$, we consider the subdivision of $[0, 1]$ given by $\{0, s_n, \dots, s_0\}$, with $s_k = m^{-k}$. We define the continuous random time change $(\tau_n(t), t \in [0, 1])$ by

$$(15) \quad \tau_n(t) = \begin{cases} m^n t, & t \in [0, m^{-n}], \\ |\mathbb{T}_{n-k}^*| + (m^k t - 1)(m-1)^{-1} |\mathbb{G}_{n-k+1}^*|, & t \in [m^{-k}, m^{-k+1}], 1 \leq k \leq n. \end{cases}$$

Notice that $\tau_n(t) \leq |\mathbb{T}^*|$. The set $I_{[\tau_n(t)]}^*$, with $t \in [0, 1]$, corresponds to the elements of \mathbb{T}_{n-k}^* , with $k = \lfloor -\frac{\log(t)}{\log(m)} \rfloor + 1$, and the “first” fraction $(m^k t - 1)/(m - 1)$ of the elements of generation \mathbb{G}_{n-k+1}^* .

Recall (9). For the sake of simplicity, for any real $x \geq 0$, we will write $M_x^*(f)$ instead of $M_{I_{[x]}^*}(f)$, with the convention that $M_0^*(f) = 0$.

Proposition 4.1. *Let F satisfy (i)-(vi), $f \in F$ and $t \in [0, 1]$. The sequence $(m^{-n} M_{\tau_n(t)}^*(f), n \in \mathbb{N}^*)$ converges to $\langle \mu, f \rangle m(m-1)^{-1} W t$ in L^2 .*

Proof. We first consider the case $\langle \mu, f \rangle = 0$. If $t = 0$, then $\tau_n(t) = 0$ and $M_0^*(f) = 0$ by convention. Let $t \in (0, 1]$. There exists a unique $k \geq 1$ such that $m^{-k} < t \leq m^{-k+1}$. For $n \geq k$, we have, using (15) and that $\tilde{\Pi}$ preserves the order on \mathbb{T}^* ,

$$M_{\tau_n(t)}^*(f) = \sum_{i \in I_{[\tau_n(t)]}^*} f(X_i) = \sum_{i=1}^{\lfloor \tau_n(t) \rfloor} f(X_{\tilde{\Pi}(i)}) = M_{\mathbb{T}_{n-k}^*}(f) + M_{J_n}(f),$$

where $J_n = \{\tilde{\Pi}(i), |\mathbb{T}_{n-k}^*| < i \leq \lfloor \tau_n(t) \rfloor\}$. Notice that $J_n = \emptyset$ if $|\mathbb{T}_{n-k+1}^*| = 0$ and that, by convention, we then have $M_{J_n}(f) = 0$. Theorem 3.3 implies that $m^{-n} M_{\mathbb{T}_{n-k}^*}(f)$ converges to 0 in L^2 as n goes to ∞ . Recall \mathcal{G} is the σ -field generated by $(X_i, i \in \mathbb{T})$. Since $J_n \subset \mathbb{G}_{n-k+1}^*$, we have

$$\mathbb{E}[M_{J_n}(f)^2 | \mathcal{G}] = \sum_{i,j \in \mathbb{G}_{n-k+1}^*} f(X_i) f(X_j) \mathbb{E}[\mathbf{1}_{\{i,j \in J_n\}} | \mathbb{T}^*].$$

Thanks to the definition of $\tilde{\Pi}$, we have for $i, j \in \mathbb{G}_{n-k+1}$

$$\mathbf{1}_{\{i,j \in \mathbb{G}_{n-k+1}^*\}} \mathbb{E}[\mathbf{1}_{\{i,j \in J_n\}} | \mathbb{T}^*] = \mathbf{1}_{\{i,j \in \mathbb{G}_{n-k+1}^*\}} (\mathbf{1}_{\{i \neq j\}} \chi_2 + \mathbf{1}_{\{i=j\}} \chi_1),$$

where, with $a = \lfloor (m^k t - 1)(m - 1)^{-1} |\mathbb{G}_{n-k+1}^*| \rfloor$,

$$\chi_1 = \frac{a}{|\mathbb{G}_{n-k+1}^*|} \quad \text{and} \quad \chi_2 = \frac{a(a-1)}{|\mathbb{G}_{n-k+1}^*| (|\mathbb{G}_{n-k+1}^*| - 1)}.$$

Thus, we get

$$\begin{aligned} \mathbb{E}[M_{J_n}(f)^2 | \mathcal{G}] &= \chi_2 \sum_{i,j \in \mathbb{G}_{n-k+1}^*} f(X_i) f(X_j) + (\chi_1 - \chi_2) \sum_{i \in \mathbb{G}_{n-k+1}^*} f^2(X_i) \\ &= \chi_2 M_{\mathbb{G}_{n-k+1}^*}(f)^2 + (\chi_1 - \chi_2) M_{\mathbb{G}_{n-k+1}^*}(f^2) \\ &\leq M_{\mathbb{G}_{n-k+1}^*}(f)^2 + M_{\mathbb{G}_{n-k+1}^*}(f^2), \end{aligned}$$

as $0 \leq \chi_2 \leq \chi_1 \leq 1$. We have

$$(16) \quad \|m^{-n} M_{J_n}(f)^2\|_{L^2}^2 \leq \|m^{-n} M_{\mathbb{G}_{n-k+1}^*}(f)\|_{L^2}^2 + m^{-n} \|m^{-n} M_{\mathbb{G}_{n-k+1}^*}(f^2)\|_{L^1}.$$

The first term of the right hand-side of (16) converges to $\langle \mu, f \rangle W = 0$ as n goes to infinity, thanks to Theorem 3.1. The same Theorem entails that $\|m^{-n} M_{\mathbb{G}_{n-k+1}^*}(f^2)\|_{L^1}$ converges to $\mathbb{E}[\langle \mu, f^2 \rangle W]$, and consequently the second term of the right hand-side of (16) also converges to 0 as n goes to infinity. We deduce that the sequence $(m^{-n} M_{J_n}(f)^2, m \in \mathbb{N}^*)$ converges to 0 in L^2 .

Since $m^{-n} M_{\tau_n(t)}^*(f) = m^{-n} M_{\mathbb{T}_{n-k}^*}(f) + m^{-n} M_{J_n}(f)$, the sequence $(m^{-n} M_{\tau_n(t)}^*(f), n \in \mathbb{N}^*)$ converges to 0 in L^2 .

Next, we consider the case $\langle \mu, f \rangle \neq 0$. We set $g = f - \langle \mu, f \rangle$. Since $m^{-n}M_{\tau_n(t)}^*(f) = m^{-n}M_{\tau_n(t)}^*(g) + \langle \mu, f \rangle m^{-n}[\tau_n(t)]$, the Proposition will be proved as soon as we check that $(m^{-n}[\tau_n(t)], n \in \mathbb{N}^*)$ converges to $m(m-1)^{-1}tW$ in L^2 .

The case $t = 0$ is obvious. For $t \in (0, 1]$, there exists a unique $k \geq 1$ such that $m^{-k} < t \leq m^{-k+1}$. We deduce from (15) that, for $1 \leq k \leq n$,

$$m^{-n}\tau_n(t) = (m-1)^{-1} \left(\frac{|\mathbb{T}_{n-k}^*|}{t_{n-k}} (m^{-k+1} - \frac{1}{m^n}) + \frac{|\mathbb{G}_{n-k+1}^*|}{m^{n-k+1}} (mt - m^{-k+1}) \right).$$

Since both $m^{-n}|\mathbb{G}_n^*|$ and $t_n^{-1}|\mathbb{T}_n^*|$ converges to W in L^2 , we finally obtain that $m^{-n}\tau_n(t)$ converges to $m(m-1)^{-1}tW$ in L^2 . \square

We deduce from (15) and (7), that for $t \in (0, 1]$, $n \geq k$, where $k = \lfloor -\frac{\log(t)}{\log(m)} \rfloor + 1$, we have

$$\mathbb{E}[\tau_n(t)] = t_{n-k} + (m^k t - 1)(m-1)^{-1}m^{n-k+1} = (m^{n+1}t - 1)(m-1)^{-1}.$$

Thus, Proposition 4.1 implies that $(\mathbb{E}[\tau_n(t)]^{-1}M_{\tau_n(t)}^*(f), n \in \mathbb{N}^*)$ converges to $\langle \mu, f \rangle W$ in L^2 for all $t \in [0, 1]$, which generalizes Theorem 3.3.

In fact the convergence in Proposition 4.1 is uniform in t .

Corollary 4.2. *Let F satisfy (i)-(vi), $f \in F$ s.t. $|f| \in F$. We set $R_n(t) = m^{-n}M_{\tau_n(t)}^*(f) - \langle \mu, f \rangle m(m-1)^{-1}tW$. The sequence $(\sup_{t \in [0,1]} |R_n(t)|, n \in \mathbb{N}^*)$ converges to 0 in L^2 .*

Proof. Let $f \in F$ s.t. $|f| \in F$. We set $f^+ = \max(0, f)$ and $f^- = \max(0, -f)$. As F is a vector space, we get that $f^+ = (f + |f|)/2$ and $f^- = f^+ - f$ belong to F . Notice that $|R_n(t)| \leq |R_n^+(t)| + |R_n^-(t)|$, where $R_n^\delta(t) = m^{-n}M_{\tau_n(t)}^*(f^\delta) - \langle \mu, f^\delta \rangle m(m-1)^{-1}tW$ for $\delta \in \{+, -\}$. So it is enough to prove the Corollary for f non-negative. As $t \mapsto m^{-n}M_{\tau_n(t)}^*(f)$ and $t \mapsto \langle \mu, f \rangle m(m-1)^{-1}tW$ are non-decreasing and $R_n(0) = 0$, we get that for $N \geq 1$,

$$\sup_{t \in [0,1]} |R_n(t)| \leq \frac{1}{N} \langle \mu, f \rangle m(m-1)^{-1}W + \sum_{k=1}^N |R_n(k/N)|.$$

Now, use that $W \in L^2$ and that $R_n(t)$ goes to 0 in L^2 for all $t \in [0, 1]$ to get the result. \square

We have a version of Proposition 4.1 and Corollary 4.2 for functions defined on \bar{S}^3 .

Proposition 4.3. *Let F satisfy (i)-(vi), $g \in \mathcal{B}(\bar{S}^3)$ s.t. P^*g and $P^*(g^2)$ exist and belong to F . Let $t \in [0, 1]$. The sequence $(m^{-n}M_{\tau_n(t)}^*(g), n \in \mathbb{N}^*)$ converges to $\langle \mu, P^*g \rangle m(m-1)^{-1}tW$ in L^2 .*

Furthermore, if $P^|g|$ and $P^*(g|g|)$ also belong to F then $(\sup_{t \in [0,1]} |R_n(t)|, n \in \mathbb{N}^*)$ converges to 0 in L^2 , where $R_n(t) = m^{-n}M_{\tau_n(t)}^*(g) - \langle \mu, P^*g \rangle m(m-1)^{-1}tW$, for $t \in [0, 1]$.*

Proof. The proof of the first part is similar to the proof of Theorem 3.4. The proof of the second part is similar to the proof of Corollary 4.2. \square

5. FLUCTUATIONS

Recall (9). For any real $x \geq 0$, using notations from the previous Section, we will write $M_x^*(f)$ for $M_{I_{[x]}^*}(f)$, with the convention that $M_0^*(f) = 0$, where I_n^* is defined by (14). We shall prove a central limit theorem for the sequence $(M_n^*(f), n \geq 1)$, based on martingale theorems.

We set $\mathcal{H}_n = \sigma(\Delta_{\tilde{\Pi}(k)}, 1 \leq k \leq n \wedge |\mathbb{T}^*|) \vee \sigma(\tilde{\Pi}(k), 1 \leq k \leq n+1)$ for $n \geq 1$, $\mathcal{H}_0 = \sigma(X_\emptyset)$ and $\mathcal{H} = (\mathcal{H}_n, n \in \mathbb{N})$ for the corresponding filtration. With the convention that $X_{\partial_{\mathbb{T}}} = \partial$, we notice that $X_{\tilde{\Pi}(n+1)}$ is \mathcal{H}_n -measurable. Indeed, given $(\tilde{\Pi}(k), 1 \leq k \leq n+1)$, if $\tilde{\Pi}(n+1) \neq \partial_{\mathbb{T}}$, we have $\tilde{\Pi}(n+1) = \tilde{\Pi}(j)i$ for some $j \in \{1, \dots, n\}$ and $i \in \{0, 1\}$, and as $\Delta_{\tilde{\Pi}(j)} = (X_{\tilde{\Pi}(j)}, X_{\tilde{\Pi}(j)0}, X_{\tilde{\Pi}(j)1}) \in \mathcal{H}_n$, we deduce that $X_{\tilde{\Pi}(n+1)}$ is \mathcal{H}_n -measurable. In particular, as $\{|\mathbb{T}^*| \geq n+1\} \in \mathcal{H}_n$, this implies that $\mathbf{1}_{\{|\mathbb{T}^*| \geq n+1\}} \mathbb{E}[f(\Delta_{\tilde{\Pi}(n+1)}) | \mathcal{H}_n] = \mathbf{1}_{\{|\mathbb{T}^*| \geq n+1\}} P^* f(X_{\tilde{\Pi}(n+1)})$, for any $f \in \mathcal{B}(\bar{S}^3)$ such that $P^* f$ is well defined. If in addition $P^* f = 0$, then $(M_n^*(f), n \in \mathbb{N})$ is an \mathcal{H} -martingale.

We shall first recall a slightly weaker version of Theorem 4.3 from [15] on martingale convergence. (Theorem 4.3 from [15] is stated for filtrations which may vary with n .)

For $u \in \mathbb{R}^d$, we denote by u' its transpose. Let $\mathcal{H} = (\mathcal{H}_i, i \in \mathbb{N})$ be a filtration. If $(D_i, i \in \mathbb{N})$ is a sequence of vector valued random variables \mathcal{H} -adapted and such that $\mathbb{E}[D_{i+1} | \mathcal{H}_i] = 0$ for all $i \in \mathbb{N}$, then $(D_i, i \in \mathbb{N})$ is called an \mathcal{H} -martingale difference.

Theorem 5.1 (Theorem 4.3 from [15]). *Let $\mathcal{H} = (\mathcal{H}_i, i \in \mathbb{N})$ be a filtration. For all $n \in \mathbb{N}^*$, let $(D_{n,i} = (D_{n,i}^{(1)}, \dots, D_{n,i}^{(d)}), i \in \mathbb{N})$ be a sequence of \mathbb{R}^d -valued random vectors and an \mathcal{H} -martingale difference. For each $n \in \mathbb{N}$, let $(\tau_n(t), t \in [0, 1])$ be a non-decreasing càdlàg function s.t. $\tau_n(t)$ is a \mathcal{H} -stopping time for all $t \in [0, 1]$. Let $(\mathcal{T}(t), t \in [0, 1])$ be a $\mathbb{R}^{d \times d}$ -valued continuous, possibly random, function. We assume the following two conditions hold:*

- (1) *Convergence of the timescales. For all $t \in [0, 1]$, we have the following convergence in probability:*

$$\sum_{i=1}^{\tau_n(t)} \mathbb{E} [D_{n,i}(D_{n,i})' | \mathcal{H}_{i-1}] \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \mathcal{T}(t).$$

- (2) *Lindeberg condition. For all $\varepsilon > 0$, $1 \leq \ell \leq d$, we have the following convergence in probability:*

$$\sum_{i=1}^{\tau_n(1)} \mathbb{E} \left[(D_{n,i}^{(\ell)})^2 \mathbf{1}_{\{|D_{n,i}^{(\ell)}| > \varepsilon\}} | \mathcal{H}_{i-1} \right] \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0.$$

Then $(\sum_{i=1}^{\lfloor \tau_n(\cdot) \rfloor} D_{n,i}, n \in \mathbb{N}^*)$ converges in distribution to $B_{\mathcal{T}}$ in the Skorohod space $\mathbb{D}([0, 1])^d$ of \mathbb{R}^d -valued càdlàg functions defined on $[0, 1]$, where, conditionally on \mathcal{T} , $(B_{\mathcal{T}}(t), t \geq 0)$ is a Gaussian process with independent increments and $B_{\mathcal{T}}(t)$ has zero mean and variance $\mathcal{T}(t)$.

Furthermore the convergence is stable: if $(Y_n, n \in \mathbb{N})$ converges in probability to Y , then $((\sum_{i=1}^{\lfloor \tau_n(\cdot) \rfloor} D_{n,i}, Y_n), n \in \mathbb{N})$ converges in distribution to $(B_{\mathcal{T}}, Y)$, where $B_{\mathcal{T}}$ is conditionally on (\mathcal{T}, Y) distributed as $B_{\mathcal{T}}$ conditionally on \mathcal{T} , and the distribution of (\mathcal{T}, Y) is determined by the following convergence

$$\left(\sum_{i=1}^{\tau_n(\cdot)} \mathbb{E} [D_{n,i}(D_{n,i})' | \mathcal{H}_{i-1}], Y_n \right) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} (\mathcal{T}, Y).$$

For the sake of simplicity, we will write $P^* h^k$ for $P^*(h^k)$, and if $h = (h_1, \dots, h_d)$ is an \mathbb{R}^d valued function, we will write $P^* h$ for $(P^* h_1, \dots, P^* h_d)$ and $\langle \mu, h \rangle$ for $(\langle \mu, h_1 \rangle, \dots, \langle \mu, h_d \rangle)$.

Theorem 5.2. *Let $(X_i, i \in \mathbb{T}^*)$ be a super-critical spatially homogeneous P^* -BMC on a GW tree and τ_n be defined by (15). Let F satisfy (i)-(vi). Let $d \geq 1$, $d' \geq 1$, $f = (f_1, \dots, f_d) \in \mathcal{B}(\bar{S}^3)^d$, $g = (g_1, \dots, g_{d'}) \in \mathcal{B}(\bar{S}^3)^{d'}$ such that $P^* f_\ell^k$, exist and belong to F for all $1 \leq \ell \leq d$ and $1 \leq k \leq 4$, $P^* g_\ell$, $P^* |g_\ell|$, $P^* g_\ell^2$ and $P^* g_\ell |g_\ell|$ exist and belong to F for all $1 \leq \ell \leq d'$. Let*

Σ be a square root of the symmetric positive matrix $m(m-1)^{-1}\langle\mu, P^*(ff') - (P^*f)(P^*f)'\rangle$ and $\gamma = m(m-1)^{-1}\langle\mu, P^*g\rangle$.

Then, the sequence $(m^{-n/2}M_{\tau_n(\cdot)}^*(f - P^*f), m^{-n}M_{\tau_n(\cdot)}^*(g))$ converges in distribution in the Skorohod space $\mathbb{D}([0, 1], \mathbb{R}^{d+d'})$ of $\mathbb{R}^{d+d'}$ -valued càdlàg functions defined on $[0, 1]$, to the process $(\Sigma\sqrt{W}B, \gamma Wh_0)$, where B is a d -dimensional Brownian motion independent of W , defined by (8), and h_0 is the identity function $t \mapsto t$.

Proof. Notice that τ_n defined by (15) is a non-decreasing continuous function s.t. $\tau_n(t)$ is a \mathcal{H} -stopping time for all $t \in [0, 1]$. We set for all $n, i \in \mathbb{N}^*$,

$$D_{n,i} = m^{-n/2} \left(f(\Delta_{\tilde{\Pi}(i)}) - P^*f(X_{\tilde{\Pi}(i)}) \right) \mathbf{1}_{\{i \leq |\mathbb{T}^*|\}},$$

so that $(D_{n,i}, i \in \mathbb{N})$ is an \mathcal{H} -martingale difference. Notice the matrix $\langle\mu, P^*ff' - (P^*f)(P^*f)'\rangle$ is indeed symmetric and positive, so that Σ is well defined.

Notice that

$$\mathbb{E} [D_{n,i}(D_{n,i})' | \mathcal{H}_{i-1}] = m^{-n} \left(P^*(ff')(X_{\tilde{\Pi}(i)}) - (P^*f)(X_{\tilde{\Pi}(i)})(P^*f)'(X_{\tilde{\Pi}(i)}) \right) \mathbf{1}_{\{i \leq |\mathbb{T}^*|\}}.$$

The convergence of the timescales condition of Theorem 5.1 with $\mathcal{T}(t) = \Sigma^2 Wt$, is then a direct application of Proposition 4.1

For $1 \leq \ell \leq d$, we have

$$\mathbb{E} \left[(D_{n,i}^{(\ell)})^2 \mathbf{1}_{\{|D_{n,i}^{(\ell)}| > \varepsilon\}} | \mathcal{H}_{i-1} \right] \leq \varepsilon^{-2} \mathbb{E} \left[(D_{n,i}^{(\ell)})^4 | \mathcal{H}_{i-1} \right] = \varepsilon^{-2} m^{-2n} P^*(f_\ell - P^*f_\ell)^4(X_{\tilde{\Pi}(i)}) \mathbf{1}_{\{i \leq |\mathbb{T}^*|\}}.$$

The Lindeberg condition of Theorem 5.1 is then a direct application of Proposition 4.1

Notice the second part of Proposition 4.3. implies the convergence of $Y_n = m^{-n}M_{\tau_n(\cdot)}^*(g)$ to γWh_0 in probability in the Skorohod space. We then deduce the result from Theorem 5.1. \square

The following result is an immediate consequence of Theorem 5.2.

Corollary 5.3. *Let $(X_i, i \in \mathbb{T}^*)$ be a super-critical spatially homogeneous P^* -BMC on a GW tree. Let F satisfy (i)-(vi). Let $f \in \mathcal{B}(\bar{S}^3)$ such that P^*f^k exists and belongs to F for all $1 \leq k \leq 4$. Let $\sigma^2 = \langle\mu, P^*f^2 - (P^*f)^2\rangle$.*

Then we have the following convergence in distribution:

$$\mathbf{1}_{\{|\mathbb{G}_n^*| > 0\}} |\mathbb{T}_n^*|^{-1/2} \sum_{i \in \mathbb{T}_n^*} f(\Delta_i) - P^*f(X_i) \xrightarrow[n \rightarrow \infty]{(d)} \mathbf{1}_{\{W \neq 0\}} \sigma G,$$

where G is a Gaussian random variable with mean zero and variance 1 independent of W , which is defined by (8).

Proof. Notice that $\sum_{i \in \mathbb{T}_n^*} f(\Delta_i) - P^*f(X_i) = M_{\tau_n(1)}^*(f) - M_{\tau_n(1)}^*(P^*f)$, $|\mathbb{T}_n^*| = M_{\tau_n(1)}^*(\mathbf{1})$ and that $\mathbf{1}_{\{|\mathbb{G}_n^*| > 0\}}$ converges a.s. to $\mathbf{1}_{\{W \neq 0\}}$. Then, to conclude, use the stable convergence of Theorem 5.2, and the fact that the marginals at time 1 converge since the limit is continuous. \square

The next result gives that the fluctuations over each generation are asymptotically independent.

Corollary 5.4. *Let $(X_i, i \in \mathbb{T}^*)$ be a super-critical spatially homogeneous P^* -BMC on a GW tree. Let F satisfy (i)-(vi). Let $d \geq 1$, $f_1, \dots, f_d \in \mathcal{B}(\bar{S}^3)$ such that $P^*f_\ell^k$ exist and belong to F for all $1 \leq \ell \leq d$ and $1 \leq k \leq 4$. Let $\sigma_\ell^2 = \langle\mu, P^*f_\ell^2 - (P^*f_\ell)^2\rangle$ for $1 \leq \ell \leq d$.*

We set for $f \in \mathcal{B}(\bar{S}^3)$

$$N_n(f) = |\mathbb{G}_n^*|^{-1/2} (M_{\mathbb{G}_n^*}(f - P^*f)).$$

Then we have the following convergence in distribution:

$$(N_n(f_1), \dots, N_{n-d+1}(f_d)) \mathbf{1}_{\{|\mathbb{G}_n^*| > 0\}} \xrightarrow[n \rightarrow \infty]{(d)} \mathbf{1}_{\{W \neq 0\}} (\sigma_1 G_1, \dots, \sigma_d G_d),$$

where G_1, \dots, G_d are independent Gaussian random variables with mean zero and variance 1 and are independent of W , which is defined by (8).

Proof. Notice that for $n > k \geq 0$,

$$N_{n-k}(f) \mathbf{1}_{\{|\mathbb{G}_n^*| > 0\}} = \frac{M_{\tau_n(m-k)}^*(f - P^*f) - M_{\tau_n(m-k-1)}^*(f - P^*f)}{\sqrt{M_{\tau_n(m-k)}^*(\mathbf{1}) - M_{\tau_n(m-k-1)}^*(\mathbf{1})}} \mathbf{1}_{\{|\mathbb{G}_n^*| > 0\}}$$

and $\mathbf{1}_{\{|\mathbb{G}_n^*| > 0\}}$ converges a.s. to $\mathbf{1}_{\{W \neq 0\}}$. To conclude, use the stable convergence of Theorem 5.2 and that the increments of the Brownian motion are independent. \square

The extension of the two previous Corollaries to vector-valued function can be proved in a very similar way.

6. ESTIMATION AND TESTS FOR THE ASYMMETRIC AUTO-REGRESSIVE MODEL

We consider the asymmetric auto-regressive model given in Section 1.1. Notice that the process $(X_i, i \in \mathbb{T})$ defined in Section 1.1 with the convention that $X_i = \partial$ if the cell i is dead or non existing is a spatially homogeneous BMC on a GW tree. We shall assume it is super-critical, that is $2p_{1,0} + p_1 + p_0 > 1$.

We compute the maximum likelihood estimator (MLE)

$$\hat{\theta}^n = (\hat{\alpha}_0^n, \hat{\beta}_0^n, \hat{\alpha}_1^n, \hat{\beta}_1^n, \hat{\alpha}_0^m, \hat{\beta}_0^m, \hat{\alpha}_1^m, \hat{\beta}_1^m, \hat{p}_{1,0}^n, \hat{p}_0^n, \hat{p}_1^n)$$

of θ given by (4) and $\kappa^n = (\hat{\sigma}^n, \hat{\rho}^n, \hat{\sigma}_0^n, \hat{\sigma}_1^n)$ of $\kappa = (\sigma, \rho, \sigma_0, \sigma_1)$, based on the observation of a sub-tree \mathbb{T}_{n+1}^* . Let $\mathbb{T}_n^{1,0}$ be the set of cells in \mathbb{T}_n^* with two living daughters, \mathbb{T}_n^0 (resp. \mathbb{T}_n^1) be the set of cells of \mathbb{T}_n^* with only the new (resp. old) pole daughter alive:

$$\mathbb{T}_n^{1,0} = \{i \in \mathbb{T}_n^*, \Delta_i \in S^3\}, \mathbb{T}_n^0 = \{i \in \mathbb{T}_n^*, \Delta_i \in S^2 \times \{\partial\}\} \text{ and } \mathbb{T}_n^1 = \{i \in \mathbb{T}_n^*, \Delta_i \in S \times \{\partial\} \times S\}.$$

It is elementary to get that for $\delta \in \{0, 1\}$,

$$(17) \quad \hat{\alpha}_\delta^n = \frac{|\mathbb{T}_n^{1,0}|^{-1} \sum_{i \in \mathbb{T}_n^{1,0}} X_i X_{i\delta} - (|\mathbb{T}_n^{1,0}|^{-1} \sum_{i \in \mathbb{T}_n^{1,0}} X_i)(|\mathbb{T}_n^{1,0}|^{-1} \sum_{i \in \mathbb{T}_n^{1,0}} X_{i\delta})}{|\mathbb{T}_n^{1,0}|^{-1} \sum_{i \in \mathbb{T}_n^{1,0}} X_i^2 - (|\mathbb{T}_n^{1,0}|^{-1} \sum_{i \in \mathbb{T}_n^{1,0}} X_i)^2},$$

$$(18) \quad \hat{\beta}_\delta^n = |\mathbb{T}_n^{1,0}|^{-1} \sum_{i \in \mathbb{T}_n^{1,0}} X_{i\delta} - \hat{\alpha}_\delta^n |\mathbb{T}_n^{1,0}|^{-1} \sum_{i \in \mathbb{T}_n^{1,0}} X_i,$$

$$\hat{\alpha}_\delta^m = \frac{|\mathbb{T}_n^\delta|^{-1} \sum_{i \in \mathbb{T}_n^\delta} X_i X_{i\delta} - (|\mathbb{T}_n^\delta|^{-1} \sum_{i \in \mathbb{T}_n^\delta} X_i)(|\mathbb{T}_n^\delta|^{-1} \sum_{i \in \mathbb{T}_n^\delta} X_{i\delta})}{|\mathbb{T}_n^\delta|^{-1} \sum_{i \in \mathbb{T}_n^\delta} X_i^2 - (|\mathbb{T}_n^\delta|^{-1} \sum_{i \in \mathbb{T}_n^\delta} X_i)^2},$$

$$\hat{\beta}_\delta^m = |\mathbb{T}_n^\delta|^{-1} \sum_{i \in \mathbb{T}_n^\delta} X_{i\delta} - \hat{\alpha}_\delta^m |\mathbb{T}_n^\delta|^{-1} \sum_{i \in \mathbb{T}_n^\delta} X_i,$$

$$(19) \quad \hat{p}_{1,0}^n = \frac{|\mathbb{T}_n^{1,0}|}{|\mathbb{T}_n^*|}, \quad \hat{p}_\delta^n = \frac{|\mathbb{T}_n^\delta|}{|\mathbb{T}_n^*|},$$

and

$$(\hat{\sigma}^n)^2 = \frac{1}{2|\mathbb{T}_n^{1,0}|} \sum_{i \in \mathbb{T}_n^{1,0}} (\hat{\varepsilon}_{i0}^2 + \hat{\varepsilon}_{i1}^2), \quad \hat{\rho}^n = \frac{1}{(\hat{\sigma}^n)^2 |\mathbb{T}_n^{1,0}|} \sum_{i \in \mathbb{T}_n^{1,0}} \hat{\varepsilon}_{i0} \hat{\varepsilon}_{i1}, \quad \text{and} \quad (\hat{\sigma}_\delta^n)^2 = \frac{1}{|\mathbb{T}_n^\delta|} \sum_{i \in \mathbb{T}_n^\delta} \hat{\varepsilon}'_{i\delta}.$$

The residues are

$$\hat{\varepsilon}_{i\delta} = X_{i\delta} - \hat{\alpha}_\delta^n X_i - \hat{\beta}_\delta^n \quad \text{for } i \in \mathbb{T}_n^{1,0}, \quad \text{and} \quad \hat{\varepsilon}'_{i\delta} = X_{i\delta} - \hat{\alpha}_\delta^m X_i - \hat{\beta}_\delta^m \quad \text{for } i \in \mathbb{T}_n^\delta.$$

Notice that those MLE are based on polynomial functions of the observations. In order to use the results of Sections 3 and 5, we first show that the set of continuous and polynomially growing functions satisfies properties (i) to (v) of Section 2. The set of continuous and polynomially growing functions $C_{\text{pol}}(\mathbb{R})$ is defined as the set of continuous real functions defined on \mathbb{R} , s.t. there exists $m \geq 0$ and $c \geq 0$ and for all $x \in \mathbb{R}$, $|f(x)| \leq c(1 + |x|^m)$. It is easy to check that $C_{\text{pol}}(\mathbb{R})$ satisfies conditions (i)-(iii). To check properties (iv) and (v), we notice that the auxiliary Markov chain $Y = (Y_n, n \in \mathbb{N})$ can be written in the following way:

$$Y_{n+1} = a_{n+1} Y_n + b_{n+1},$$

with $b_n = b'_n + s_n e_n$, where $((a_n, b'_n, s_n), n \geq 1)$ is a sequence of independent identically distributed random variables, whose common distribution is given by, for $\delta \in \{0, 1\}$,

$$(20) \quad \mathbb{P}(a_1 = \alpha_\delta, b'_1 = \beta_\delta, s_1 = \sigma) = \frac{p_{1,0}}{m} \quad \text{and} \quad \mathbb{P}(a_1 = \alpha'_\delta, b'_1 = \beta'_\delta, s_1 = \sigma_\delta) = \frac{p_\delta}{m},$$

$(e_n, n \geq 1)$ is a sequence of independent $\mathcal{N}(0, 1)$ random variables, and is independent of $((a_n, b'_n, s_n), n \geq 1)$, and both sequences are independent of Y_0 . Notice that Y_n is distributed as $Z_n = a_1 a_2 \cdots a_{n-1} a_n Y_0 + \sum_{k=1}^n a_1 a_2 \cdots a_{k-1} b_k$. Since $|a_k| \leq \max(|\alpha_0|, |\alpha_1|, |\alpha'_0|, |\alpha'_1|) < 1$ for all $k \in \mathbb{N}^*$, we get that the sequence $(Z_n, n \in \mathbb{N})$ converges a.s. to a limit Z . This implies that Y converges in distribution to Z . Following the proof of Lemma 26 in [9], we get that $C_{\text{pol}}(\mathbb{R})$ fulfills properties (iv) and (v), with μ the distribution of Z .

Proposition 6.1. *Assume that the distribution of the ancestor X_0 has finite moments of all orders. Then $(\mathbf{1}_{\{|\mathbb{G}_n^*| > 0\}} \hat{\theta}^n, n \geq 1)$ and $(\mathbf{1}_{\{|\mathbb{G}_n^*| > 0\}} \hat{\kappa}^n, n \geq 1)$ converges in probability respectively to $\mathbf{1}_{\{W \neq 0\}} \theta$ and $\mathbf{1}_{\{W \neq 0\}} \kappa$, where W is defined by (8).*

Proof. The hypothesis on the distribution of X_\emptyset implies that $C_{\text{pol}}(\mathbb{R})$ fulfills (vi). The result is then a direct consequence of Theorem 3.4. \square

Remark 6.2. Using similar arguments as in Proposition 30 and 34 of [9], it is easy to deduce from (13) and the proofs of Theorem 14 and Proposition 28 of [9] that the convergences in Proposition 6.1 hold a.s., that is the MLEs $\hat{\theta}^n$ and $\hat{\kappa}^n$ are strongly consistent.

From the definition of Z_n , we deduce that in distribution $Z \stackrel{(d)}{=} a_1 Z' + b_1$, where Z' is distributed as Z and is independent of (a_1, b_1) (see (20) for the distribution of (a_1, b_1)). This equality in distribution entails that

$$(21) \quad \mu_1 = \mathbb{E}[Z] = \frac{\bar{\beta}}{1 - \bar{\alpha}} \quad \text{and} \quad \mu_2 = \mathbb{E}[Z^2] = \frac{2\bar{\alpha}\bar{\beta}/(1 - \bar{\alpha}) + \bar{\beta}^2 + \bar{\sigma}^2}{1 - \bar{\alpha}^2},$$

where $\bar{\alpha} = \mathbb{E}[a_1]$, $\bar{\alpha}^2 = \mathbb{E}[a_1^2]$, $\bar{\beta} = \mathbb{E}[b_1]$, $\bar{\beta}^2 = \mathbb{E}[b_1^2]$, $\bar{\alpha}\bar{\beta} = \mathbb{E}[a_1 b_1]$ and $\bar{\sigma}^2 = \mathbb{E}[s_1^2]$.

We can now state one of the main result of this section.

Proposition 6.3. *Assume that the distribution of the ancestor X_\emptyset has finite moments of all orders. Then $\mathbf{1}_{\{|\mathbb{T}_n^*| > 0\}} |\mathbb{T}_n^*|^{1/2} (\hat{\theta}^n - \theta)$ converges in law to $\mathbf{1}_{\{W \neq 0\}} G_{11}$, where G_{11} is a 11-dimensional vector, independent of W defined by (8), with law $\mathcal{N}(0, \Sigma)$ where*

$$\Sigma = \begin{pmatrix} \sigma^2 K/p_{1,0} & \rho\sigma^2 K/p_{1,0} & 0 & 0 & 0 \\ \rho\sigma^2 K/p_{1,0} & \sigma^2 K/p_{1,0} & 0 & 0 & 0 \\ 0 & 0 & \sigma_0^2 K/p_0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_1^2 K/p_1 & 0 \\ 0 & 0 & 0 & 0 & \Gamma \end{pmatrix} \quad \text{with}$$

$$K = (\mu_2 - \mu_1^2)^{-1} \begin{pmatrix} 1 & -\mu_1 \\ -\mu_1 & \mu_2 \end{pmatrix} \quad \text{and} \quad \Gamma = \begin{pmatrix} p_{1,0}(1 - p_{1,0}) & -p_0 p_{1,0} & -p_1 p_{1,0} \\ -p_0 p_{1,0} & p_0(1 - p_0) & -p_0 p_1 \\ -p_1 p_{1,0} & -p_0 p_1 & p_1(1 - p_1) \end{pmatrix}.$$

The proof of Proposition 6.3 relies on Theorem 5.2 and mimics the proof of Proposition 33 of [9]. It is left to the reader.

Remark 6.4. Proposition 6.3 deals with the asymptotic normality of the MLE of θ based on the observation of the sub-tree \mathbb{T}_{n+1}^* . If $L(X_i, i \in \mathbb{T}_{n+1}^*, \theta)$ denotes the corresponding log-likelihood function for θ , the Fisher information, say I_{n+1} , is given by

$$I_{n+1} = -\mathbb{E} \left[\frac{\partial^2 L(X_i, i \in \mathbb{T}_{n+1}^*, \theta)}{\partial \theta \partial \theta'} \right].$$

Using Theorem 3.4, one can check that $\lim_{n \rightarrow \infty} I_{n+1} / \mathbb{E}[|\mathbb{T}_{n+1}^*|] = \Sigma^{-1}$. This is the analogue of the well-known asymptotic efficiency of the MLE for parametric sample of i.i.d. random variables.

Let $\theta_{1,0}$ (resp. $\hat{\theta}_{1,0}^n$) stand for $(\alpha_0, \beta_0, \alpha_1, \beta_1)$ (resp. $(\hat{\alpha}_0^n, \hat{\beta}_0^n, \hat{\alpha}_1^n, \hat{\beta}_1^n)$).

Remark 6.5. Proposition 6.3 is quite similar to Proposition 33 in [9]. One of the main differences comes from the factor $p_{1,0}^{-1}$ in front of the matrix K in the asymptotic covariance matrix for the estimation of $\theta_{1,0}$ with $\hat{\theta}_{1,0}^n$. As a matter of fact, this factor comes from the normalization by $|\mathbb{T}_n^*|^{1/2}$, number of living cells up to generation n , whereas this estimation is related to the cells with two living daughters, which would induce a normalization by $|\mathbb{T}_n^{1,0}|^{1/2}$. Since $\mathbf{1}_{\{|\mathbb{T}_n^*| > 0\}} |\mathbb{T}_n^{1,0}| / |\mathbb{T}_n^*|$ converges in probability to $p_{1,0} \mathbf{1}_{\{W \neq 0\}}$, such a normalization would suppress the factor $p_{1,0}^{-1}$, see the following Corollary.

Corollary 6.6. *Assume that the distribution of the ancestor X_\emptyset has finite moments of all orders. Then $\mathbf{1}_{\{|\mathbb{G}_n^*|>0\}}|\mathbb{T}_n^{1,0}|^{1/2}(\hat{\theta}_{1,0}^n - \theta_{1,0})$ converges in law to $\mathbf{1}_{\{W \neq 0\}}G_4$, where G_4 is a 4-dimensional vector, independent of W defined by (8), with law $\mathcal{N}(0, \Sigma')$ where*

$$\Sigma' = \sigma^2 \begin{pmatrix} K & \rho K \\ \rho K & K \end{pmatrix} \quad \text{with} \quad K = (\mu_2 - \mu_1^2)^{-1} \begin{pmatrix} 1 & -\mu_1 \\ -\mu_1 & \mu_2 \end{pmatrix}.$$

This result is formally the same as Proposition 33 of [9], but one should notice that μ_1 and μ_2 are not defined the same way as in [9], since here they also depends on the parameters concerning cells with dead sisters see equations (2), (3), (20) and (21).

In order to detect cellular aging, see [9] in the case of no death ($m = 2$), we consider the null hypothesis $H_0 = \{(\alpha_0, \beta_0) = (\alpha_1, \beta_1)\}$, which corresponds to no aging and its alternative $H_1 = \{(\alpha_0, \beta_0) \neq (\alpha_1, \beta_1)\}$. Notice that $\theta \mapsto \mu_1(\theta)$ and $(\theta, \kappa) \mapsto \mu_2(\theta, \kappa)$ given by (21) are continuous functions defined respectively on $\Theta = ((-1, 1) \times \mathbb{R})^4 \times ([0, 1]^3 \setminus \{0, 0, 0\})$ and $\Theta \times]0, +\infty[^3$. We set $\hat{\mu}_1^n = \mu_1(\hat{\theta}^n)$ and $\hat{\mu}_2^n = \mu_2(\hat{\theta}^n, \hat{\kappa}^n)$. Proposition 6.7 allows to build a test for H_0 against H_1 . Its proof, which is left to the reader, follows the proof of Proposition 35 of [9] and uses Corollary 6.6, the value of the extinction probability $\eta = \mathbb{P}(W = 0) = 1 - \frac{m-1}{p_{1,0}}$, where W is defined by (8) and Remark 6.2.

Proposition 6.7. *Let U and V be two independent random variables, with U distributed as a χ^2 with two degrees of freedom and V a Bernoulli random variable with parameter $1 - \eta$.*

Assume that the distribution of the ancestor X_\emptyset has finite moments of all orders and define the test statistic

$$\zeta_n = \frac{|\mathbb{T}_n^{1,0}|}{2(\hat{\sigma}^n)^2(1 - \hat{\rho}^n)} \left\{ (\hat{\alpha}_0^n - \hat{\alpha}_1^n)^2 (\hat{\mu}_2^n - (\hat{\mu}_1^n)^2) + \left((\hat{\alpha}_0^n - \hat{\alpha}_1^n) \hat{\mu}_1^n + \hat{\beta}_0^n - \hat{\beta}_1^n \right)^2 \right\}.$$

Then, the statistics $\mathbf{1}_{\{|\mathbb{G}_n^|>0\}}\zeta_n$ converges under H_0 in distribution to UV , and under H_1 a.s. to 0 on $\{V = 0\}$ and $+\infty$ on $\{V = 1\}$.*

Remark 6.8. Let us assume that:

- Death occurs, that is $m \in (1, 2)$.
- There is no difference for the marginal distribution of a daughter according to her sister is dead or alive; that is $\alpha'_\delta = \alpha_\delta$ and $\beta'_\delta = \beta_\delta$ for $\delta \in \{0, 1\}$.
- For simplicity, the death probability is symmetric, that is $p_0 = p_1$.

If one uses the statistics given by Proposition 33 in [9] with all the available data, that is if one uses

- Formula (17) and (18) with $\mathbb{T}_n^{1,0}$ replaced by $\mathbb{T}_n^{1,0} \cup \mathbb{T}_n^\delta$;
- The variance estimator:

$$(\hat{\sigma}^n)^2 = \frac{1}{|\mathbb{T}_{n+1}^*| - 1} \left(\sum_{i \in \mathbb{T}_n^{1,0}} (\hat{\varepsilon}_{i0}^2 + \hat{\varepsilon}_{i1}^2) + \sum_{i \in \mathbb{T}_n^0} \hat{\varepsilon}_{i0}^2 + \sum_{i \in \mathbb{T}_n^1} \hat{\varepsilon}_{i1}^2 \right);$$

(Notice that we divide by $|\mathbb{T}_{n+1}^*| - 1$ as this is equal to the total number of data:

$$2|\mathbb{T}_n^{1,0}| + |\mathbb{T}_n^0| + |\mathbb{T}_n^1|.)$$

- Keep the same estimation of the correlation: $\hat{\rho}^n = \frac{1}{(\hat{\sigma}^n)^2 |\mathbb{T}_n^{1,0}|} \sum_{i \in \mathbb{T}_n^{1,0}} \hat{\varepsilon}_{i0} \hat{\varepsilon}_{i1}$;

then one check that, as n goes to infinity, $\mathbf{1}_{\{|\mathbb{G}_n^*|>0\}}|\mathbb{T}_n^*|^{1/2}(\hat{\theta}^n - \theta)$ converges in distribution to $\mathbf{1}_{\{W \neq 0\}}G$, where G is a centered Gaussian vector with covariance matrix

$$\sigma^2 (p_{1,0} + p_1)^{-1} \begin{pmatrix} K & \rho p_{1,0} (p_{1,0} + p_1)^{-1} K \\ \rho p_{1,0} (p_{1,0} + p_1)^{-1} K & K \end{pmatrix},$$

where K is as in Proposition 6.3; and G is independent of W , which is defined by (8). Then, it is not difficult to check that the statistics proposed by Guyon in Proposition 35 of [9], converges under H_0 towards cUV , with U and V as in Proposition 6.7 and

$$c = \frac{(p_{1,0} + p_1)^{-1}(1 - \rho p_{1,0}(p_{1,0} + p_1)^{-1})}{(1 - \rho)}.$$

As $\rho \in [-1, 1]$, $p_{1,0} + p_1 > 1/2$ (because $m > 1$ and $p_0 = p_1$) and $2p_1 + p_{1,0} \leq 1$, one can check that $c > 1$. In particular, using the test statistic designed for cells with no death to data of cells with death leads to a non-conservative test.

REFERENCES

- [1] M. ACKERMANN, L. CHAO, C. T. BERGSTROM, and M. BOEBELI. On the evolutionary of aging. *Aging Cell*, 6(2):235–244, 2007.
- [2] M. ACKERMANN, S. STEARNS, and U. JENAL. Senescence in a bacterium with asymmetric division. *Science*, 300:1920, 2003.
- [3] K. B. ATHREYA and P. E. NEY. *Branching processes*. Springer-Verlag, New York, 1972.
- [4] V. BANSAYE. Cell contamination and branching process in random environment with immigration. *ArXiv:0805.1863*, 2008.
- [5] V. BANSAYE. Proliferating parasites in dividing cells: Kimmel’s branching model revisited. *Ann. Appl. Probab.*, 18(3):967–996, 2008.
- [6] B. BERCU, B. DE SAPORTA, and A. GÉGOUT-PETIT. Asymptotic analysis for bifurcating autoregressive processes via a martingale approach. *arXiv*, 0807.0528v1, 2008.
- [7] R. COWAN and R. G. STAUDTE. The bifurcating autoregression model in cell lineage. *Biometrics*, 42(4):769–783, 1986.
- [8] S. N. EVANS and D. STEINSALTZ. Damage segregation at fissioning may increase growth rates: A superprocess model. *Theor. Pop. Biol.*, 71(4):473–490, 2007.
- [9] J. GUYON. Limit theorems for bifurcating Markov chains. Application to the detection of cellular aging. *Ann. Appl. Probab.*, 17(5-6):1538–1569, 2007.
- [10] T. HARRIS. *The theory of branching process*. Springer-Verlag, Berlin, 1963.
- [11] S. Y. HWANG, I. V. BASAWA, and I. K. YEO. Local asymptotic normality for bifurcating autoregressive processes and related asymptotic inference. *Statistical Methodology*, 2008.
- [12] A. B. LINDNER, R. MADDEN, A. DEMAREZ, E. J. STEWART, and F. TADDEI. Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *PNAS*, 105(8):3076–3081, 2008.
- [13] F. MAAOUIA and A. TOUATI. Identification of multitype branching processes. *Comptes Rendus de l’Academie des Sciences Series I Mathematics*, 331:923–928, 2000.
- [14] S. P. MEYN and R. L. TWEEDIE. *Markov chains and stochastic stability*. Communications and Control Engineering Series. Springer-Verlag London Ltd., London, 1993.
- [15] H. ROOTZÉN. Central limit theory for martingales via random change of time. In *Probability and mathematical statistics*, pages 154–189. Uppsala Univ., Uppsala, 1983.
- [16] E. J. STEWART, R. MADDEN, P. G., and F. TADDEI. Aging and death in an organism that reproduces by morphologically symmetric division. *PLoS Biol.*, 3(2):e45, 2005.

JEAN-FRANÇOIS DELMAS, CERMICS, UNIV. PARIS-EST, 6-8 AV. BLAISE PASCAL, CHAMPS-SUR-MARNE, 77455 MARNE LA VALLÉE, FRANCE.

E-mail address: delmas@cermics.enpc.fr

LAURENCE MARSALLE, LAB. P. PAINLEVÉ, CNRS UMR 8524, BÂT. M2, UNIV. LILLE 1, CITÉ SCIENTIFIQUE, 59655 VILLENEUVE D’ASCQ CEDEX, FRANCE.

E-mail address: Laurence.Marsalle@univ-lille1.fr