CENTRAL LIMIT THEOREM FOR BIFURCATING MARKOV CHAINS: THE MOTHER-DAUGHTERS TRIANGLES CASE

S. VALÈRE BITSEKI PENDA

Abstract. The main objective of this article is to establish a central limit theorem for additive three-variable functionals of bifurcating Markov chains. We thus extend the central limit theorem under point-wise ergodic conditions studied in Bitseki-Delmas (2022) and to a lesser extent, the results of Bitseki-Delmas (2022) on central limit theorem under $L^2$ ergodic conditions. Our results also extend and complement those of Guyon (2007) and Delmas and Marsalle (2010). In particular, when the ergodic rate of convergence is greater than $1/\sqrt{2}$, we have, for certain class of functions, that the asymptotic variance is non-zero at a speed faster than the usual central limit theorem studied by Guyon and Delmas-Marsalle.

Keywords: Bifurcating Markov chains, binary trees.

Mathematics Subject Classification (2010): 60F05, 60J80.

1. Introduction

This article is devoted to the extension of the central limit theorem for bifurcating Markov chains given in Bitseki-Delmas [3] when the functions do not only depend on one variable $x$ say, but on the triplet, $xx_0x_1 := (x, x_0, x_1)$ say. Using the terminology given in [8], we will talk about the mother-daughters triangles case. The study of mother-daughters triangles for bifurcating Markov chains models is particularly important to make statistical inference (see for e.g [8, 7, 4, 6, 1]). The results given here allow us to extend those obtained by Guyon [8] and Delmas and Marsalle [7]. Indeed, in their works, the authors in [7, 8] considered conditionally centered additive functionals (see Remark 3.2 for more details), which is equivalent to study increments of martingale. As in [3], we prove the existence of three regimes for the central limit theorem. However, these three regimes disappear when considering conditionally centered additive functionals (see Section 3 and in particular Theorem 3.1 for more details). We stress that with appropriate hypothesis, the results obtained in this paper also hold in the $L^2$ ergodic case (see [2] for more details). Now, before giving the plan of the paper, we introduce useful definitions, notations and assumptions.

1.1. Bifurcating Markov chains. If $(E, \mathcal{E})$ is a measurable space, then $\mathcal{B}(E)$ (resp. $\mathcal{B}_b(E)$, resp. $\mathcal{B}_+(E)$) denotes the set of (resp. bounded, resp. non-negative) $\mathbb{R}$-valued measurable functions defined on $E$. For $f \in \mathcal{B}(E)$, we set $\|f\|_{\infty} = \sup\{|f(x)|, x \in E\}$. For a finite measure $\lambda$ on $(E, \mathcal{E})$ and $f \in \mathcal{B}(E)$ we shall write $\langle \lambda, f \rangle$ for $\int f(x) d\lambda(x)$ whenever this integral is well defined. If $(E, d)$ is a metric space, then $\mathcal{E}$ will denote its Borel $\sigma$-field and the set $\mathcal{C}_b(E)$ (resp. $\mathcal{C}_+(E)$) denotes the set of bounded (resp. non-negative) $\mathbb{R}$-valued continuous functions defined on $E$.

Let $(S, \mathcal{F})$ be a measurable space. Let $Q$ be a probability kernel on $S \times \mathcal{F}$, that is: $Q(\cdot, A)$ is measurable for all $A \in \mathcal{F}$, and $Q(x, \cdot)$ is a probability measure on $(S, \mathcal{F})$ for all $x \in S$. For any
where we set $xx$ that $(I, n - 1)$ is chosen independently from $P$ non-negative integers and $N_k$ for $X$. We define $(Qg)$, or simply $Qg$, for $g \in \mathcal{B}(S)$ as soon as the integral $[\mathcal{I}]$ is well defined, and we have $Qg \in \mathcal{B}(S)$. For $n \in \mathbb{N}$, we denote by $Q^n$ the $n$-th iterate of $Q$ defined by $Q^0 = I_d$, the identity map on $\mathcal{B}(S), Q^{n+1} = Q^n(Qg)$ for $g \in \mathcal{B}(S)$.

Let $P$ be a probability kernel on $S \times \mathcal{F}^2$, that is: $P(\cdot, A)$ is measurable for all $A \in \mathcal{F}^2$, and $P(x, \cdot)$ is a probability measure on $(S^2, \mathcal{F}^2)$ for all $x \in S$. For any $g \in \mathcal{B}(S^3)$ and $h \in \mathcal{B}(S^2)$, we set for $x \in S$:

\[
(Qf)(x) = \int_S f(y) Q(x, dy).
\]

We define $(Pg)$ (resp. $(Ph)$), or simply $Pg$ for $g \in \mathcal{B}(S^3)$ (resp. $Ph$ for $h \in \mathcal{B}(S^2)$), as soon as the corresponding integral $[\mathcal{Q}]$ is well defined, and we have that $Pg$ and $Ph$ belong to $\mathcal{B}(S)$.

We now introduce some notations related to the regular binary tree. Recall that $\mathbb{N}$ is the set of non-negative integers and $\mathbb{N}^* = \mathbb{N}\setminus \{0\}$. We set $T_0 = G_0 = \{\emptyset\}, G_k = \{0, 1\}^k$ and $T_k = \bigcup_{0 \leq r \leq k} G_r$ for $k \in \mathbb{N}$, and $T = \bigcup_{k \in \mathbb{N}} G_r$. The set $G_k$ corresponds to the $k$-th generation, $T_k$ to the tree up the $k$-th generation, and $T$ the complete binary tree. For $i \in T$, we denote by $|i|$ the generation of $i$ ($|i| = k$ if and only if $i \in G_k$) and $IA = \{ij; j \in A\}$ for $A \subset T$, where $ij$ is the concatenation of the two sequences $i, j \in T$, with the convention that $\emptyset i = i \emptyset = i$.

We recall the definition of bifurcating Markov chain (BMC) from Guyon [8].

**Definition 1.1.** We say a stochastic process indexed by $T$, $X = (X_i, i \in T)$, is a bifurcating Markov chain on a measurable space $(S, \mathcal{F})$ with initial probability distribution $\nu$ on $(S, \mathcal{F})$ and probability kernel $P$ on $S \times \mathcal{F}^2$, a BMC in short, if:

- (Initial distribution.) The random variable $X_0$ is distributed as $\nu$.
- (Branching Markov property.) For a sequence $(g_i, i \in T)$ of functions belonging to $\mathcal{B}(S^3)$, we have for all $k \geq 0$,

\[
\mathbb{E}
\left[
\prod_{i \in G_k} g_i(x_i, x_{i0}, x_{i1}) | \sigma(x_j; j \in T_k)
\right]
= \prod_{i \in G_k} P_{g_i}(x_i).
\]

We define three probability kernels $P_0, P_1$ and $\Omega$ on $S \times \mathcal{F}$ by:

\[
P_0(x, A) = \mathcal{P}(x, A \times S), \quad P_1(x, A) = \mathcal{P}(x, S \times A) \quad \text{for } (x, A) \in S \times \mathcal{F}, \quad \text{and} \quad \Omega = \frac{1}{2}(P_0 + P_1).
\]

Notice that $P_0$ (resp. $P_1$) is the restriction of the first (resp. second) marginal of $\mathcal{P}$ to $S$. Following Guyon [9], we introduce an auxiliary Markov chain $Y = (Y_n, n \in \mathbb{N})$ on $(S, \mathcal{F})$ with $Y_0$ distributed as $X_0$ and transition kernel $Q$. The distribution of $Y_n$ corresponds to the distribution of $X_I$, where $I$ is chosen independently from $X$ and uniformly at random in generation $G_n$. We shall write $\mathbb{E}_x$ when $X_0 = x$ (i.e. the initial distribution $\nu$ is the Dirac mass at $x \in S$).

For all $u \in T$, we denote by $X_u = (X_u, X_{u0}, X_{u1})$ the mother-daughters triangle. One can check that $(X_u, u \in T)$ is a bifurcating Markov chain on $S^3$ with transition probability $P^\Delta$ defined by

\[
P^\Delta(x0x1, dy0y1, dz0z1) = \delta_{x0}(dy) \mathcal{P}(y, dy0, dy1) \delta_{x1}(dz) \mathcal{P}(z, dz0, dz1),
\]

where we set $xx0x1 = (x, x0, x1)$. The first and the second marginal of $P^\Delta$ are defined by:

\[
P_0^\Delta(xx0x1, dy0y1) = \delta_{x0}(dy) \mathcal{P}(y, dy0, dy1) \quad \text{and} \quad P_1^\Delta(xx0x1, dz0z1) = \delta_{x1}(dy) \mathcal{P}(z, dz0, dz1).
\]
We consider the sequence \((Y^\Delta_n, n \in \mathbb{N})\) defined recursively as follow: we set \(Y^\Delta_0 = X^\Delta_0\) and for all \(n \in \mathbb{N}\), if we are on a vertex \(u \in G_n\) of the tree, then with probability 1/2, we choose the vertex \(u0\) and we set \(Y^\Delta_{n+1} = X^\Delta_{u0}\) or we choose the vertex \(u1\) and we set \(Y^\Delta_{n+1} = X^\Delta_{u1}\). One can easily see that the sequence \((Y^\Delta_n, n \in \mathbb{N})\) is a Markov chain on \(S^3\) whose transition probability \(\Omega^\Delta\) is defined by

\[
\Omega^\Delta(x x_0 x_1, dy_0 y_1) = \frac{1}{2}(\delta_{x_0}(dy) + \delta_{x_1}(dy))\mathcal{P}(y, dy_0, dy_1).
\]

Let \(i, j \in \mathbb{T}\). We write \(i \preceq j\) if \(j \in i \mathbb{T}\). We denote by \(i \wedge j\) the most recent common ancestor of \(i\) and \(j\), which is defined as the only \(u \in \mathbb{T}\) such that if \(v \in \mathbb{T}\) and \(v \preceq i\), \(v \preceq j\) then \(v \preceq u\). We also define the lexicographic order \(i \preceq j\) if either \(i \prec j\) or \(v0 \preceq i\) and \(v1 \preceq j\) for \(v = i \wedge j\). Let \(X = (X_i, i \in \mathbb{T})\) be a BMC with kernel \(\mathcal{P}\) and initial measure \(\nu\). For \(i \in \mathbb{T}\), we define the \(\sigma\)-field:

\[
\mathcal{F}_i = \{X_u; u \in \mathbb{T}\text{ such that } u \preceq i\}.
\]

By convention, the \(\sigma\)-fields \((\mathcal{F}_i; i \in \mathbb{T})\) are nested as \(\mathcal{F}_i \subset \mathcal{F}_j\) for \(i \preceq j\).

By convention, for \(f, g \in \mathcal{B}(S)\), we define the function \(f \otimes g\) by \((f \otimes g)(x, y) = f(x)g(y)\) for \(x, y \in S\) and

\[
f \otimes \text{sym } g = \frac{1}{2}(f \otimes g + g \otimes f) \quad \text{and} \quad f \otimes 2 = f \otimes f.
\]

For \(f \in \mathcal{B}(S^3)\) and a finite subset \(A \subset \mathbb{T}\), we define:

\[
M_A(f) = \sum_{u \in A} f(X^\Delta_u).
\]

In the sequel we will also use the following notation: let \(g\) and \(h\) be two functions which depend on one variable, \(x\) say; we denote by \(g \oplus h\) the function of three variables, \(xx_0 x_1 = (x, x_0, x_1)\) say, defined by

\[
(g \oplus h)(x, x_0, x_1) = g(x_0) + h(x_1).
\]

1.2. Assumptions. Let \(X = (X_u, u \in \mathbb{T})\) be a BMC on \((S, \mathcal{S})\) with initial probability distribution \(\nu\), and probability kernel \(\mathcal{P}\). Recall that \(\mathcal{Q}\) is the induced Markov kernel. We will present the results according to the following hypothesis.

For a set \(F \subset \mathcal{B}(S)\) of \(\mathbb{R}\)-valued functions, we write \(F^2 = \{f^2; f \in F\}\), \(F \otimes F = \{f_0 \otimes f_1; f_0, f_1 \in F\}\), and \(P(E) = \{Pf; f \in E\}\) whenever a kernel \(P\) act on a set of functions \(E\). We state first a structural assumption on the set of functions we shall consider.

**Assumption 1.2.** Let \(F \subset \mathcal{B}(S)\) be a set of \(\mathbb{R}\)-valued functions such that:

(i) \(F\) is a vector subspace which contains the constants;

(ii) \(F^2 \subset F\);

(iii) \(F \subset L^1(\nu)\);

(iv) \(F \otimes F \subset L^1(\mathcal{P}(x, \cdot))\) for all \(x \in S\), and \(\mathcal{P}(F \otimes F) \subset F\).

The condition (iv) implies that \(P_0(F) \subset F\), \(P_1(F) \subset F\) as well as \(\mathcal{Q}(F) \subset F\). Notice that if \(f \in F\), then even if \(|f|\) does not belong to \(F\), using conditions (i) and (ii), we get, with \(g = (1 + f^2)/2\), that \(|f| \leq g\) and \(g \in F\).

We consider the following ergodic properties for \(\mathcal{Q}\).

**Assumption 1.3.** There exists a probability measure \(\mu\) on \((S, \mathcal{S})\) such that \(F \subset L^1(\mu)\) and for all \(f \in F\), we have the point-wise convergence \(\lim_{n \to \infty} \mathcal{Q}^n f = (\mu, f)\) and there exists \(g \in F\) with:

\[
|\mathcal{Q}^n(f)| \leq g \quad \text{for all } n \in \mathbb{N}.
\]
We consider also the following geometrical ergodicity.

**Assumption 1.4.** There exists a probability measure \( \mu \) on \((S, \mathcal{S})\) such that \( F \subset L^1(\mu) \), and \( \alpha \in (0, 1) \) such that for all \( f \in F \) there exists \( g \in F \) such that:

\[
|Q^n f - \langle \mu, f \rangle| \leq \alpha^n g \quad \text{for all } n \in \mathbb{N}.
\]

A sequence \( f = (f_\ell, \ell \in \mathbb{N}) \) of elements of \( F \) satisfies uniformly \( (\ref{eq:asympt}) \) and \( (\ref{eq:asympt2}) \) if there is \( g \in F \) such that:

\[
|Q^n(f_\ell)| \leq g \quad \text{and} \quad |Q^n f_\ell - \langle \mu, f_\ell \rangle| \leq \alpha^n g \quad \text{for all } n, \ell \in \mathbb{N}.
\]

This implies in particular that \( |f_\ell| \leq g \) and \( |\langle \mu, f_\ell \rangle| \leq \langle \mu, g \rangle \). Notice that \( (\ref{eq:asympt2}) \) trivially holds if \( f \) takes finitely distinct values (i.e. the subset \( \{f_\ell; \ell \in \mathbb{N}\} \) of \( F \) is finite) each satisfying \( (\ref{eq:asympt}) \) and \( (\ref{eq:asympt2}) \).

We consider the stronger ergodic property based on a second spectral gap.

**Assumption 1.5.** There exists a probability measure \( \mu \) on \((S, \mathcal{S})\) such that \( F \subset L^1(\mu) \), and \( \alpha \in (0, 1) \), a finite non-empty set \( J \) of indices, distinct complex eigenvalues \( \{\alpha_j, j \in J\} \) of the operator \( Q \) with \( |\alpha_j| = \alpha \), non-zero complex projectors \( \{R_j, j \in J\} \) defined on \( \mathbb{C}F \), the \( \mathbb{C} \)-vector space spanned by \( F \), such that \( R_j \circ R_{j'} = R_{j'} \circ R_j = 0 \) for all \( j \neq j' \) (so that \( \sum_{j \in J} R_j \) is also a projector defined on \( \mathbb{C}F \)) and a positive sequence \( (\beta_n, n \in \mathbb{N}) \), non-increasing, bounded from above by \( 1 \) and converging to zero, such that for all \( f \in F \), there exists \( g \in F \) and, with \( \theta_j = \alpha_j/\alpha \):

\[
|Q^n(f) - \langle \mu, f \rangle - \alpha^n \sum_{j \in J} \theta^n_j R_j(f)| \leq \beta_n \alpha^n g \quad \text{for all } n \in \mathbb{N}.
\]

We shall consider sequences \( f = (f_\ell, \ell \in \mathbb{N}) \) of elements of \( F \) which satisfies Assumption 1.5 uniformly, that is such that there exists \( g \in F \) with:

\[
|Q^n(f_\ell)| \leq g, \quad |Q^n(f_\ell)| \leq \alpha^n g \quad \text{and} \quad |Q^n(f_\ell)| \leq \beta_n \alpha^n g \quad \text{for all } n, \ell \in \mathbb{N},
\]

where

\[
\hat{f} = f - \sum_{j \in J} R_j(f) \quad \text{with} \quad \tilde{f} = f - \langle \mu, f \rangle.
\]

**Remark 1.6.** In \( \ref{sec:ergod2} \), only the ergodic Assumptions 1.2 and 1.3 were assumed. Indeed, as we will see in Section 3, if the sequence \( f = (f_\ell, \ell \in \mathbb{N}) \) is such that \( \mathcal{P}(f_\ell) = 0 \) for all \( \ell \in \mathbb{N} \), then Assumption 1.5 is not needed.

**Remark 1.7.** We recall that \( \mu \) is the invariant probability measure of \( Q \). We consider the probability measure \( \mu^\alpha \) defined on \( S^3 \) by

\[
\mu^\alpha(dx_0 x_1) = \mu(dx)\mathcal{P}(x, dx_0, dx_1).
\]

Then, for all \( f \in \mathcal{B}(S^3) \), one can easily check that \( \mu^\alpha \mathcal{Q}^\alpha f = \langle \mu, f \rangle \), that is \( \mu^\alpha \) is the invariant probability measure of \( \mathcal{Q}^\alpha \). One can also check the following: for all \( n \in \mathbb{N}^* \), we have

\[
(\mathcal{Q}^\alpha)^n f = \frac{1}{2}(\mathcal{Q}^{n-1} f \oplus \mathcal{Q}^{n-1} \mathcal{P} f).
\]

In the sequel, for all \( f \in \mathcal{B}(S^3) \), if \( \langle \mu, \mathcal{P} f \rangle \) is well defined, we will also set

\[
\hat{f} = f - \langle \mu^\alpha, f \rangle = f - \langle \mu, \mathcal{P} f \rangle.
\]

The paper is organised as follows. In Section 2, we state the main result in the sub-critical case, that is when \( 2\alpha^2 < 1 \). In Section 3, we study the special case of conditionally centered functions, that is when \( \mathcal{P}(f) = 0 \). We will see in particular that in this special case, the value of the ergodicity rate \( \alpha \in (0, 1) \) does not have any influence on the fluctuations. In Section 4, we study the critical and the super-critical cases, that is when \( 2\alpha^2 = 1 \) and \( 2\alpha^2 > 1 \).
2. The sub-critical case: $2\alpha^2 < 1$.

Assume that Assumptions 1.2 and 1.4 hold. We shall consider sequences $f = (f, \ell \in \mathbb{N})$ of elements of $\mathcal{B}(S^3)$ such that for all $\ell \in \mathbb{N}$, $\mathcal{P} f_\ell$, $\mathcal{P} f_\ell^2$, $\ell \in \mathbb{N}$ and $\mathcal{P} f_\ell^4$ exist and $(\mathcal{P} f_\ell, \ell \in \mathbb{N})$, $(\mathcal{P} f_\ell^2, \ell \in \mathbb{N})$ and $(\mathcal{P} f_\ell^4, \ell \in \mathbb{N})$ are sequences of elements of $F$. We shall also assume that (5) holds for the sequences $(\mathcal{P} f_\ell, \ell \in \mathbb{N})$, $(\mathcal{P} f_\ell^2, \ell \in \mathbb{N})$ and $(\mathcal{P} f_\ell^4, \ell \in \mathbb{N})$. For a sequence $f = (f, \ell \in \mathbb{N})$ of elements of $\mathcal{B}(S^3)$, we set

$$N_{n,0}(f) = |G_n|^{-1/2} \sum_{\ell=0}^{n} M_{G_{n-\ell}}(\hat{f}_\ell).$$

where we set $\hat{f}_\ell = f_\ell - \langle \mu, f_\ell \rangle = f_\ell - (\mu, f_\ell)$. We have the following result.

**Theorem 2.1.** Assume that Assumptions 1.2 and 1.4 hold with $\alpha \in (0, 1/\sqrt{2})$. For all sequence $f = (f, \ell \in \mathbb{N})$ of elements of $\mathcal{B}(S^3)$ such that for all $\ell \in \mathbb{N}$, $\mathcal{P} f_\ell$, $\mathcal{P} f_\ell^2$ and $\mathcal{P} f_\ell^4$ exist, $(\mathcal{P} f_\ell, \ell \in \mathbb{N})$, $(\mathcal{P} f_\ell^2, \ell \in \mathbb{N})$ and $(\mathcal{P} f_\ell^4, \ell \in \mathbb{N})$ are sequences of elements of $F$ which satisfy (5) for some $g \in F$, we have the following convergence in distribution

$$N_{n,0}(f) \xrightarrow{d, n \to \infty} G,$$

where $G$ is a Gaussian real-valued random variable with covariance $\Sigma_{\alpha, \text{sub}}(f)$ defined by

$$\Sigma_{\alpha, \text{sub}}(f) = \Sigma_{1, \text{sub}}(f) + \Sigma_{2, \text{sub}}(f),$$

with $\Sigma_{1, \text{sub}}(f)$ and $\Sigma_{2, \text{sub}}(f)$ are defined by:

$$\Sigma_{1, \text{sub}}(f) = \sum_{\ell=0}^{\infty} 2^{-\ell} \langle \mu, \mathcal{P}(\hat{f}_\ell^2) \rangle + \sum_{\ell \geq 0, k \geq 0} 2^{k-\ell} \langle \mu, \mathcal{P}(\mathcal{Q}^k(\mathcal{P}^k \hat{f}_\ell) \otimes \mathcal{Q}^k(\mathcal{P}^k \hat{f}_\ell)) \rangle,$$

and

$$\Sigma_{2, \text{sub}}(f) = \sum_{0 \leq \ell < k} 2^{-\ell-1} \langle \mu, \mathcal{P}(\mathcal{Q}^k \mathcal{Q}^{k-\ell-1} \mathcal{P} \hat{f}_\ell) \otimes \mathcal{Q}^{k-\ell-1} \mathcal{P} \hat{f}_\ell) \rangle)$$

$$\Sigma_{2, \text{sub}}(f) = \sum_{0 \leq \ell < k} 2^{r-\ell} \langle \mu, \mathcal{P}(\mathcal{Q}^r(\mathcal{P} \hat{f}_k) \otimes \mathcal{Q}^r(\mathcal{P} \hat{f}_k)) \rangle).$$

**Remark 2.2.** Let $f \in \mathcal{B}(S^3)$ such that $\mathcal{P} f$ and $\mathcal{P}(f)^2$ exist and belong to $F$. If we take $f = (f, 0, 0, \ldots)$, the infinite sequence where only the first component is non-zero, we obtain

$$|G_n|^{-1/2} M_{G_n}(f) \xrightarrow{d, n \to \infty} G_1,$$

where $G_1$ is a Gaussian real-valued random variable with covariance $\Sigma_{G, \text{sub}}(f)$ defined by

$$\Sigma_{G, \text{sub}}(f) = \langle \mu, \mathcal{P}(\hat{f}^2) \rangle + \sum_{k=0}^{\infty} 2^k \langle \mu, \mathcal{P}(\mathcal{Q}^k(\mathcal{P} \hat{f}) \otimes \mathcal{Q}^k(\mathcal{P} \hat{f}) \rangle).$$

If we take $f \in \mathcal{B}(S^3)$ such that $\mathcal{P}(f) = 0$ and $f = (f, f, \ldots)$, we recover the results of Guyon (see Theorem 19 and Corollary 20 in [8]). In fact, in this special case, we will see in the next section that the fluctuations do not depend on the values of the ergodicity rate $\alpha \in (0, 1)$. 
Moreover, for two reals $a$ and $b$, if we set $f_{a,b} = a(f - \mathcal{P}(f)) + b\mathcal{P}\hat{f}$, we have

$$\left|G_n\right|^{-1/2}M_{G_n}(f_{a,b}) \xrightarrow{(d) \ n \to \infty} G_{1,a,b}$$

where $G_{1,a,b}$ is a Gaussian real-valued random variable with covariance $\Sigma_{G}^{\triangle, \text{sub}}(f_{a,b})$ defined by

$$\Sigma_{G}^{\triangle, \text{sub}}(f_{a,b}) = a^2(\mu, \mathcal{P}f^2 - (\mathcal{P}f)^2) + b^2 \Sigma_{G}^{\text{sub}}(\mathcal{P}f),$$

where

$$(8) \quad \Sigma_{G}^{\text{sub}}(\mathcal{P}f) = \langle \mu, (\mathcal{P}\hat{f})^2 \rangle + \sum_{k=0}^{\infty} 2^k \langle \mu, \mathcal{P}(Q^k(\mathcal{P}\hat{f} \otimes Q^k(\mathcal{P}\hat{f}))) \rangle.$$ 

The latter limit implies that

$$\left(\left|G_n\right|^{-1/2}M_{G_n}(f - \mathcal{P}f) \right) \xrightarrow{(d) \ n \to \infty} G_2,$$

where $G_2$ is a two-dimensional Gaussian vector with zero mean and covariance matrix $\Sigma_{G}^{\triangle,2}(f)$ defined by

$$\Sigma_{G}^{\triangle,2}(f) = \begin{pmatrix} \langle \mu, \mathcal{P}f^2 - (\mathcal{P}f)^2 \rangle & 0 \\ 0 & \Sigma_{G}^{\text{sub}}(\mathcal{P}f) \end{pmatrix},$$

where $\Sigma_{G}^{\text{sub}}(\mathcal{P}f)$ is defined in [8]. More generally, for the subtree $T_n$, we have

$$\left(\left|T_n\right|^{-1/2}M_{T_n}(f - \mathcal{P}f) \right) \xrightarrow{(d) \ n \to \infty} G_{2,T},$$

where $G_{2,T}$ is a two-dimensional Gaussian vector with zero mean and covariance matrix $\Sigma_{T}^{\triangle, \text{sub}}(f)$ defined by

$$\Sigma_{T}^{\triangle, \text{sub}}(f) = \begin{pmatrix} \Sigma_{1,1}(f) & \Sigma_{1,2}(f) \\ \Sigma_{2,1}(f) & \Sigma_{2,2}(f) \end{pmatrix},$$

with

$$\Sigma_{1,1}(f) = \langle \mu, \mathcal{P}((f)^2 - (\mathcal{P}f)^2) \rangle,$$

$$\Sigma_{1,2}(f) = \Sigma_{2,1}(f) = \sum_{0 \leq \ell < k} 2^{-\ell} \langle \mu, \mathcal{P}((f - \mathcal{P}f)(Q^{k-\ell-1}\mathcal{P}\hat{f} \oplus Q^{k-\ell-1}\mathcal{P}\hat{f}))) \rangle,$$

$$\Sigma_{2,2}(f) = \Sigma_{G}^{\text{sub}}(\mathcal{P}f) + \sum_{0 \leq \ell < k} 2^{-\ell} \langle \mu, (\mathcal{P}\hat{f})Q^{k-\ell}\mathcal{P}\hat{f} \rangle$$

$$+ \sum_{0 \leq \ell < k, \tau \geq 0} 2^{-\ell} \langle \mu, \mathcal{P}(Q^{\tau}(\mathcal{P}\hat{f} \otimes \text{sym}, Q^{\tau+k-\ell}(\mathcal{P}\hat{f})) \rangle,$$

where $\Sigma_{G}^{\text{sub}}(\mathcal{P}f)$ is defined in [8].

Now, if we set $f_{a,b} = a(f - \mathcal{P}(f)) + b\mathcal{P}\hat{f}$ (where we recall that $\hat{f} = f - \langle \mu, \mathcal{P}f \rangle$), then we have

$$\left|G_n\right|^{-1/2}M_{G_n}(f_{a,b}) \xrightarrow{(d) \ n \to \infty} G_{1,a,b}$$

where $G_{1,a,b}$ is a Gaussian real-valued random variable with covariance $\Sigma_{G}^{\triangle, \text{sub}}'(f_{a,b})$ defined by

$$\Sigma_{G}^{\triangle, \text{sub}}'(f_{a,b}) = a^2 \langle \mu, \mathcal{P}((f)^2 - (\mathcal{P}f)^2) \rangle + b^2 \Sigma_{G}^{\text{sub}}(\mathcal{P}f) + 2ab \langle \mu, \mathcal{P}((f)^2 - (\mathcal{P}f)^2) \rangle,$$
where $\Sigma_{G,2}^{\Delta}(\mathcal{P} f)$ is given in (3). The latter limit implies that
\[
\left(\frac{|G_n|^{-1/2} M_{G_n}(f - \mathcal{P} f)}{|G_n|^{-1/2} M_{G_n}(\hat{f})}\right) \xrightarrow{(d)}_{n \to \infty} G_2',
\]
where $G_2'$ is a two-dimensional Gaussian vector with zero mean and covariance matrix $\Sigma_{G,2}^{\Delta}(\hat{f})$ defined by
\[
\Sigma_{G,2}^{\Delta}(\hat{f}) = \begin{pmatrix}
\langle \mu, \mathcal{P}(f)^2 - (\mathcal{P} f)^2 \rangle & 2\langle \mu, \mathcal{P}((f)^2 - (\mathcal{P} f)^2) \rangle \\
2\langle \mu, \mathcal{P}((f)^2 - (\mathcal{P} f)^2) \rangle & \Sigma_{G,2}^{\Delta}(\mathcal{P} f)
\end{pmatrix}.
\]
As a consequence, for $f \in \mathcal{B}(S^3)$ such that $\mathcal{P} f$ and $\mathcal{P}(f^2)$ exist and belong to $F$, we have that $M_{G_n}(f - \mathcal{P} f)$ and $M_{G_n}(\mathcal{P} f)$ are asymptotically independent, which is not the case for $M_{T_n}(f - \mathcal{P} f)$ and $M_{T_n}(\mathcal{P} f)$, and for $M_{G_n}(f - \mathcal{P} f)$ and $M_{G_n}(\mathcal{P} f)$.

**Proof of Theorem 2.7.** Note that (7) and (4) imply that for all $f \in \mathcal{B}(S^3)$ such that $\mathcal{P}(f) \in F$, there exists $g \in \mathcal{B}(S^3)$ such that
\[
\mathcal{P}(g) \in F \quad \text{and} \quad \|(Q^\Delta)^n f - \langle \mu^\Delta, f \rangle\| \leq \alpha n g \quad \text{for all } n \in \mathbb{N}.
\]
Applying the results of Theorem 3.1 in Bitseki-Delmas [3] to the bifurcating Markov chain $(X_u, u \in \mathbb{T})$, we get
\[
\Sigma_{G,2}^{\Delta}(f) = \sum_{0 \leq \ell < k} 2^{-\ell} \langle \mu^\Delta, \hat{f}_k(Q^\Delta)^k - \ell \hat{f}_\ell \rangle + \sum_{0 \leq \ell < k} 2^{-\ell} \langle \mu^\Delta, \mathcal{P}(Q^\Delta)^k \otimes \text{sym} (Q^\Delta)^{k-\ell} \hat{f}_\ell \rangle
\]
and
\[
\Sigma_{G,2}^{\Delta}(f) = \sum_{\ell \geq 0} 2^{-\ell} \langle \mu^\Delta, (\hat{f}_\ell)^2 \rangle + \sum_{\ell \geq 0, k \geq 0} 2^{-\ell} \langle \mu^\Delta, \mathcal{P}(Q^\Delta)^k \otimes (Q^\Delta)^k \hat{f}_\ell \rangle.
\]
Using (7), we obtain
\[
\langle \mu^\Delta, \mathcal{P}(Q^\Delta)^r \hat{f}_k \otimes (Q^\Delta)^k \hat{f}_\ell \rangle = \frac{1}{4} \langle \mu^\Delta, \mathcal{P}(Q^{r-1} f_{\hat{k}} \oplus Q^{r-1} f_{\hat{\ell}}) \otimes (Q^{k-\ell+r-1} f_{\hat{k}} \oplus Q^{k-\ell+r-1} f_{\hat{\ell}}) \rangle.
\]
Next, using the definition of $\mathcal{P}(\cdot)$, we obtain
\[
\frac{1}{4} \langle \mathcal{P}(Q^{r-1} f_{\hat{k}} \oplus Q^{r-1} f_{\hat{\ell}}) \otimes (Q^{k-\ell+r-1} f_{\hat{k}} \oplus Q^{k-\ell+r-1} f_{\hat{\ell}}) \rangle
\]
\[
= \frac{1}{4} \langle 2Q(Q^{r-1} f_{\hat{k}}) \otimes 2Q(Q^{\ell-r+1} f_{\hat{k}}) \rangle = Q^{r}(f_{\hat{k}}) \otimes Q^{k-\ell+r}(f_{\hat{\ell}}).
\]
From the two previous equalities and the definition of $\mu^\Delta$ we are led to
\[
\langle \mu^\Delta, \mathcal{P}(Q^\Delta)^r \hat{f}_k \otimes (Q^\Delta)^k \hat{f}_\ell \rangle = \langle \mu^\Delta, \mathcal{P}(Q^\Delta)^r \hat{f}_k \otimes Q^{k-\ell+r}(f_{\hat{\ell}}) \rangle
\]
\[
= \langle \mu, \mathcal{P}(Q^\Delta)^r \hat{f}_k \otimes Q^{k-\ell+r}(f_{\hat{\ell}}) \rangle.
\]
In the same way, from (7) and using the definition of $\mu^\Delta$, we are led to
\[
\langle \mu^\Delta, \hat{f}_k(Q^\Delta)^k \hat{f}_\ell \rangle = \langle \mu^\Delta, \hat{f}_k((1/2)(Q^{k-\ell-1} f_{\hat{\ell}}) \oplus Q^{k-\ell-1} f_{\hat{\ell}}) \rangle
\]
\[
= (1/2) \langle \mu, (Q^{k-\ell-1} f_{\hat{\ell}}) \rangle.
\]
Now, from [9] and (11) we obtain the definition of \(\Sigma_{2}^{\Delta,\text{sub}}(f)\) given in Theorem 2.1. We can obtain \(\Sigma_{1}^{\Delta,\text{sub}}(f)\) in the same way and this ends the proof.

\[
\square
\]

3. The special case of sequences \((f_\ell, \ell \in \mathbb{N})\) such that \(P f_\ell = 0\) for all \(\ell \in \mathbb{N}\).

Under the assumptions of Theorem 2.1 let \(f = (f_\ell, \ell \in \mathbb{N})\) be a sequence such that \(P f_\ell = 0\) for all \(\ell \in \mathbb{N}\). We will show that in this special case, the fluctuations does not depend to the value of the ergodicity rate \(\alpha \in (0, 1)\). For that purpose we have the following result.

**Theorem 3.1.** Under the assumptions of Theorem 2.1, let \(f = (f_\ell, \ell \in \mathbb{N})\) be a sequence such that \(P f_\ell = 0\) for all \(\ell \in \mathbb{N}\). Then for all \(\alpha \in (0, 1)\), we have the following convergence in distribution:

\[
N_{n,0}(f) \xrightarrow{(d)} G,
\]

where \(G\) is a Gaussian real-valued random variable with covariance \(\Sigma(f)\) defined by (10).

**Remark 3.2.** We stress that Theorem 3.1 allows to recover the results of Guyon [8] and the results of Delmas and Marsalle [7]. In fact, on the one hand, if we take \(f = (f, f, \cdots)\), the infinite sequence of the same function \(f\), we get

\[
|G_n|^{-1/2} M_{\log}(f) \xrightarrow{(d)} G,
\]

where \(G\) is the Gaussian real-valued random variable with mean \(\mu\) and variance \(\Sigma^2 = 2\mu, Pf^2\). In particular, using the fact that \(\lim_{n \to \infty} (\bar{T}_n/G_n) = 2\), the latter result implies that

\[
|T_n|^{-1/2} M_{\log}(f - Pf) \xrightarrow{(d)} \langle \mu, Pf^2 \rangle
\]

and we thus recover the results in [8]. Next, let \(d \in \mathbb{N}^*\). If in Theorem 3.1 we take \(f = (f_1, \cdots, f_d, 0, 0, \cdots)\) where \((f_1, \cdots, f_d) \in \mathbb{B}(S^d)^d\), with \(P f_\ell = 0\) and \(P(f_\ell^2) \in F\) for all \(\ell \in \{1, \cdots, d\}\), we get

\[
|G_n|^{-1/2} \sum_{\ell=0}^{d-1} M_{\log}(f_{\ell+1}) \xrightarrow{(d)} G,
\]

where \(G\) is a Gaussian real-valued random variable with mean \(\mu\) and variance

\[
\Sigma_G^2 = \sum_{\ell=0}^{d-1} 2^{-\ell} \langle \mu, Pf((f_{\ell+1})^2) \rangle.
\]

We thus recover, as is [7], that the fluctuations over different generations are asymptotically independent.

**Proof.** Let \((p_n, n \in \mathbb{N}^*)\) be an increasing sequence of elements of \(\mathbb{N}^*\) which satisfies for all \(\lambda > 0\):

\[
(11) \quad p_n < n, \quad \lim_{n \to \infty} \frac{p_n}{n} = 1 \quad \text{and} \quad \lim_{n \to \infty} n - p_n - \lambda \log(n) = +\infty.
\]

We have

\[
N_{n,0}(f) = R_0(n) + \Delta_n(f),
\]

where

\[
R_0(n) = |G_n|^{-1/2} \sum_{\ell=0}^{n-p-1} \sum_{u \in G_\ell} f_{n-\ell}(X_u^\Delta) \quad \text{and} \quad \Delta_n(f) = \sum_{i \in G_{n-p}} \Delta_{n,i}(f),
\]
CLT FOR BMC: EXTENSION TO THE MOTHER-DAUGHTERS TRIANGLE 9

with the $\mathcal{F}_t$-martingale increments $\Delta_{n,t}(f)$ defined by

$$\Delta_{n,t}(f) = |G_n|^{-1/2} \sum_{\ell=0}^p \sum_{u \in G_{n,\ell}} f\ell(X_u^\delta).$$

We have

$$\mathbb{E}[R_0(n)^2] = \frac{1}{|G_n|} \mathbb{E}[(\sum_{\ell=0}^{n-1} \sum_{u \in G_{n,\ell}} \tilde{f}_{n-\ell}(X_u^\delta))^2]$$

$$\leq \frac{1}{|G_n|} (\sum_{\ell=0}^{n-1} \mathbb{E}[(\sum_{u \in G_{n,\ell}} \tilde{f}_{n-\ell}(X_u^\delta))^2])^{1/2},$$

where we use the Minkowski inequality for the first inequality. By developing the term in the expectation, we get

$$\mathbb{E}[(\sum_{u \in G_{n,\ell}} \tilde{f}_{n-\ell}(X_u^\delta))^2] = \mathbb{E}[\sum_{u \neq v \in G_{n,\ell}} \tilde{f}_{n-\ell}(X_u^\delta)\tilde{f}_{n-\ell}(X_v^\delta) | X_u, X_v] + \mathbb{E}[\sum_{u \in G_{n,\ell}} \mathbb{E}[\tilde{f}_{n-\ell}^2(X_u^\delta) | X_u]]$$

$$= \mathbb{E}[\sum_{u \neq v \in G_{n,\ell}} \mathbb{P}[\tilde{f}_{n-\ell}(X_u)\tilde{f}_{n-\ell}(X_v)] + \mathbb{E}[\sum_{u \in G_{n,\ell}} \mathbb{P}(\tilde{f}_{n-\ell}^2)(X_u)]$$

$$= \mathbb{E}[(\sum_{u \in G_{n,\ell}} \mathbb{P}[\tilde{f}_{n-\ell}(X_u)])^2] + \mathbb{E}[\sum_{u \in G_{n,\ell}} (\mathbb{P}(\tilde{f}_{n-\ell}^2) - (\mathbb{P}f_{n-\ell})^2)(X_u)],$$

where we used the branching Markov property for the second inequality and the fact that $\mathbb{P}(\tilde{f}_{n-\ell}^2) - (\mathbb{P}f_{n-\ell})^2$ for the third equality. Using this last equality in (12) and using the inequalities $\sqrt{a + b} \leq \sqrt{a} + \sqrt{b}$ and $(a + b)^2 \leq 2a^2 + 2b^2$, we get

$$\mathbb{E}[R_0(n)^2] \leq \frac{1}{|G_n|} (\sum_{\ell=0}^{n-1} \mathbb{E}[(\sum_{u \neq v \in G_{n,\ell}} \tilde{f}_{n-\ell}(X_u)^2)^{1/2}] + \mathbb{E}[\sum_{u \in G_{n,\ell}} (\mathbb{P}(f_{n-\ell}^2) - (\mathbb{P}f_{n-\ell})^2)(X_u)^{1/2})])^2$$

$$\leq \frac{2}{|G_n|} (\sum_{\ell=0}^{n-1} \mathbb{E}[M_{G_{n,\ell}}(\mathbb{P}(f_{n-\ell}^2))^{1/2}]^2) + \frac{2}{|G_n|} (\sum_{\ell=0}^{n-1} \mathbb{E}[M_{G_{n,\ell}}(\mathbb{P}(f_{n-\ell})^2) - (\mathbb{P}f_{n-\ell})^2)^{1/2}]^2).$$

On the one hand, by replacing $\tilde{f}_{n-\ell}$ with $\mathbb{P}(\tilde{f}_{n-\ell})$ in the Lemma 5.2 we get

$$\lim_{n \to +\infty} \frac{2}{|G_n|} (\sum_{\ell=0}^{n-1} \mathbb{E}[M_{G_{n,\ell}}(\mathbb{P}(f_{n-\ell})^2)]^{1/2})^2 = 0.$$

On the other hand, from (19), (5) and (iii) of Assumption 1.2 we get that there is a positive constant $c$ independent of $n$ such that

$$\mathbb{E}[M_{G_{n,\ell}}(\mathbb{P}(f_{n-\ell}^2) - (\mathbb{P}f_{n-\ell})^2)] \leq c^\ell.$$

The latter inequality implies that

$$\frac{2}{|G_n|} (\sum_{\ell=0}^{n-1} \mathbb{E}[M_{G_{n,\ell}}(\mathbb{P}(f_{n-\ell}^2) - (\mathbb{P}f_{n-\ell})^2)])^{1/2}]^2 \leq \frac{2}{|G_n|} (\sum_{\ell=0}^{n-1} c^{1/2})^{1/2} \leq \frac{2c}{(\sqrt{2} - 1)^{1/2}} 2^{-p}.$$
Since \( \lim_{n \to +\infty} p_n = +\infty \), we deduce that

\[
\lim_{n \to +\infty} \frac{2}{|G_n|} \sum_{\ell=0}^{n-p-1} \mathbb{E}[M_{G_{\ell}}(\mathcal{P}(f_{n-\ell}^2) - (\mathcal{P}f_{n-\ell}^2)^2)]^{1/2} = 0.
\]

Finally, from (13) and (14), we get that

\[
\lim_{n \to \infty} \mathbb{E}[(R_0(n))^2] = 0 \quad \text{for all } \alpha \in (0, 1).
\]

Now, we consider the bracket

\[ V(n) = \sum_{i \in G_{n-p}} \mathbb{E}[(\Delta_{n,i}(f))^2 | \mathcal{F}_i]. \]

For all \( i \in G_{n-p} \), we have

\[
(\Delta_{n,i}(f))^2 = |G_n|^{-1} \sum_{\ell=0}^{p} M_{G_{n-\ell}}(f_{\ell})^2 + 2|G_n|^{-1} \sum_{0 \leq \ell < k \leq p} M_{G_{n-\ell}}(f_{\ell}) M_{G_{n-k}}(f_k).
\]

Using the branching Markov chain property, we have, for all \( 0 \leq \ell < k \leq p \):

\[
\mathbb{E}[M_{G_{n-\ell}}(f_{\ell}) M_{G_{n-k}}(f_k) | \mathcal{F}_i] = \mathbb{E}_{X_i} [M_{G_{n-k}}(f_k) M_{G_{n-\ell}}(\mathcal{P}f_{\ell})] = 0,
\]

where we used the fact that \( \mathcal{P}(f_{\ell}) = 0 \). On the other hand, using again the branching Markov property twice and (19), we have for all \( 0 \leq \ell \leq p \),

\[
\mathbb{E}[M_{G_{n-\ell}}(f_{\ell})^2 | \mathcal{F}_i] = \mathbb{E}_{X_i} [M_{G_{n-\ell}}(f_{\ell}^2)] = 2^{p-\ell} \mathcal{Q}^{p-\ell} \mathcal{P} f_{\ell}^2 (X_i).
\]

From the two latter equalities, it follows that

\[
V_n = |G_{n-p}|^{-1} \sum_{i \in G_{n-p}} \sum_{\ell=0}^{p} 2^{-\ell} \mathcal{Q}^{p-\ell} \mathcal{P} f_{\ell}^2 (X_i).
\]

We set

\[
\Sigma(f) = \sum_{\ell=0}^{\infty} 2^{-\ell} \langle \mu, \mathcal{P} f_{\ell}^2 \rangle \quad \text{and} \quad \Sigma_n(f) = \sum_{\ell=0}^{p} 2^{-\ell} \langle \mu, \mathcal{P} f_{\ell}^2 \rangle.
\]

On the one hand, we have

\[
\mathbb{E}[|V_n - \Sigma_n(f)|] \leq \sum_{\ell=0}^{p} 2^{-\ell} |G_{n-p}|^{-1} \mathbb{E}[M_{G_{n-p}}(|\mathcal{Q}^{p-\ell}(\mathcal{P} f_{\ell}^2 - \langle \mu, \mathcal{P} f_{\ell}^2 \rangle)|)]
\]

\[
\leq \sum_{\ell=0}^{p} 2^{-\ell} \alpha^{p-\ell} \langle \nu, g \rangle
\]

\[
\leq c \sum_{\ell=0}^{p} 2^{-\ell} \alpha^{p-\ell},
\]

where we used (14) for the second inequality and (iii) of Assumption 1.2 for the last inequality. Now, since

\[
\sum_{\ell=0}^{\infty} 2^{-\ell} < \infty \quad \text{and} \quad \lim_{n \to \infty} \alpha^{p-\ell} = 0 \quad \text{for all fixed } \ell,
\]

we get, using the dominated convergence theorem and (15),

\[
\lim_{n \to \infty} \mathbb{E}[|V_n - \Sigma_n(f)|] = 0.
\]
On the other hand, since \((\mathcal{P}f^k_\ell, \ell \in \mathbb{N})\) satisfies uniformly \([3]\) and \([4]\), it follows that
\[
\lim_{n \to \infty} |\Sigma(f) - \Sigma_n(f)| = 0.
\]

From the foregoing, we deduce that, for all \(\alpha \in (0, 1)\),
\[
\lim_{n \to \infty} V_n = \Sigma(f) \quad \text{in probability.} \tag{16}
\]

Finally, in order to check the Lindeberg condition, we consider \(R_3(n) = \sum_{i \in \mathbb{G}_{n-p}} E[|\Delta_{n,i}(f_i)|^4]\).

Using \((\sum_{k=0}^p a_k)^4 \leq (p + 1)^3 \sum_{k=0}^p a_k^4\), we get
\[
R_3(n) \leq (p + 1)^3 |G_n|^{-2} \sum_{\ell=0}^p \sum_{i \in \mathbb{G}_{n-p}} E[(M_{\mathbb{G}_{n-p}}(f_\ell))^4]. \tag{17}
\]

For all \(i \in \mathbb{G}_{n-p}\) and for all \(0 \leq \ell \leq p\), we have, using the branching Markov property (see the calculus in \([3]\), Remark 2.3 for more details), \([19]\), \([20]\) and \([3]\):
\[
E[M_{\mathbb{G}_{n-p}}(f_\ell)^4] \leq E[M_{\mathbb{G}_{n-p}}(\mathcal{P}f^4_\ell)] + 6 E[M_{\mathbb{G}_{n-p}}(\mathcal{P}f^2_\ell)^2
\]
\[
= 2^{p-\ell} E[Q^{p-\ell} \mathcal{P}f^4_\ell(X_i)] + 6 \mathbb{E}[2^{p-\ell} Q^{p-\ell} (\mathcal{P}f^2_\ell)^2(X_i)]
\]
\[
+ 6 \mathbb{E} \sum_{k=0}^{p-\ell-1} 2^{p-\ell-k} Q^{p-\ell-k-1}(\mathcal{P}(Q^k \mathcal{P}f^2_\ell)\otimes 2)(X_i)
\]
\[
\leq c 2^{p-\ell} \mathbb{E}[(g_1(X_i) + 2^{p-\ell} g_2(X_i))],
\]
where \(g_1, g_2 \in F\). Using \([17]\), the latter inequality, \([19]\) and \((iii)\) of Assumption \([12]\) we get
\[
R_3(n) \leq c_3 p^3 |G_n|^{-2} \sum_{\ell=0}^p 2^{p-\ell} \mathbb{E}[M_{\mathbb{G}_{n-p}}(g_1)] + c_3 p^3 |G_n|^{-2} \sum_{\ell=0}^p 2^{2(p-\ell)} \mathbb{E}[M_{\mathbb{G}_{n-p}}(g_2)]
\]
\[
\leq c_3 p^3 2^{-n} + c_3 p^3 2^{-n+p}.
\]

Using \([11]\), it follows from the foregoing that for all \(\alpha \in (0, 1)\),
\[
\lim_{n \to \infty} R_3(n) = 0.
\]

\[\square\]

4. THE CRITICAL AND THE SUPER-CRITICAL CASES: \(2\alpha^2 = 1\) AND \(2\alpha^2 > 1\).

First note that
\[
N_{n,\emptyset}(f) = |G_n|^{-1/2} \sum_{\ell=0}^n M_{\mathbb{G}_{n-\ell}}(f_\ell - \mathcal{P}f_\ell) + |G_n|^{-1/2} \sum_{\ell=0}^n M_{\mathbb{G}_{n-\ell}}(\mathcal{P}f_\ell - \langle \mu, \mathcal{P}f_\ell \rangle).
\]

Next, let \((s_n, n \in \mathbb{N})\) be a sequence of real numbers converging to 0. From Theorem \([34]\), we have
\[
s_n |G_n|^{-1/2} \sum_{\ell=0}^n M_{\mathbb{G}_{n-\ell}}(f_\ell - \mathcal{P}f_\ell) \xrightarrow{n \to \infty} 0 \quad \text{in probability,}\tag{18}
\]
so that in the critical and the super-critical case, the study of the asymptotic law of \(N_{n,\emptyset}(f)\) is reduced to the study of the asymptotic law of \(|G_n|^{-1/2} \sum_{\ell=0}^n M_{\mathbb{G}_{n-\ell}}(\mathcal{P}f_\ell - \langle \mu, \mathcal{P}f_\ell \rangle)\).

We introduce the following notation for \(k, \ell \in \mathbb{N}\):
\[
\mathcal{P}f^k_\ell = \sum_{j \in J} \theta_j^{k-\ell} \mathfrak{P}_j(\mathcal{P}f_k) \otimes_{sym} \mathfrak{P}_j(\mathcal{P}f_\ell).
\]
Then, from Theorem 4.3 in Bitseki-Delmas and Theorem 3.4 in Bitseki-Delmas applied to the functions $(\mathcal{P}f_{\ell}, \ell \in \mathbb{N})$, we get the following.

**Theorem 4.1.** Assume that Assumptions 1.2 and 1.5 hold with $\alpha = 1/\sqrt{2}$. For all sequences $f = (f_\ell, \ell \in \mathbb{N})$ of elements of $\mathcal{B}(S^3)$ such that for all $\ell \in \mathbb{N}$, $\mathcal{P}f_\ell$ and $\mathcal{P}f_{\ell}^2$ exist, $(\mathcal{P}f_{\ell}, \ell \in \mathbb{N})$, $(\mathcal{P}f_{\ell}^2, \ell \in \mathbb{N})$ and $(\mathcal{P}f_{\ell}^4, \ell \in \mathbb{N})$ are sequences of elements of $F$ which satisfy (6) for some $g \in F$, we have the following convergence in distribution:

$$n^{-1/2}N_{n,0}(f) \xrightarrow{d} G,$$

where $G$ is a Gaussian real-valued random variable with covariance $\Sigma_{\ell, \text{crit}}(f)$ defined by

$$\Sigma_{\ell, \text{crit}}(f) = \Sigma_{1, \text{crit}}(f) + \Sigma_{2, \text{crit}}(f),$$

where

$$\Sigma_{1, \text{crit}}(f) = \sum_{\ell=0}^{+\infty} 2^{-\ell} \langle \mu, \mathcal{P}f_{\ell}^* \rangle \quad \text{and} \quad \Sigma_{2, \text{crit}}(f) = \sum_{0 \leq k < \ell} 2^{-(k+\ell)/2} \langle \mu, \mathcal{P}f_{k,\ell}^* \rangle.$$

**Remark 4.2.** We stress that the covariance $\Sigma_{\ell, \text{crit}}(f)$ can take the value 0. This is the case if for all $\ell \in \mathbb{N}$, the orthogonal projection of $\mathcal{P}f_{\ell}$ on the eigen-spaces corresponding to the eigenvalues 1 and $\alpha_j$, $j \in J$, equal 0. However, the case where $\mathcal{P}f_{\ell} = 0$ for all $\ell \in \mathbb{N}$ is treated in Theorem 3.1. Indeed, we have seen that in this case the good normalisation is not $(n|\mathcal{G}_n|)^{-1/2}$, but $|\mathcal{G}_n|^{-1/2}$.

Next, from Theorem 3.9 in Bitseki-Delmas applied to the functions $(\mathcal{P}f_{\ell}, \ell \in \mathbb{N})$, we get the following result.

**Theorem 4.3.** Assume that Assumptions 1.2 and 1.5 hold with $\alpha > 1/\sqrt{2}$. For all sequences $f = (f_\ell, \ell \in \mathbb{N})$ of elements of $\mathcal{B}(S^3)$ such that for all $\ell \in \mathbb{N}$, $\mathcal{P}f_\ell$ and $\mathcal{P}f_{\ell}^2$ exist, $(\mathcal{P}f_{\ell}, \ell \in \mathbb{N})$ and $(\mathcal{P}f_{\ell}^2, \ell \in \mathbb{N})$ are sequences of elements of $F$ which satisfy (6) for some $g \in F$, we have the following convergence in probability:

$$(2\alpha^2)^{-n/2}N_{n,0}(f) \rightarrow \lim_{n \rightarrow \infty} (2\alpha)^{-\ell} \sum_{j \in J} \theta_j^{-\ell} M_{\infty,j}(\mathcal{P}(f_\ell)) \xrightarrow{P} 0.$$

**Remark 4.4.** If the sequence $f = (f_\ell, \ell \in \mathbb{N})$ is such that for all $\ell \in \mathbb{N}$, the orthogonal projection of $\mathcal{P}(f_\ell)$ on the eigen-spaces corresponding to the eigenvalues 1 and $\alpha_j$, $j \in J$, equal 0, then we have the following convergence in probability:

$$(2\alpha^2)^{-n/2}N_{n,0}(f) \xrightarrow{P} 0.$$

Once again, the case where $\mathcal{P}(f_\ell) = 0$ for all $\ell \in \mathbb{N}$ is treated in Theorem 3.1. Indeed, we have seen that in this case the good normalisation is not $(2\alpha^2)^n|\mathcal{G}_n|^{-1/2}$, but $|\mathcal{G}_n|^{-1/2}$.

5. **APPENDIX**

**Moments formula for BMC.** Let $X = (X_i, i \in \mathbb{T})$ be a BMC on $(S, \mathcal{F})$ with probability kernel $\mathcal{P}$. Recall that $|\mathcal{G}_n| = 2^n$ and $M_{\alpha,n}(f) = \sum_{x \in \mathcal{G}_n} f(X_i)$. We also recall that $2\Omega(x, A) = \mathcal{P}(x, A \times S) + \mathcal{P}(x, S \times A)$ for $A \in \mathcal{F}$. We use the convention that $\sum_\emptyset = 0$.

We recall the following well known and easy to establish many-to-one formulas for BMC.
Lemma 5.1. Let $f, g \in \mathcal{B}(S), x \in S$ and $n \geq m \geq 0$. Assuming that all the quantities below are well defined, we have:

\begin{align}
(19) \quad E_x [M_{G_n}(f)] &= |G_n| Q^n f(x) = 2^n Q^n f(x), \\
(20) \quad E_x [M_{G_n}(f)^2] &= 2^n Q^n(f^2)(x) + \sum_{k=0}^{n-1} 2^{n+k} Q^{n-k-1} (\mathbb{P}(\Omega^k f \otimes \Omega^k f))(x). 
\end{align}

We have the following lemma.

Lemma 5.2. Under the assumptions of Theorem 2.1, let $f = (f_\ell, \ell \in \mathbb{N})$ be a sequence of elements of $F$ satisfying Assumptions [7,4] uniformly, that is (7) for some $g \in F$. Let $(p_n, n \in \mathbb{N}^*)$ be an increasing sequence of elements of $\mathbb{N}^*$ which satisfies (11). Then, for all $\alpha \in (0, 1),$

$$\lim_{n \to \infty} |G_n|^{-1/2} \sum_{k=0}^{n-p-1} E[M_{G_k}(f_{n-k})^2]^{1/2} = 0$$

Proof. For all $k \geq 1$, we have:

$$E_x [M_{G_k}(f_{n-k})^2] \leq 2^k g_1(x) + \sum_{\ell=0}^{k-1} 2^{k+\ell} \alpha^{2\ell} Q^{k-\ell-1} (\mathbb{P}(g_2 \otimes g_2))(x)$$

$$\leq 2^k g_1(x) + 2^k \sum_{\ell=0}^{k-1} (2\alpha^2)^\ell g_3(x)$$

$$\leq 2^k g_4(x) \times \begin{cases} 1 & \text{if } 2\alpha^2 < 1 \\
2 & \text{if } 2\alpha^2 = 1 \\
(2\alpha^2)^k & \text{if } 2\alpha^2 > 1, \end{cases}$$

with $g_1, g_2, g_3, g_4 \in F$ and where we used (20), (3) twice and (4) twice (with $f$ and $g$ replaced by $2(g^2 + \langle \mu, g \rangle^2)$ and $g_1$, and with $f$ and $g$ replaced by $g$ and $g_2$) for the first inequality, (3) (with $f$ and $g$ replaced by $\mathbb{P}(g_2 \otimes g_2_2)$ and $g_3$) for the second, and $g_4 = g_1 + (1 + |1 - 2\alpha^2|^{-1})g_3$ for the last. We deduce that:

$$|G_n|^{-1/2} \sum_{k=0}^{n-p-1} E[M_{G_k}(f_{n-k})^2]^{1/2} \leq c 2^{-n/2} \times \begin{cases} \sum_{k=0}^{n-p-1} 2^{k/2} & \text{if } 2\alpha^2 < 1 \\
\sum_{k=0}^{n-p-1} k^{1/2} 2^{k/2} & \text{if } 2\alpha^2 = 1 \\
\sum_{k=0}^{n-p-1} (2\alpha^2)^{k/2} 2^{k/2} & \text{if } 2\alpha^2 > 1 \end{cases}$$

$$\leq 3c \begin{cases} 2^{-p/2} & \text{if } 2\alpha^2 = 1 \\
(n-p)^{1/2} 2^{-p/2} & \text{if } 2\alpha^2 = 1 \\
(2\alpha^2)^{(n-p)/2} 2^{-p/2} & \text{if } 2\alpha^2 > 1. \end{cases}$$

Use that $\lim_{n \to \infty} p_n = \infty$ and $\lim_{n \to \infty} p_n/n = 1$ to conclude. \hfill \Box

REFERENCES

[1] B. Bercu, B. De Saporta, and A. Gégout-Petit. Asymptotic analysis for bifurcating autoregressive processes via a martingale approach. *Electronic Journal of Probability*, 14:2492–2526, 2009.

[2] S. V. Bitseki Penda and J.-F. Delmas. Central limit theorem for bifurcating Markov chains under $L^2$ ergodic conditions. *Advances in Applied Probability*, pages 1–33, 2022.

[3] S. V. Bitseki Penda and J.-F. Delmas. (in press) Central limit theorem for bifurcating Markov chains under pointwise ergodic conditions. *Ann. Appl. Probab.*, 2022.

[4] S. V. Bitseki Penda and H. Djellout. Deviation inequalities and moderate deviations for estimators of parameters in bifurcating autoregressive models. *Ann. Inst. H. Poincaré Probab. Statist.*, 50(3):806–844, 08 2014.
[5] S. V. Bitseki Penda, H. Djellout, and A. Guillin. Deviation inequalities, moderate deviations and some limit theorems for bifurcating Markov chains with application. *Ann. Appl. Probab.*, 24(1):235–291, 2014.
[6] S. V. Bitseki Penda and A. Olivier. Autoregressive functions estimation in nonlinear bifurcating autoregressive models. *Stat. Inference Stoch. Process.*, 20(2):179–210, 2017.
[7] J.-F. Delmas and L. Marsalle. Detection of cellular aging in a Galton-Watson process. *Stochastic Process. Appl.*, 120(12):2495–2519, 2010.
[8] J. Guyon. Limit theorems for bifurcating Markov chains. Application to the detection of cellular aging. *Ann. Appl. Probab.*, 17(5-6):1538–1569, 2007.

S. Valère Bitseki Penda, IMB, CNRS-UMR 5584, Université Bourgogne Franche-Comté, 9 avenue Alain Savary, 21078 Dijon Cedex, France.

*Email address:* simeon-valere.bitseki-penda@u-bourgogne.fr