



**HAL**  
open science

## Deviation inequalities, Moderate deviations and some limit theorems for bifurcating Markov chains with application.

Valère Bitseki Penda, Hacène Djellout, Arnaud Guillin

### ► To cite this version:

Valère Bitseki Penda, Hacène Djellout, Arnaud Guillin. Deviation inequalities, Moderate deviations and some limit theorems for bifurcating Markov chains with application.. *The Annals of Applied Probability*, 2014, 24 (1), pp.235-291. hal-00646588v1

**HAL Id: hal-00646588**

**<https://hal.science/hal-00646588v1>**

Submitted on 30 Nov 2011 (v1), last revised 17 Jan 2014 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# DEVIATION INEQUALITIES, MODERATE DEVIATIONS AND SOME LIMIT THEOREMS FOR BIFURCATING MARKOV CHAINS WITH APPLICATION.

S.VALÈRE BITSEKI PENDA, HACÈNE DJELLOUT, AND ARNAUD GUILLIN

ABSTRACT. Firstly, under geometric ergodicity assumption, we provide some limit theorems and some probability inequalities for bifurcating Markov chains introduced by Guyon to detect cellular aging from cell lineage, thus completing the work of Guyon. This probability inequalities are then applied to derive a first result on moderate deviation principle for a functional of bifurcating Markov chains with a restricted range of speed, but with a function which can be unbounded. Next, under uniform geometric ergodicity assumption, we provide deviation inequalities for bifurcating Markov chains and apply them to derive a second result on moderate deviation principle for bounded functional of bifurcating Markov chains with a more larger range of speed. As statistical applications, we provide superexponential convergence in probability and deviation inequalities (under the gaussian setting or the bounded setting), and moderate deviation principle for least square estimators of the parameters of a first order bifurcating autoregressive process.

*Key words:* Bifurcating Markov chains, limit theorems, ergodicity, deviation inequalities, moderate deviation, martingale, first order bifurcating autoregressive process, cellular aging.

*AMS 2000 subject classifications.* Primary 60F05, 60F10, 60F15, 60E15; Secondary 60G42, 60J05, 62M02, 62M05, 62P10

## 1. INTRODUCTION

Bifurcating Markov chains (BMC) are an adaptation of (usual) Markov chains to the data of regular binary tree (see below for a more precise definition). In other terms, it is a Markov chain for which the index set is a regular binary tree. They are appropriate for example in modeling of cell lineage data when each cell in one generation gives birth to two offspring in the next one. Recently, they have received a great attention because of the experiments made by biologists on aging of the *Escherichia Coli* (see [15], [13]). *E. Coli* is a rod-shaped bacterium which reproduces by dividing in the middle, thus producing two cells, one which already existed and that we call old pole progeny, and the other which is new and that we call new pole progeny. The aim of their experiments were to look for evidence of aging in *E. Coli*. In this section, we will introduce the model that allowed the authors of [13] to study the aging of *E. Coli* and we refer to their works for further motivations and insights on the datas leading to the model studied here. This model is a typical example of bifurcating markovian dynamics and he has been the motivation for the study of BMC in [12]. This also motivates in the sequel Section 2 and Section 3, where we give a rigorous asymptotic (and non asymptotic) study of BMC under geometric ergodicity and uniform geometric ergodicity assumptions.

### 1.1. The model.

Let  $\mathbb{T}$  be a binary regular tree in which each vertex is seen as a positive integer different from 0, see Figure 1. For  $r \in \mathbb{N}$ , let

$$\mathbb{G}_r = \{2^r, 2^r + 1, \dots, 2^{r+1} - 1\}, \quad \mathbb{T}_r = \bigcup_{q=0}^r \mathbb{G}_q,$$

which denote respectively the  $r$ -th column and the first  $(r + 1)$  columns of the tree.

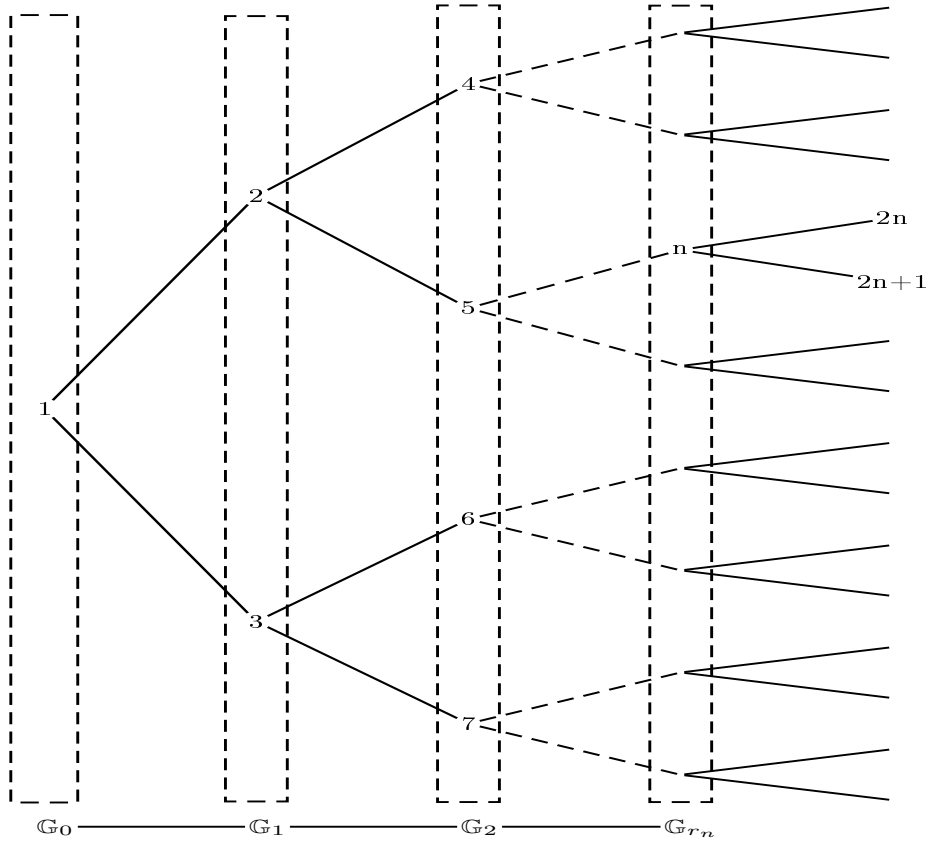


FIGURE 1. The binary tree  $\mathbb{T}$

Then, the cardinality  $|\mathbb{G}_r|$  of  $\mathbb{G}_r$  is  $2^r$  and that of  $\mathbb{T}_r$  is  $|\mathbb{T}_r| = 2^{r+1} - 1$ . A column of a given integer  $n$  is  $\mathbb{G}_{r_n}$  with  $r_n = \lfloor \log_2 n \rfloor$ , where  $\lfloor x \rfloor$  denotes the integer part of the real number  $x$ . The genealogy of the cells is described by this tree. In the sequel we will thus see  $\mathbb{T}$  as a given population. Then the vertex  $n$ , the column  $\mathbb{G}_r$  and the first  $(r + 1)$  columns  $\mathbb{T}_r$  designate respectively individual  $n$ , the  $r$ -th generation and the first  $(r + 1)$  generations. The initial individual is denoted 1.

Guyon & Al. ([13], [12]) have proposed the following linear gaussian model to describe the evolution of the growth rate of the population of cells derived from an initial individual

$$\mathcal{L}(X_1) = \nu, \quad \text{and} \quad \forall n \geq 1, \quad \begin{cases} X_{2n} = \alpha_0 X_n + \beta_0 + \varepsilon_{2n} \\ X_{2n+1} = \alpha_1 X_n + \beta_1 + \varepsilon_{2n+1}, \end{cases} \quad (1.1)$$

where  $X_n$  is the growth rate of individual  $n$ ,  $n$  is the mother of  $2n$  (the new pole progeny cell) and  $2n + 1$  (the old pole progeny cell),  $\nu$  is a distribution probability on  $\mathbb{R}$ ,  $\alpha_0, \alpha_1 \in (-1, 1)$ ;  $\beta_0, \beta_1 \in \mathbb{R}$  and  $((\varepsilon_{2n}, \varepsilon_{2n+1}), n \geq 1)$  forms a sequence of i.i.d bivariate random variables with law  $\mathcal{N}_2(0, \Gamma)$ , where

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

The processes  $(X_n)$  define by (1.1) is a typical example of BMC which is called first order bifurcating autoregressive processes (BAR(1)).

In [12], Guyon, using the theory of BMC, gives laws of large numbers and central limit theorem for the least-square estimators  $\hat{\theta}^r = (\hat{\alpha}_0^r, \hat{\beta}_0^r, \hat{\alpha}_1^r, \hat{\beta}_1^r)$  of the 4-dimensional parameter  $\theta = (\alpha_0, \beta_0, \alpha_1, \beta_1)$ , (see section (4) for a more precise definition). He also gives some statistical tests which allow to test if the model is symmetric or not, and if the new pole and the old poles populations are even distinct in mean, which allows him to conclude a statistical evidence in aging in E. Coli. Let us also mention [3], where Bercu & Al. using the martingale approach give asymptotic analysis of the least squares estimators of the unknown parameters of a general asymmetric  $p$ th-order BAR processes.

In this paper, we will give moderate deviation principle (MDP) for this estimator and the statistical tests done by Guyon. We will also give deviation inequalities for  $\hat{\theta}^r - \theta$ , which are important for a rigorous (non asymptotic) statistical study. This will be done in two cases: the gaussian case as described above and the case where the noise is assumed to take values in a compact set. Note that the latter case implies that the BAR(1) process defined by (1.1) values in compact set.

We are now going to give a rigorous definition of BMC. We refer to [12] for more details.

## 1.2. Definitions.

For an individual  $n \in \mathbb{T}$ , we are interested in the quantity  $X_n$  (it may be the weight, the growth rate,  $\dots$ ) with values in the metric space  $S$  endowed with its Borel  $\sigma$ -field  $\mathcal{S}$ .

**Definition 1.1** ( $\mathbb{T}$ -transition probability, see ([12])). *We call  $\mathbb{T}$ -transition probability any mapping  $P : S \times \mathcal{S}^2 \rightarrow [0, 1]$  such that*

- $P(\cdot, A)$  is measurable for all  $A \in \mathcal{S}^2$ ,
- $P(x, \cdot)$  is a probability measure on  $(\mathcal{S}^2, \mathcal{S}^2)$  for all  $x \in S$ .

For a  $\mathbb{T}$ -transition probability  $P$  on  $S \times \mathcal{S}^2$ , we denote by  $P_0, P_1$  and  $Q$  respectively the first and the second marginal of  $P$ , and the mean of  $P_0$  and  $P_1$ , that is  $P_0(x, B) = P(x, B \times S)$ ,  $P_1(x, B) = P(x, S \times B)$  for all  $x \in S$  and  $B \in \mathcal{S}$  and  $Q = \frac{P_0 + P_1}{2}$ .

For  $p \geq 1$ , we denote by  $\mathcal{B}(S^p)$  (resp.  $\mathcal{B}_b(S^p), \mathcal{C}(S^p), \mathcal{C}_b(S^p)$ ) the set of all  $S^p$ -measurable (resp.  $S^p$ -measurable and bounded, continuous, continuous and bounded) mapping  $f : S^p \rightarrow \mathbb{R}$ . For  $f \in \mathcal{B}(S^3)$ , when it is defined, we denote by  $Pf \in \mathcal{B}(S)$  the function

$$x \mapsto Pf(x) = \int_{S^2} f(x, y, z)P(x, dydz).$$

**Definition 1.2** (Bifurcating Markov Chains, see ([12])). *Let  $(X_n, n \in \mathbb{T})$  be a family of  $S$ -valued random variables defined on a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_r, r \in \mathbb{N}), \mathbb{P})$ . Let  $\nu$  be a probability on  $(S, \mathcal{S})$  and  $P$  be a  $\mathbb{T}$ -transition probability. We say that  $(X_n, n \in \mathbb{T})$  is a  $(\mathcal{F}_r)$ -bifurcating Markov chain with initial distribution  $\nu$  and  $\mathbb{T}$ -transition probability  $P$  if*

- $X_n$  is  $\mathcal{F}_{r_n}$ -measurable for all  $n \in \mathbb{T}$ ,
- $\mathcal{L}(X_1) = \nu$ ,
- for all  $r \in \mathbb{N}$  and for all family  $(f_n, n \in \mathbb{G}_r) \subseteq \mathcal{B}_b(S^3)$

$$\mathbb{E} \left[ \prod_{n \in \mathbb{G}_r} f_n(X_n, X_{2n}, X_{2n+1}) / \mathcal{F}_r \right] = \prod_{n \in \mathbb{G}_r} Pf_n(X_n)$$

In the following, when unprecised, the filtration implicitly used will be  $\mathcal{F}_r = \sigma(X_i, i \in \mathbb{T}_r)$ . We denote by  $(Y_r, r \in \mathbb{N})$  the Markov chain on  $S$  with  $Y_0 = X_1$  and transition probability  $Q$ . The chain  $(Y_r, r \in \mathbb{N})$  corresponds to a random lineage taken in the population.

We denote by  $\Xi$  the set of all permutations of  $\mathbb{N}^*$  that leaves each  $\mathbb{G}_r$  invariant. We draw a permutation  $\pi$  uniformly on  $\Xi$ , independently of  $X = (X_n, n \in \mathbb{T})$ .  $\pi$  allows to define a random order on  $\mathbb{T}$  which preserves the genealogical order.

For all  $i \in \mathbb{T}$ , set  $\Delta_i = (X_i, X_{2i}, X_{2i+1})$  and define the following empirical quantities

$$\begin{cases} \overline{M}_{\mathbb{G}_r}(f) = \frac{1}{|\mathbb{G}_r|} \sum_{i \in \mathbb{G}_r} f(X_i) \\ \overline{M}_{\mathbb{T}_r}(f) = \frac{1}{|\mathbb{T}_r|} \sum_{i \in \mathbb{T}_r} f(X_i) \\ \overline{M}_n^\pi(f) = \frac{1}{n} \sum_{i=1}^n f(X_{\pi(i)}) \end{cases} \quad \text{if } f \in \mathcal{B}(S) \quad (1.2)$$

and

$$\begin{cases} \overline{M}_{\mathbb{G}_r}(f) = \frac{1}{|\mathbb{G}_r|} \sum_{i \in \mathbb{G}_r} f(\Delta_i) \\ \overline{M}_{\mathbb{T}_r}(f) = \frac{1}{|\mathbb{T}_r|} \sum_{i \in \mathbb{T}_r} f(\Delta_i) \\ \overline{M}_n^\pi(f) = \frac{1}{n} \sum_{i=1}^n f(\Delta_{\pi(i)}) \end{cases} \quad \text{if } f \in \mathcal{B}(S^3). \quad (1.3)$$

Guyon in [12] studied limit theorems of the empirical means (1.2), (1.3), namely law of large numbers (in  $L^2$  and for some almost surely) and central limit theorems for (1.3) with a conditional recentering. An extension of the BMC have been proposed in [6], in which the authors study a model of the BMC with missing data. To take into account the possibility for a cell to die, the authors in [6] use Galton-Watson tree instead of a regular tree. And they give a weak law of large number, an invariance principle and the central limit result for the average over one generation or up to one generation. One can also mention the work of De Saporta & Al. [5] dealing with bifurcating autoregressive processes with missing data in the estimation procedure of the parameters of the asymmetric BAR process. They use a two type Galton-Watson process to model the genealogy and give convergence and asymptotic

normality of their estimators. Now, the previous work do not take into account the study of the convergence speed.

### 1.3. Objectives.

Our objectives in this paper are:

- to give some limit theorems for BMC that complete those done in [12] (CLT, LIL,...);
- to give probability inequalities and deviation inequalities for the empirical means (1.2) and (1.3), i.e. for  $f \in \mathcal{B}(S)$  and all  $x > 0$

$$\mathbb{P}(\overline{M}_{\mathbb{T}_r} - (\mu, f) \geq x) \leq e^{-C(x)}$$

where  $C(x)$  will crucially depends on our set of assumptions on  $f$  and on the ergodic property of  $P$ ;

- to study moderate deviation principle (MDP) for BMC, i.e. for some range of speed  $\sqrt{r} \ll b_r \ll r$  (depending on assumptions) and for  $f \in \mathcal{C}_b(S^3)$  with  $Pf = 0$

$$\frac{b_{\mathbb{T}_r}^2}{\mathbb{T}_r} \log \mathbb{P}\left(\frac{1}{b_{\mathbb{T}_r}} M_{\mathbb{T}_r} \geq x\right) \sim -\frac{x^2}{2\sigma^2}.$$

- to obtain the MDP and deviation inequalities for the estimator of bifurcating autoregressive process, which are important for a rigorous statistical study.

All these results will be obtained under hypothesis of geometric ergodicity or uniform geometric ergodicity.

Limit theorems, given in this paper, include strong law of large numbers for the empirical average  $\overline{M}_n^n(f)$  with  $f \in \mathcal{B}(S)$  (this case is not studied in [12]), the law of the iterated logarithm and the almost sure functional central limit theorem. Strong law of large numbers will be done via calculation of 4th order moments. We thus generalize the computation of 2nd order moments made by Guyon in [12]. It will be noted that the technique we will use can be applied to compute the other higher order moments but at the price of huge and tedious computations.

Deviation inequalities will be obtained in the setting of unbounded functions, by using the classical Markov inequality and under geometric ergodicity assumption. We will restrict ourselves to the use of 2nd order moment (refing results of Guyon), sufficient to obtain the MDP for a restricted range of speed, but it will be very clear that these inequalities may be improved using higher order moments.

Exponential deviation inequalities will be done in the setting of bounded functions and under uniform geometric ergodicity assumption, using intensively Azuma-Bennet-Hoeffding inequality [1], [2], [14], which we can only apply on bounded random variables. Extension to unbounded functions and weaker ergodicity assumptions will be done in a further work.

The MDP will be deducted in large part from these inequalities. The speed will depend from the fact that we are under hypothesis of geometric ergodicity or uniform geometric ergodicity.

Before the presentation of the plan of our paper, let us recall the definition of a moderate deviation principle (MDP): let  $(b_n)_{n \geq 0}$  be a positive sequence such that

$$\frac{b_n}{n} \xrightarrow{n \rightarrow \infty} 0, \quad \text{and} \quad \frac{b_n^2}{n} \xrightarrow{n \rightarrow \infty} \infty.$$

We say that a sequence of centered random variables  $(M_n)_n$  with topological state space  $(S, \mathcal{S})$  satisfies a MDP with speed  $\frac{b_n^2}{n}$  and rate function  $I : S \rightarrow \mathbb{R}_+^*$  if for each  $A \in \mathcal{S}$ ,

$$-\inf_{x \in A^\circ} I(x) \leq \liminf_{n \rightarrow \infty} \frac{n}{b_n^2} \log \mathbb{P} \left( \frac{n}{b_n} M_n \in A \right) \leq \limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \log \mathbb{P} \left( \frac{n}{b_n} M_n \in A \right) \leq -\inf_{x \in \bar{A}} I(x),$$

here  $A^\circ$  and  $\bar{A}$  denote the interior and closure of  $A$  respectively.

The MDP is an intermediate behavior between the central limit theorem ( $b_n = b\sqrt{n}$ ) and Large deviation ( $b_n = bn$ ). Usually, MDP has a simpler rate function inherited from the approximated Gaussian process, and holds for a larger class of dependent random variables than the large deviation principle.

The paper is then organised as follows. Section 2 states the moments control and its consequences. We shall state in this section a first result on the MDP for BMC in a general framework, but with small speed. Let us note that 2nd order moments are sufficient to obtain this result. Section 3 deals with the exponential inequalities and its consequences. In this section, we shall generalize the MDP done in section 2, but in a restricted framework. In section 4, we will interest particularly to the first order bifurcating autoregressive processes. Section 5 is devoted to recalling some definitions and some limit theorems for martingales used intensively in the paper, we have included them here for completeness.

## 2. MOMENTS CONTROL AND CONSEQUENCES

Let  $F$  a vector subspace of  $\mathcal{B}(S)$  such that

- (i)  $F$  contains the constants,
- (ii)  $F^2 \subset F$ ,
- (iii)  $F \otimes F \subset L^1(P(x, \cdot))$  for all  $x \in S$ , and  $P(F \otimes F) \subset F$ ,
- (iv) there exists a probability  $\mu$  on  $(S, \mathcal{S})$  such that  $F \subset L^1(\mu)$  and  $\lim_{r \rightarrow \infty} \mathbb{E}_x [f(Y_r)] = (\mu, f)$  for all  $x \in S$  and  $f \in F$ ,
- (v) for all  $f \in F$ , there exists  $g \in F$  such that for all  $r \in \mathbb{N}$ ,  $|Q^r f| \leq g$ ,
- (vi)  $F \subset L^1(\nu)$

The following hypothesis is about the geometric ergodicity of  $Q$ :

**(H1)** Assume that for all  $f \in F$  such that  $(\mu, f) = 0$ , there exists  $g \in F$  such that for all  $r \in \mathbb{N}$  and for all  $x \in S$ ,  $|Q^r f(x)| \leq \alpha^r g(x)$  for some  $\alpha \in (0, 1)$ , that is the Markov chain  $(Y_r, r \in \mathbb{N})$  is geometrically ergodic.

Recall that under this hypothesis, Guyon [12] has shown the weak law of large numbers for the three empirical average  $\overline{M}_{\mathbb{G}_r}(f)$ ,  $\overline{M}_{\mathbb{T}_r}(f)$  and  $\overline{M}_n^\pi(f)$  (in [12]: see theorem 11 when  $f \in F$  and

theorem 12 when  $f \in \mathcal{B}(S^3)$ ) and the strong law of large numbers only for  $\overline{M}_{\mathbb{G}_r}(f)$ ,  $\overline{M}_{\mathbb{T}_r}(f)$  (in [12]: see theorem 14 and corollary 15 when  $f \in F$  and theorem 18 when  $f \in \mathcal{B}(S^3)$ ).

When  $f \in \mathcal{B}(S^3)$  and under the additional hypothesis  $Pf^2$  and  $Pf^4$  exist and belong to  $F$ , he proved the central limit theorem for  $\overline{M}_{\mathbb{T}_r}(f)$  and  $\overline{M}_n^\pi(f)$  (in [12]: see theorem 19 and corollary 21). Recall that the central limit theorem for the three empirical means (1.2) when  $f \in \mathcal{B}(S)$  is still an open question, see [6] for more precision.

In this section, we complete these results by showing the strong law of large numbers for  $\overline{M}_n^\pi(f)$ , when  $f \in F$ . We prove also the law of the iterated logarithm (LIL) and almost sure functional central limit theorem (ASFCLT) for  $\overline{M}_n^\pi(f)$  when  $f \in \mathcal{B}(S^3)$ .

## 2.1. Control of the 4-th order moments.

In order to establish limit theorems below, let us state the following

**Theorem 2.1.** *Let  $F$  satisfy (i)-(vi). Let  $f \in F$  such that  $(\mu, f) = 0$ . We assume hypothesis **(H1)**. Then for all  $r \in \mathbb{N}$*

$$\mathbb{E} \left[ \overline{M}_{\mathbb{G}_r}(f)^4 \right] \leq \begin{cases} c \left(\frac{1}{4}\right)^r & \text{if } \alpha^2 < \frac{1}{2} \\ cr^2 \left(\frac{1}{4}\right)^r & \text{if } \alpha^2 = \frac{1}{2} \\ c\alpha^{4r} & \text{if } \alpha^2 > \frac{1}{2} \end{cases} \quad (2.1)$$

where the positive constant  $c$  depends on  $\alpha$  and  $f$  (and may differ line by line).

*Proof.* First note that  $f(X_i) \in L^4$  for all  $i \in \mathbb{G}_r$ . Indeed, let  $(z_1, \dots, z_r) \in \{0, 1\}^r$  the unique path in the binary tree from the root 1 to  $i$ . Then,

$$\mathbb{E} \left[ f^4(X_i) \right] = \nu P_{z_1} \cdots P_{z_r} f^4,$$

and from hypothesis (ii), (iii) and (vi) we conclude that  $\nu P_{z_1} \cdots P_{z_r} f^4 < \infty$ .

Now, the rest of the proof divides into two parts.

**Part 1. Computation of  $\mathbb{E} \left[ \overline{M}_{\mathbb{G}_r}(f)^4 \right]$ .**

Independently on  $X$ , let us draw four independent indices  $I_r, J_r, K_r$  and  $L_r$  uniformly from  $\mathbb{G}_r$ . Then

$$\mathbb{E} \left[ \overline{M}_{\mathbb{G}_r}(f)^4 \right] = \mathbb{E} \left[ f(X_{I_r}) f(X_{J_r}) f(X_{K_r}) f(X_{L_r}) \right].$$

For all  $p \in \{0, \dots, r\}$ , let us define the following events:

- $E_0^p$ : "The ancestors of  $I_r, J_r, K_r$  and  $L_r$  are different in  $\mathbb{G}_p$ ".
- $E_1^p$ : "Exactly two of  $I_r, J_r, K_r$  and  $L_r$  have the same ancestor in  $\mathbb{G}_p$ ".
- $E_2^p$ : " $I_r, J_r, K_r$  and  $L_r$  have the same ancestor two by two in  $\mathbb{G}_p$ ".
- $E_3^p$ : "Exactly three of  $I_r, J_r, K_r$  and  $L_r$  have the same ancestor in  $\mathbb{G}_p$ ".
- $E_4^p$ : " $I_r, J_r, K_r$  and  $L_r$  have the same ancestor in  $\mathbb{G}_p$ ".



In the sequel we do the convention that  $E_0^{r+1}$  is a certain event. Then after successive conditioning by events  $E_i^p$  for  $p \in \{0, \dots, r\}$  and  $i \in \{0, \dots, 4\}$ , we have

$$\begin{aligned} \mathbb{E}\left[f(X_{I_r})f(X_{J_r})f(X_{K_r})f(X_{L_r})\right] &= \mathbb{E}\left[f(X_{I_r})f(X_{J_r})f(X_{K_r})f(X_{L_r})/E_0^2\right] \times \mathbb{P}(E_0^2) \quad (2.2) \\ &+ \sum_{p=2}^r \mathbb{E}\left[f(X_{I_r})f(X_{J_r})f(X_{K_r})f(X_{L_r})/E_0^{p+1}, E_1^p\right] \times \mathbb{P}(E_1^p)\mathbb{P}(E_0^{p+1}/E_1^p) \\ &+ \sum_{p=2}^r \mathbb{E}\left[f(X_{I_r})f(X_{J_r})f(X_{K_r})f(X_{L_r})/E_0^{p+1}, E_2^p\right] \times \mathbb{P}(E_2^p)\mathbb{P}(E_0^{p+1}/E_2^p) \\ &\quad + \mathbb{E}\left[f(X_{I_r})f(X_{J_r})f(X_{K_r})f(X_{L_r})/E_3^r\right] \times \mathbb{P}(E_3^r) \\ &\quad + \mathbb{E}\left[f(X_{I_r})f(X_{J_r})f(X_{K_r})f(X_{L_r})/E_4^r\right] \times \mathbb{P}(E_4^r). \end{aligned}$$

Now it is easy to see that

$$\mathbb{P}(E_0^2) = \frac{3}{32}; \quad \mathbb{P}(E_1^r) = 3\frac{2^r - 1}{2^{2r}}; \quad \mathbb{P}(E_2^r) = 6\frac{2^r - 1}{2^{3r}}; \quad \mathbb{P}(E_3^r) = 4\frac{2^r - 1}{2^{3r}}; \quad \mathbb{P}(E_4^r) = 6\frac{1}{2^{3r}};$$

and for  $p \in \{2, \dots, r-1\}$ ,

$$\mathbb{P}(E_1^p)\mathbb{P}(E_0^{p+1}/E_1^p) = \frac{3}{2}\frac{2^p - 1}{2^{2p}} \quad \text{and} \quad \mathbb{P}(E_2^p)\mathbb{P}(E_0^{p+1}/E_2^p) = \frac{6}{4}\frac{2^p - 1}{2^{3p}}.$$

Indeed, the realization of " $E_1^p \cap E_0^{p+1}$ " can be seen as "draw uniformly four independent indices from  $\mathbb{G}_p$  such that two are the same and others are different and the two indices which are the same take different paths at  $\mathbb{G}_{p+1}$ ". Similarly, the realization of " $E_2^p \cap E_0^{p+1}$ " may be interpreted as "draw uniformly four independent indices from  $\mathbb{G}_p$  which are the same two by two and all the indices take different path at  $\mathbb{G}_{p+1}$ ".

We are going now compute each term which appears in (2.2). We do the following convention  $P(Q^{-1}f \otimes Q^{-1}f) = f^2$ .

$$(a) \quad \mathbb{E}\left[f(X_{I_r})f(X_{J_r})f(X_{K_r})f(X_{L_r})/E_4^r\right] = \nu Q^r f^4.$$

(b) Conditionally on  $E_3^r$ , we may assume that the indices  $I_r$ ,  $K_r$  and  $L_r$  are the same. We then have

$$\begin{aligned} \mathbb{E}\left[f(X_{I_r})f(X_{J_r})f(X_{K_r})f(X_{L_r})/E_3^r\right] &= \mathbb{E}\left[f^3(X_{I_r})f(X_{J_r})/E_3^r\right] \\ &= \frac{1}{2} \sum_{p=0}^{r-1} 2^{-p-1} \nu Q^p P(Q^{r-p-1}f^3 \otimes Q^{r-p-1}f) + \frac{1}{2} \sum_{p=0}^{r-1} 2^{-p-1} \nu Q^p P(Q^{r-p-1}f \otimes Q^{r-p-1}f^3). \end{aligned}$$

(c) Let  $p \in \{2, \dots, r\}$ . Conditionally on  $E_2^p$  and  $E_0^{p+1}$  we may assume that  $I_r$  and  $J_r$  have the same ancestor at  $\mathbb{G}_p$  and  $K_r$  and  $L_r$  have the same ancestor at  $\mathbb{G}_p$ . We thus have



$$\begin{aligned}
& + \nu P (Q^{p-1} P (Q^{r-p-1} f \otimes Q^{r-p-1} f) \otimes P (Q^{r-2} f \otimes Q^{r-2} f)) \\
& + \nu P (P (Q^{r-2} f \otimes Q^{p-2} P (Q^{r-p-1} f \otimes Q^{r-p-1} f)) \otimes Q^{r-1} f) \\
& + \nu P (P (Q^{p-2} P (Q^{r-p-1} f \otimes Q^{r-p-1} f) \otimes Q^{r-2} f) \otimes Q^{r-1} f) \\
& + \nu P (Q^{r-1} f \otimes P (Q^{r-2} f \otimes P (Q^{r-p-1} f \otimes Q^{r-p-1} f))) \\
& + \nu P (Q^{r-1} f \otimes P (Q^{p-2} P (Q^{r-p-1} f \otimes Q^{r-p-1} f) \otimes Q^{r-2} f)) \Big\}.
\end{aligned}$$

(e) Finally,

$$\begin{aligned}
& \mathbb{E} \left[ f(X_{I_r}) f(X_{J_r}) f(X_{K_r}) f(X_{L_r}) / E_0^2 \right] \\
& = \mathbb{E} \left[ \mathbb{E} \left[ \mathbb{E} \left[ f(X_{I_r}) f(X_{J_r}) f(X_{K_r}) f(X_{L_r}) / \mathcal{F}_2 \right] / \mathcal{F}_1 \right] / E_0^2 \right] \\
& = \mathbb{E} \left[ P (Q^{r-2} f \otimes Q^{r-2} f) (X_2) P (Q^{r-2} f \otimes Q^{r-2} f) (X_3) / E_0^2 \right] \\
& = \nu P (P (Q^{r-2} f \otimes Q^{r-2} f) \otimes P (Q^{r-2} f \otimes Q^{r-2} f)).
\end{aligned}$$

Gathering together all these terms each multiplied by its probability, we obtain an explicit expression for  $\mathbb{E} \left[ \overline{M}_{\mathbb{G}_r}(f)^4 \right]$ .

## Part 2. Rate.

We are now going to give some rate for the different terms that appear in the expression of  $\mathbb{E} \left[ \overline{M}_{\mathbb{G}_r}(f)^4 \right]$ . In the sequel  $c$  denotes a positive constant which depends on  $f$ ; it may vary from one line to another.  $c_1$  and  $c_2$  are positives constants which depend on  $\alpha$ . Then we have

$$(a) \mathbb{E} \left[ f(X_{I_r}) f(X_{J_r}) f(X_{K_r}) f(X_{L_r}) / E_4^r \right] \times \mathbb{P}(E_4^r) \leq c \frac{1}{2^{3r}},$$

where from (ii), (v) and (vi)  $c$  is such that  $\nu Q^r f^4 < c$ .

$$(b) \mathbb{E} \left[ f(X_{I_r}) f(X_{J_r}) f(X_{K_r}) f(X_{L_r}) / E_3^r \right] \times \mathbb{P}(E_3^r) \leq c \frac{1}{2\alpha} \alpha^r \left( \frac{1}{4} \right)^r \sum_{p=0}^{r-1} \left( \frac{1}{2\alpha} \right)^p,$$

where from (ii), (iii), (v) and (vi),  $c$  is such for all  $p, q \in \mathbb{N}$

$$\max (\nu Q^p P (Q^q f^3 \otimes g), \nu Q^p P (g \otimes Q^q f^3)) < c,$$

and from hypothesis **(H1)**,  $g$  is such that for all  $p \in \{1, \dots, r-1\}$

$$Q^{r-p-1} f \leq \alpha^{r-p-1} g. \quad (2.3)$$

- If  $\alpha \neq \frac{1}{2}$ , then  $\mathbb{E} \left[ f(X_{I_r}) f(X_{J_r}) f(X_{K_r}) f(X_{L_r}) / E_3^r \right] \times \mathbb{P}(E_3^r) \leq \frac{4}{|2\alpha-1|} c \left( \left( \frac{\alpha}{4} \right)^r + \left( \frac{1}{2^3} \right)^r \right)$ .
- If  $\alpha = \frac{1}{2}$  then,  $\mathbb{E} \left[ f(X_{I_r}) f(X_{J_r}) f(X_{K_r}) f(X_{L_r}) / E_3^r \right] \times \mathbb{P}(E_3^r) \leq cr \left( \frac{1}{2^3} \right)^r$ .

$$(c) \mathbb{E} \left[ f(X_{I_r}) f(X_{J_r}) f(X_{K_r}) f(X_{L_r}) / E_0^2 \right] \times \mathbb{P}(E_0^2) \leq \frac{3}{32} \frac{1}{\alpha^8} \alpha^{4r}.$$

$$\begin{aligned}
 \text{(d)} \quad \sum_{p=2}^r \mathbb{E} \left[ f(X_{I_r})f(X_{J_r})f(X_{K_r})f(X_{L_r})/E_0^{p+1}, E_2^p \right] &\times \mathbb{P}(E_2^p)\mathbb{P}(E_0^{p+1}/E_2^p) \\
 &\leq 6c \left(\frac{1}{4}\right)^r + \frac{3c}{2\alpha^4} \alpha^{4r} \sum_{p=2}^{r-1} \left(\frac{1}{4\alpha^4}\right)^p.
 \end{aligned}$$

where from (ii), (iii), (v) and (vi)  $c$  is such that for all  $p \in \{2, \dots, r-1\}$ ,  $q \in \{0, \dots, r-1\}$ ,  $l \in \{0, \dots, p-1\}$

$$\max \left( \nu Q^q P \left( Q^{r-p-1} f^2 \otimes Q^{r-p-1} f^2 \right), \nu Q^l P \left( Q^{p-l-1} P(g \otimes g) \otimes Q^{p-l-1} P(g \otimes g) \right) \right) < c,$$

and  $g$  is defined as before (2.3).

- If  $\alpha^2 \neq \frac{1}{2}$  then

$$\begin{aligned}
 \sum_{p=2}^r \mathbb{E} \left[ f(X_{I_r})f(X_{J_r})f(X_{K_r})f(X_{L_r})/E_0^{p+1}, E_2^p \right] &\times \mathbb{P}(E_2^p)\mathbb{P}(E_0^{p+1}/E_2^p) \\
 &\leq c \left( 6 \left(\frac{1}{4}\right)^r + \frac{3}{8\alpha^8|4\alpha^4-1|} \alpha^{4r} + \frac{6}{|4\alpha^4-1|} \left(\frac{1}{4}\right)^r \right).
 \end{aligned}$$

- If  $\alpha^2 = \frac{1}{2}$  then

$$\begin{aligned}
 \sum_{p=2}^r \mathbb{E} \left[ f(X_{I_r})f(X_{J_r})f(X_{K_r})f(X_{L_r})/E_0^{p+1}, E_2^p \right] &\times \mathbb{P}(E_2^p)\mathbb{P}(E_0^{p+1}/E_2^p) \\
 &\leq 6c(r-1) \left(\frac{1}{4}\right)^r.
 \end{aligned}$$

(e)

- If  $\alpha = \frac{1}{2}$  then

$$\mathbb{E} \left[ f(X_{I_r})f(X_{J_r})f(X_{K_r})f(X_{L_r})/E_1^r \right] \times \mathbb{P}(E_1^r) \leq 12c \left(\frac{1}{4}\right)^r.$$

- If  $\alpha \neq \frac{1}{2}$ 
  - if  $\alpha^2 = \frac{1}{2}$  then

$$\mathbb{E} \left[ f(X_{I_r})f(X_{J_r})f(X_{K_r})f(X_{L_r})/E_1^r \right] \times \mathbb{P}(E_1^r) \leq \frac{3}{4}c_2(r-1) \left(\frac{1}{4}\right)^r;$$

- if  $\alpha^2 \neq \frac{1}{2}$  then

$$\begin{aligned}
 \mathbb{E} \left[ f(X_{I_r})f(X_{J_r})f(X_{K_r})f(X_{L_r})/E_1^r \right] \times \mathbb{P}(E_1^r) &\leq \frac{3c_2c}{2\alpha^4|2\alpha^2-1|} \left(\frac{\alpha^2}{2}\right)^r + \frac{9c}{8\alpha^4} \left(\frac{\alpha^2}{2}\right)^r \\
 &\quad + \frac{6c_2c}{|2\alpha^2-1|} \left(\frac{1}{4}\right)^r,
 \end{aligned}$$

where from (ii), (iii), (v) and (vi)  $c$  is such that for all  $l \in \{2, \dots, r-1\}$ ,  $q \in \{0, \dots, l-1\}$

$$\max \left( \nu Q^p P \left( Q^{l-q-1} P(g \otimes g) \otimes Q^{r-q-1} f^2 \right), \nu Q^q P \left( Q^{l-q-1} P(g \otimes Q^{r-l-1} f^2) \otimes g \right) \right) < c,$$

and  $g$  is defined as before (2.3).

(f)

- If  $\alpha = \frac{1}{2}$  then

$$\sum_{p=2}^{r-1} \mathbb{E} \left[ f(X_{L_r}) f(X_{J_r}) f(X_{K_r}) f(X_{L_r}) / E_0^{p+1}, E_1^p \right] \times \mathbb{P}(E_1^p) \mathbb{P}(E_0^{p+1} / E_1^p) \leq 12c \left( \frac{1}{4} \right)^r.$$

- If  $\alpha \neq \frac{1}{2}$   
– if  $\alpha^2 = \frac{1}{2}$  then

$$\sum_{p=2}^{r-1} \mathbb{E} \left[ f(X_{L_r}) f(X_{J_r}) f(X_{K_r}) f(X_{L_r}) / E_0^{p+1}, E_1^p \right] \times \mathbb{P}(E_1^p) \mathbb{P}(E_0^{p+1} / E_1^p) \leq 3c_1 c r^2 \left( \frac{1}{4} \right)^r;$$

- if  $\alpha^2 \neq \frac{1}{2}$  then

$$\begin{aligned} & \sum_{p=2}^{r-1} \mathbb{E} \left[ f(X_{L_r}) f(X_{J_r}) f(X_{K_r}) f(X_{L_r}) / E_0^{p+1}, E_1^p \right] \times \mathbb{P}(E_1^p) \mathbb{P}(E_0^{p+1} / E_1^p) \\ & \leq \left( \frac{3c_1 c}{8\alpha^8(2\alpha^2-1)^2} + \frac{3c_1 c}{4\alpha^6(2\alpha^2-1)^2(2\alpha^2+1)} + \frac{3c(2\alpha+1)}{32\alpha^8|2\alpha^2-1|} \right) \alpha^{4r} \\ & \quad + \left( \frac{3c_1 c}{2\alpha^2(2\alpha^2-1)^2} + \frac{c}{8\alpha^4|2\alpha^2-1|} \right) \left( \frac{\alpha^2}{2} \right)^r + \frac{12\alpha^2 c_1 c}{(2\alpha^2-1)^2(2\alpha^2+1)} \left( \frac{1}{4} \right)^r. \end{aligned}$$

Where  $c$  is define in the same way as before.

Now the results of the theorem 2.1 follow from (a)-(f) of **Part 2**.  $\square$

We give an extension of theorem 2.1 to the two empirical averages  $\overline{M}_{\mathbb{T}_r}(f)$  and  $\overline{M}_n^\pi(f)$ .

**Corollary 2.2.** *Let  $F$  satisfy (i)-(vi). Let  $f \in F$  such that  $(\mu, f) = 0$ . We assume that hypothesis **(H1)** is fulfilled. Then for all  $r \in \mathbb{N}$  and  $n \in \mathbb{N}$*

$$\mathbb{E} \left[ \overline{M}_{\mathbb{T}_r}(f)^4 \right] \leq \begin{cases} c \left( \frac{1}{4} \right)^{r+1} & \text{if } \alpha^2 < \frac{1}{2} \\ c r^2 \left( \frac{1}{4} \right)^{r+1} & \text{if } \alpha^2 = \frac{1}{2} \\ c \alpha^{4(r+1)} & \text{if } \alpha^2 > \frac{1}{2} \end{cases} \quad (2.4)$$

and

$$\mathbb{E} \left[ \overline{M}_n^\pi(f)^4 \right] \leq \begin{cases} c \left( \frac{1}{4} \right)^{r_n+1} & \text{if } \alpha^2 < \frac{1}{2} \\ c r_n^2 \left( \frac{1}{4} \right)^{r_n+1} & \text{if } \alpha^2 = \frac{1}{2} \\ c \alpha^{4(r_n+1)} & \text{if } \alpha^2 > \frac{1}{2} \end{cases} \quad (2.5)$$

where the positive constant  $c$  depends on  $\alpha$  and  $f$  and may differ line by line.

*Proof.* The proof follows the same steps as in the proof of **Part 2** and **Part 3** of theorem 2.10, and uses the results of the proof of theorem 2.4 to get the control of the 4th order moment in incomplete generation. See section 2.2 and 2.5 for more details.  $\square$

**Remark 2.3.** If  $f \in \mathcal{B}(S^3)$  is such that  $Pf^2$  and  $Pf^4$  exist and belong to  $F$ , with  $Pf = 0$ , then we have for some positive constant  $c$

$$\mathbb{E} \left[ \overline{M}_{\mathbb{G}_r}(f)^4 \right] \leq \frac{c}{|\mathbb{G}_r|^2}. \quad (2.6)$$

Indeed, we have

$$\begin{aligned} \mathbb{E} \left[ M_{\mathbb{G}_r}(f)^4 \right] &= \mathbb{E} \left[ M_{\mathbb{G}_r}(f^4) \right] + 6\mathbb{E} \left[ \sum_{i \neq j \in \mathbb{G}_r} f^2(\Delta_i) f^2(\Delta_j) \right] + 4\mathbb{E} \left[ \sum_{i \neq j \in \mathbb{G}_r} f^3(\Delta_i) f(\Delta_j) \right] \\ &\quad + 12\mathbb{E} \left[ \sum_{i \neq j \neq k \in \mathbb{G}_r} f^2(\Delta_i) f(\Delta_j) f(\Delta_k) \right] + 24\mathbb{E} \left[ \sum_{i \neq j \neq k \neq l \in \mathbb{G}_r} f(\Delta_i) f(\Delta_j) f(\Delta_k) f(\Delta_l) \right] \\ &= \mathbb{E} \left[ \sum_{i \in \mathbb{G}_r} Pf^4(X_i) \right] + 6\mathbb{E} \left[ \sum_{i \neq j \in \mathbb{G}_r} Pf^2(X_i) Pf^2(X_j) \right], \end{aligned}$$

where the last equality was obtained after conditioning by  $\mathcal{F}_{r-1}$  and using the fact that  $Pf = 0$ . Now, dividing by  $|\mathbb{G}_r|^4$  leads us to

$$\begin{aligned} \mathbb{E} \left[ \overline{M}_{\mathbb{G}_r}(f)^4 \right] &= \frac{6}{|\mathbb{G}_r|^2} \mathbb{E} \left[ \frac{1}{|\mathbb{G}_r|^2} \sum_{i \neq j \in \mathbb{G}_r} Pf^2(X_i) Pf^2(X_j) \right] + \frac{1}{|\mathbb{G}_r|^3} \mathbb{E} \left[ \frac{1}{|\mathbb{G}_r|} \sum_{i \in \mathbb{G}_r} Pf^4(X_i) \right] \\ &\leq \frac{12(\mu, Pf^2)^2}{|\mathbb{G}_r|^2} + \frac{12}{|\mathbb{G}_r|^2} \mathbb{E} \left[ \left( \frac{1}{|\mathbb{G}_r|} \sum_{i \in \mathbb{G}_r} Pf^2(X_i) - (\mu, Pf^2) \right)^2 \right] + \frac{(\mu, Pf^4)}{|\mathbb{G}_r|^3} \end{aligned}$$

and (2.6) then follows using **Part 1** of the proof of theorem 2.10.

## 2.2. Strong law of large numbers on incomplete subtree.

We now turn to prove strong law of large numbers for  $\overline{M}_n^\pi(f)$ , completing the work of Guyon [12], where the LLN was proved only for the two averages  $\overline{M}_{\mathbb{T}_r}(f)$  and  $\overline{M}_{\mathbb{G}_r}(f)$ .

**Theorem 2.4.** *Let  $F$  satisfy (i)-(vi). Let  $f \in F$  such that  $(\mu, f) = 0$ . We assume that hypothesis **(H1)** is fulfilled with  $\alpha \in \left(0, \frac{\sqrt{8}}{2}\right)$ . Then  $\overline{M}_n^\pi(f)$  almost surely converges to 0 as  $n$  goes to  $\infty$ .*

*Proof.* From the decomposition

$$\overline{M}_n^\pi(f) = \sum_{q=0}^{r_n-1} \frac{2^q}{n} \overline{M}_{\mathbb{G}_q}(f) + \frac{1}{n} \sum_{i=2^{r_n}}^n f(X_{\pi(i)}),$$

it is enough to check that

$$\sum_{n=1}^{\infty} \mathbb{E} \left[ \left( \frac{1}{n} \sum_{i=2^{r_n}}^n f(X_{\pi(i)}) \right)^4 \right] < \infty.$$

Indeed, since  $\overline{M}_{\mathbb{G}_q}(f)$  almost surely converges to 0 (corollary 15 in [12]), we deduce that the first term of the right hand side of the previous decomposition almost surely converges to 0 (lemma 13 in [12]).

We have

$$\begin{aligned} \mathbb{E} \left[ \left( \frac{1}{n} \sum_{i=2^{r_n}}^n f(X_{\pi(i)}) \right)^4 \right] &= c \frac{1}{n^4} \mathbb{E} \left[ \sum_{i=2^{r_n}}^n f^4(X_{\pi(i)}) \right] \\ &\quad + c \frac{1}{n^4} \mathbb{E} \left[ \sum_{i,j=2^{r_n}; i \neq j}^n f^2(X_{\pi(i)}) f^2(X_{\pi(j)}) \right] \\ &\quad + c \frac{1}{n^4} \mathbb{E} \left[ \sum_{i,j=2^{r_n}; i \neq j}^n f^3(X_{\pi(i)}) f(X_{\pi(j)}) \right] \\ &\quad + c \frac{1}{n^4} \mathbb{E} \left[ \sum_{i,j,k=2^{r_n}; i \neq i \neq k}^n f^2(X_{\pi(i)}) f(X_{\pi(j)}) f(X_{\pi(k)}) \right] \\ &\quad + c \frac{1}{n^4} \mathbb{E} \left[ \sum_{i,j,k,l=2^{r_n}; i \neq j \neq k \neq l}^n f(X_{\pi(i)}) f(X_{\pi(j)}) f(X_{\pi(k)}) f(X_{\pi(l)}) \right] \end{aligned}$$

where the positive constant  $c$  may differ line by line.

Recall the following: for  $i, j, k$  and  $l \in \{2^{r_n}, \dots, n\}$

- If  $i \neq j$ , then  $r_n \geq 1$ . Independently on  $(X, \pi)$ , draw two independent indices  $I_{r_n}$  and  $J_{r_n}$  uniformly from  $\mathbb{G}_{r_n}$ . Then the law of  $(\pi(i), \pi(j))$  is the conditional law of  $(I_{r_n}, J_{r_n})$  given  $\{I_{r_n} \neq J_{r_n}\}$ .
- If  $i \neq j \neq k$ , then  $r_n \geq 2$ . Independently on  $(X, \pi)$ , draw three independent indices  $I_{r_n}, J_{r_n}$  and  $K_{r_n}$  uniformly from  $\mathbb{G}_{r_n}$ . Then the law of  $(\pi(i), \pi(j), \pi(k))$  is the conditional law of  $(I_{r_n}, J_{r_n}, K_{r_n})$  given  $\{I_{r_n} \neq J_{r_n} \neq K_{r_n}\}$ .
- If  $i \neq j \neq k \neq l$ , then  $r_n \geq 2$ . Independently on  $(X, \pi)$ , draw four independent indices  $I_{r_n}, J_{r_n}, K_{r_n}$  and  $L_{r_n}$  uniformly from  $\mathbb{G}_{r_n}$ . Then the law of  $(\pi(i), \pi(j), \pi(k), \pi(l))$  is the conditional law of  $(I_{r_n}, J_{r_n}, K_{r_n}, L_{r_n})$  given  $\{I_{r_n} \neq J_{r_n} \neq K_{r_n} \neq L_{r_n}\}$ .

Then we check that

$$\frac{1}{n^4} \mathbb{E} \left[ \sum_{i=2^{r_n}}^n f^4(X_{\pi(i)}) \right] = O\left(\frac{1}{n^3}\right),$$

$$\frac{1}{n^4} \mathbb{E} \left[ \sum_{i,j=2^{r_n}; i \neq j}^n f^2(X_{\pi(i)}) f^2(X_{\pi(j)}) \right] = O\left(\frac{1}{n^2}\right),$$

and

$$\frac{1}{n^4} \mathbb{E} \left[ \sum_{i,j=2^{r_n}; i \neq j}^n f^3(X_{\pi(i)}) f(X_{\pi(j)}) \right] = o\left(\frac{1}{n^2}\right).$$

Let us deal with the remaining term

$$\frac{1}{n^4} \mathbb{E} \left[ \sum_{i,j,k=2^{r_n}; i \neq j \neq k}^n f^2(X_{\pi(i)}) f(X_{\pi(j)}) f(X_{\pi(k)}) \right] = \frac{(n - 2^{r_n} - 1)(n - 2^{r_n})(n - 2^{r_n} + 1)}{\mathbb{P}(I_{r_n} \neq J_{r_n} \neq K_{r_n}) \times n^4} \\ \times \mathbb{E} \left[ f^2(X_{I_{r_n}}) f(X_{J_{r_n}}) f(X_{K_{r_n}}) \mathbf{1}_{\{I_{r_n} \neq J_{r_n} \neq K_{r_n}\}} \right].$$

Then, we get an explicit expression for the last expectation similar to that obtained in part (d) of the calculus of  $\mathbb{E}[\overline{M}_{G_r}(f)^4]$  with a slight modification of the functions. Calculating the rate of this expression, we obtain

$$\sum_{n=4}^{\infty} \frac{1}{n^4} \mathbb{E} \left[ \sum_{i,j,k=2^{r_n}; i \neq j \neq k}^n f^2(X_{\pi(i)}) f(X_{\pi(j)}) f(X_{\pi(k)}) \right] \leq c \sum_{n=1}^{\infty} \frac{1}{n} \alpha^{2r_n} \\ + c \sum_{n=1}^{\infty} \sum_{p=2}^{r_n-1} \sum_{l=0}^{p-1} \frac{1}{n} \frac{1}{2^p} \frac{1}{2^{l+1}} \alpha^{2r_n-2p} \\ + c \sum_{n=1}^{\infty} \sum_{p=2}^{r_n-1} \sum_{l=0}^{p-1} \frac{1}{n} \frac{1}{2^p} \frac{1}{2^{l+1}} \alpha^{2r_n-p-l},$$

where the positive constant  $c$  differs term by term. Now it is no hard to see that the right hand side is finite.

Finally, to check that the series of general term

$$\frac{1}{n^4} \mathbb{E} \left[ \sum_{i,j,k,l=2^{r_n}; i \neq j \neq k \neq l}^n f(X_{\pi(i)}) f(X_{\pi(j)}) f(X_{\pi(k)}) f(X_{\pi(l)}) \right]$$

is finite, it is enough, according to the calculation of rates we have done in **Part 2** of the proof of theorem 2.1, to check that  $\sum_{n=1}^{\infty} \alpha^{4r_n} < \infty$ , which is the case if  $\alpha \in \left(0, \frac{\sqrt[4]{8}}{2}\right)$  and this ends the proof of theorem 2.4.  $\square$

**Remark 2.5.** Note that this theorem can be improved, but the price to pay is enormous computations related to the calculation of higher moments, and stronger conditions on the integrability of  $f$ . If  $f$  is bounded, this result is true for every  $\alpha \in (0, 1)$  as we will see in section 3.

### 2.3. Law of the iterated logarithm (LIL).

Now using the LIL for martingales (see Theorem 5.2 of Stout in the Appendix), we are going to prove a LIL for the BMC. This will be done when  $f$  depends on the mother-daughters triangle  $(\Delta_i)$ . We use the notations  $M_n^\pi(f) = \sum_{i=1}^n f(\Delta_{\pi(i)})$  and  $M_{\mathbb{T}_r}(f) = \sum_{i \in \mathbb{T}_r} f(\Delta_i)$ .

**Theorem 2.6.** *Let  $F$  satisfy (i)-(vi). Let  $f \in \mathcal{B}(\mathcal{S}^3)$  such that  $Pf = 0$ ,  $Pf^2$  and  $Pf^4$  exist and belong to  $F$ . We assume that hypothesis **(H1)** is fulfilled. Then*



$$\limsup_{n \rightarrow \infty} \frac{M_n^\pi(f)}{\sqrt{2\langle M^\pi(f) \rangle_n \log \log \langle M^\pi(f) \rangle_n}} = 1 \quad a.s.$$

And particularly

$$\limsup_{r \rightarrow \infty} \frac{M_{\mathbb{T}_r}(f)}{\sqrt{2|\mathbb{T}_r| \log \log |\mathbb{T}_r|}} = \sqrt{(\mu, Pf^2)} \quad a.s..$$

*Proof.* We will check the hypothesis of Stout Theorem's 5.2. Let  $f \in \mathcal{B}(\mathcal{S}^3)$ . We introduce the following filtration  $(\mathcal{H}_n)_{n \geq 0}$  defined by  $\mathcal{H}_0 = \sigma(X_1)$  and  $\mathcal{H}_n = \sigma(\Delta_{\pi(i)}, 1 \leq i \leq n)$ . Let  $(M_n^\pi(f))_{n \geq 0}$  defined by  $M_0^\pi(f) = 0$  and  $M_n^\pi(f) = \sum_{i=1}^n f(\Delta_{\pi(i)})$ . Then since  $Pf = 0$ ,  $(M_n^\pi(f))$  is a  $\mathcal{H}_n$ -martingale with  $\mathbb{E}[M_1^\pi(f)] = 0$ . We have the following decomposition

$$\frac{\langle M^\pi(f) \rangle_n}{n} = \overline{M}_n^\pi(Pf^2) = \sum_{q=0}^{r_n-1} \frac{2^q}{n} \overline{M}_{\mathbb{G}_q}(Pf^2) + \frac{1}{n} \sum_{i=2^{r_n}}^n Pf^2(X_{\pi(i)}).$$

On the one hand, we have

$$\frac{\langle M^\pi(f) \rangle_n}{n} \geq \sum_{q=0}^{r_n-1} \frac{2^q}{n} \overline{M}_{\mathbb{G}_q}(Pf^2).$$

Now, since

$$\sum_{q=0}^{r_n-1} \frac{2^q}{n} \overline{M}_{\mathbb{G}_q}(Pf^2) \xrightarrow[n \rightarrow \infty]{} c(\mu, Pf^2) \quad a.s.,$$

where  $c = \lim_{n \rightarrow \infty} \frac{2^{r_n-1}}{n}$ ; we deduce that  $n = O(\langle M^\pi(f) \rangle_n)$ , so  $\langle M^\pi(f) \rangle_n \xrightarrow[n \rightarrow \infty]{} \infty$  a.s..

On the other hand, we have

$$\frac{1}{n} \sum_{i=2^{r_n}}^n Pf^2(X_{\pi(i)}) \leq \overline{M}_{\mathbb{G}_{r_n}}(Pf^2) \xrightarrow[n \rightarrow \infty]{} (\mu, Pf^2),$$

we deduce from the previous decomposition also that  $\langle M \rangle_n = O(n)$  a.s..

Now taking  $K_n = \frac{\sqrt{2}}{\sqrt{\log \log(n)}}$  in the theorem 5.2, we have

$$\begin{aligned} R &:= \sum_{n=1}^{\infty} \frac{2 \log \log \langle M^\pi(f) \rangle_n}{K_n^2 \langle M^\pi(f) \rangle_n} \mathbb{E} \left[ f^2(\Delta_{\pi(n)}) \mathbf{1}_{\left\{ f^2(\Delta_{\pi(n)}) > \frac{K_n^2 \langle M^\pi(f) \rangle_n}{2 \log \log \langle M^\pi(f) \rangle_n} \right\}} / \mathcal{H}_{n-1} \right] \\ &\leq \sum_{n=1}^{\infty} \frac{4 (\log \log \langle M^\pi(f) \rangle_n)^2}{K_n^4 (\langle M^\pi(f) \rangle_n)^2} Pf^4(X_{\pi(n)}) \quad a.s., \end{aligned}$$

since  $\langle M \rangle_n = O(n)$  a.s. so that for  $R < \infty$  a.s., it is enough to check that

$$\sum_{n=1}^{\infty} \frac{Pf^4(X_{\pi(n)})}{n^2} < \infty \quad a.s.. \quad (2.7)$$

Now, according to hypothesis **(H1)**, there exists a positive constant  $c$  such that for all  $n \geq 1$ ,  $\mathbb{E}\left[Pf^4(X_{\pi(n)})\right] = \nu Q^{r_n} Pf^4 \leq c$ , and (2.7) follows. Applying theorem (5.2), we have

$$\limsup_{n \rightarrow \infty} \frac{M_n^\pi(f)}{\sqrt{2\langle M^\pi(f) \rangle_n \log \log \langle M^\pi(f) \rangle_n}} = 1 \quad \text{a.s.}$$

Now, for  $n = |\mathbb{T}_r|$  we have the following

$$\frac{M_{\mathbb{T}_r}(f)}{\sqrt{2\langle M(f) \rangle_{\mathbb{T}_r} \log \log \langle M(f) \rangle_{\mathbb{T}_r}}} = \sqrt{\frac{|\mathbb{T}_r| \frac{\langle M(f) \rangle_{\mathbb{T}_r}}{|\mathbb{T}_r|}}{2 \log \log \langle M(f) \rangle_{\mathbb{T}_r}}} \times \frac{M_{\mathbb{T}_r}(f)}{|\mathbb{T}_r| \frac{\langle M(f) \rangle_{\mathbb{T}_r}}{|\mathbb{T}_r|}}$$

and since  $\frac{\langle M(f) \rangle_{\mathbb{T}_r}}{|\mathbb{T}_r|} = \overline{M}_{\mathbb{T}_r}(Pf^2) \xrightarrow{r \rightarrow \infty} (\mu, Pf^2)$  a.s. (see theorem 18 in [12]), we get

$$\limsup_{r \rightarrow \infty} \frac{M_{\mathbb{T}_r}(f)}{\sqrt{2|\mathbb{T}_r| \log \log |\mathbb{T}_r|}} = \sqrt{(\mu, Pf^2)}$$

which ends the proof.  $\square$

**Remark 2.7.** Let us note that using theorem 2.4, we can prove that if hypothesis **(H1)** is fulfilled with  $\alpha \in \left(0, \frac{\sqrt[4]{8}}{2}\right)$  then,

$$\limsup_{n \rightarrow \infty} \frac{M_n^\pi(f)}{\sqrt{2n \log \log n}} = \sqrt{(\mu, Pf^2)} \quad \text{a.s.},$$

and via the computation of  $2k$ -th order moments of  $\overline{M}_{\mathbb{G}_r}(g)$ , with  $k > 2$  and  $g \in \mathcal{B}(S)$ , it is possible to prove the latter for all  $\alpha \in (0, 1)$ . But, as already emphasized, this comes at the price of enormous computations.

#### 2.4. Almost-sure functional central limit theorem (ASFCLT).

We are now going to prove an ASFCLT theorem for the BMC  $(X_n, n \in \mathbb{T})$ . This will be done when  $f$  depends on the mother-daughters triangle by using the ASFCLT for discrete martingale. We refer to theorem 5.3 of Chaabane in the Appendix for the definition of an AFSCLT.

**Theorem 2.8.** *Let  $F$  satisfy (i)-(vi). Let  $f \in \mathcal{B}(S^3)$  such that  $Pf = 0$ ,  $Pf^2$  and  $Pf^4$  exist and belong to  $F$ . We assume that hypothesis **(H1)** is fulfilled with  $\alpha \in \left(0, \frac{\sqrt[4]{8}}{2}\right)$ . Then  $M_n^\pi(f)$  verify an ASFCLT, when  $n$  goes to  $\infty$ .*

*Proof.* We use Theorem 5.3. Let  $(\mathcal{H}_n)_{n \in \mathbb{N}}$  be the filtration defined as in section (2.3). Then  $(M_n^\pi(f))$  is a  $\mathcal{H}_n$  martingale. We have to check the hypotheses of Theorem 5.3. For all  $n \geq 1$ , let  $V_n = s\sqrt{n}$  where  $s^2 = (\mu, Pf^2)$ . Then according to Theorem (2.4)

$$V_n^{-2} \langle M^\pi(f) \rangle_n = V_n^{-2} M_n^\pi(Pf^2) \xrightarrow{n \rightarrow \infty} 1 \quad \text{a.s.}$$

Let  $\varepsilon > 0$ . We have

$$\sum_{n \geq 1} \frac{1}{V_n^2} \mathbb{E} \left[ f^2(\Delta_{\pi(n)}) \mathbf{1}_{\{|f(\Delta_{\pi(n)})| > \varepsilon V_n\}} / \mathcal{H}_{n-1} \right] \leq \frac{1}{\varepsilon^2 s^4} \sum_{n \geq 1} \frac{Pf^4(X_{\pi(n)})}{n^2}.$$

According to hypothesis **(H1)**, there exists a positive constant  $c$  such that for all  $n \geq 1$ ,  $\mathbb{E}\left[Pf^4(X_{\pi(n)})\right] = \nu Q^{r_n} Pf^4 \leq c$ , and therefore,  $\forall \epsilon > 0$

$$\sum_{n \geq 1} \frac{1}{V_n^2} \mathbb{E}\left[f^2(\Delta_{\pi(n)}) \mathbf{1}_{\{|f(\Delta_{\pi(n)})| > \epsilon V_n\}} / \mathcal{H}_{n-1}\right] < \infty \quad \text{a.s.}$$

Finally, we have

$$\sum_{n \geq 1} \frac{1}{V_n^4} \mathbb{E}\left[f^4(\Delta_{\pi(n)}) \mathbf{1}_{\{|f(\Delta_{\pi(n)})| \leq V_n\}} / \mathcal{H}_{n-1}\right] \leq \frac{1}{s^4} \sum_{n \geq 1} \frac{Pf^4(X_{\pi(n)})}{n^2},$$

which as before is a.s. finite and the proof is then complete.  $\square$

**Remark 2.9.** As before, let us note that this result can be extended to the general case  $\alpha \in (0, 1)$ , but at the price of enormous computation related to the computation of  $2k$ -order moments,  $k > 2$ , for  $\overline{M}_{\mathbb{G}_r}(g)$ ,  $g \in \mathcal{B}(S)$ .

## 2.5. Some probability inequalities for BMC.

We are now going to give some probability inequalities under (i) – (vi) and **(H1)** for the empirical means (1.2) with  $(\mu, f) = 0$  and (1.3) with  $(\mu, Pf) = 0$ . This will help us in the sequel to obtain a MDP result in a general framework, that is for functional of BMC with unbounded test functions. Note that the use of 2nd order moment is sufficient for this work, since the higher moments do not improve the speed of the moderate deviation and the speed of superexponential convergence (see the Appendix for the definition). Recall that the main disadvantage here is that the range of speed for the MDP is very restricted. However, we still work under geometric ergodicity assumption and general test function, which will not be the case when we would want to extend the MDP (see section 3).

**Theorem 2.10.** *Let  $F$  satisfy conditions (i)-(vi). We assume that **(H1)** is fulfilled. Let  $f \in \mathcal{B}(S)$  such that  $(\mu, f) = 0$ . Then we have for all  $\delta > 0$  and all  $r \in \mathbb{N}$  and all  $n \in \mathbb{N}$*

$$\mathbb{P}\left(|\overline{M}_{\mathbb{G}_r}(f)| > \delta\right) \leq \begin{cases} \frac{c}{\delta^2} \left(\frac{1}{2}\right)^r & \text{if } \alpha^2 < \frac{1}{2}; \\ \frac{c}{\delta^2} r \left(\frac{1}{2}\right)^r & \text{if } \alpha^2 = \frac{1}{2}; \\ \frac{c}{\delta^2} \alpha^{2r} & \text{if } \alpha^2 > \frac{1}{2}; \end{cases} \quad (2.8)$$

$$\mathbb{P}\left(|\overline{M}_n^\pi(f)| > \delta\right) \leq \begin{cases} \frac{c}{\delta^2} \left(\frac{1}{2}\right)^{r_n+1} & \text{if } \alpha^2 < \frac{1}{2}; \\ \frac{c}{\delta^2} r_n \left(\frac{1}{2}\right)^{r_n+1} & \text{if } \alpha^2 = \frac{1}{2}; \\ \frac{c}{\delta^2} \alpha^{2(r_n+1)} & \text{if } \alpha^2 > \frac{1}{2}; \end{cases} \quad (2.9)$$

and

$$\mathbb{P}\left(\left|\overline{M}_{\mathbb{T}_r}(f)\right| > \delta\right) \leq \begin{cases} \frac{c}{\delta^2} \left(\frac{1}{2}\right)^{r+1} & \text{if } \alpha^2 < \frac{1}{2}; \\ \frac{c}{\delta^2} r \left(\frac{1}{2}\right)^{r+1} & \text{if } \alpha^2 = \frac{1}{2}; \\ \frac{c}{\delta^2} \alpha^{2(r+1)} & \text{if } \alpha^2 > \frac{1}{2}; \end{cases} \quad (2.10)$$

where the positive constant  $c$  depends on  $f$  and  $\alpha$  and may differ term by term.

*Proof.* Let  $f \in \mathcal{B}(S)$  such that  $(\mu, f) = 0$ . We shall study the three empirical averages  $\overline{M}_{\mathbb{G}_r}(f)$ ,  $\overline{M}_n^\pi(f)$  and  $\overline{M}_{\mathbb{T}_r}(f)$  successively.

**Part 1.** Let us first deal with  $\overline{M}_{\mathbb{G}_r}(f)$ . By Markov inequality, we get, for all  $\delta > 0$

$$\mathbb{P}\left(\left|\overline{M}_{\mathbb{G}_r}(f)\right| > \delta\right) = \mathbb{P}\left(\left|\overline{M}_{\mathbb{G}_r}(f)\right|^2 > \delta^2\right) \leq \frac{1}{\delta^2} \mathbb{E}\left[\left(\overline{M}_{\mathbb{G}_r}(f)\right)^2\right].$$

By Guyon (see [12]), we have

$$\mathbb{E}\left[\left(\overline{M}_{\mathbb{G}_r}(f)\right)^2\right] = \sum_{p=0}^r 2^{-p-1} \nu Q^p P(Q^{r-p-1} f \otimes Q^{r-p-1} f).$$

Hypothesis **(H1)** implies that there exists  $g \in F$  and  $\alpha \in (0, 1)$  such that for all  $p \in \{0, 1, \dots, r\}$

$$\nu Q^p P(Q^{r-p-1} f \otimes Q^{r-p-1} f) \leq \alpha^{2(r-p-1)} \nu Q^p P(g \otimes g).$$

Next, hypothesis (iii), (v) and (vi) imply that there is a positive constant  $c$  such that for all  $p \in \{0, 1, \dots, r\}$

$$\alpha^{2(r-p-1)} \nu Q^p P(g \otimes g) \leq c \alpha^{2(r-p-1)}.$$

This leads us to

$$\mathbb{E}\left[\left(\overline{M}_{\mathbb{G}_r}(f)\right)^2\right] \leq c \sum_{p=0}^r 2^{-p-1} \alpha^{2(r-p-1)} = c \left(\frac{1}{2}\right)^r + c \frac{\alpha^{2r} - \left(\frac{1}{2}\right)^r}{2\alpha^2 - 1}$$

and therefore (2.8) follows.

**Part 2.** Let us now consider  $\overline{M}_n^\pi(f)$ . By the Markov inequality, we get, for all  $\delta > 0$

$$\begin{aligned} \mathbb{P}\left(\left|\overline{M}_n^\pi(f)\right| > \delta\right) &= \mathbb{P}\left(\left|\overline{M}_n^\pi(f)\right|^2 > \delta^2\right) \leq \frac{1}{\delta^2} \mathbb{E}\left[\left(\overline{M}_n^\pi(f)\right)^2\right] \\ &\leq \frac{2}{\delta^2} \mathbb{E}\left[\left(\sum_{q=0}^{r_n-1} \frac{2^q}{n} \overline{M}_{\mathbb{G}_q}(f)\right)^2\right] + \frac{2}{\delta^2} \mathbb{E}\left[\left(\frac{1}{n} \sum_{i=2^{r_n}}^n f(X_{\pi(i)})\right)^2\right]. \end{aligned}$$

In the last inequality, we have used the decomposition

$$\overline{M}_n^\pi(f) = \sum_{q=0}^{r_n-1} \frac{2^q}{n} \overline{M}_{\mathbb{G}_q}(f) + \frac{1}{n} \sum_{i=2^{r_n}}^n f(X_{\pi(i)}).$$

In what follows, the constant  $c$  may be slightly different from that of **Part 1** and may differ term by term.

$$\mathbb{E} \left[ \left( \sum_{q=0}^{r_n-1} \frac{2^q}{n} \overline{M}_{\mathbb{G}_q}(f) \right)^2 \right] = \left\| \sum_{q=0}^{r_n-1} \frac{2^q}{n} \overline{M}_{\mathbb{G}_q}(f) \right\|_2^2 \leq \left( \sum_{q=0}^{r_n-1} \frac{2^q}{n} \|\overline{M}_{\mathbb{G}_q}(f)\|_2 \right)^2.$$

- If  $\alpha^2 < \frac{1}{2}$  then

$$\sum_{q=0}^{r_n-1} \frac{2^q}{n} \|\overline{M}_{\mathbb{G}_q}(f)\|_2 \leq \frac{c}{n} \sum_{q=0}^{r_n-1} (\sqrt{2})^q \leq c \frac{2^{\frac{r_n}{2}}}{n},$$

so that

$$\mathbb{E} \left[ \left( \sum_{q=0}^{r_n-1} \frac{2^q}{n} \overline{M}_{\mathbb{G}_q}(f) \right)^2 \right] \leq c \frac{2^{r_n}}{n^2} \leq c \left( \frac{1}{2} \right)^{r_n+1}.$$

- If  $\alpha^2 > \frac{1}{2}$  then

$$\sum_{q=0}^{r_n-1} \frac{2^q}{n} \|\overline{M}_{\mathbb{G}_q}(f)\|_2 \leq \frac{c}{n} \sum_{q=0}^{r_n-1} (2\alpha)^q \leq c\alpha^{r_n},$$

so that

$$\mathbb{E} \left[ \left( \sum_{q=0}^{r_n-1} \frac{2^q}{n} \overline{M}_{\mathbb{G}_q}(f) \right)^2 \right] \leq c\alpha^{2(r_n+1)}.$$

- If  $\alpha^2 = \frac{1}{2}$  then

$$\sum_{q=0}^{r_n-1} \frac{2^q}{n} \|\overline{M}_{\mathbb{G}_q}(f)\|_2 \leq \frac{c}{n} \sum_{q=0}^{r_n-1} q\sqrt{2}^q \leq c \frac{(r_n-1)\sqrt{2}^{r_n}}{n},$$

so that

$$\mathbb{E} \left[ \left( \sum_{q=0}^{r_n-1} \frac{2^q}{n} \overline{M}_{\mathbb{G}_q}(f) \right)^2 \right] \leq c \frac{r_n^2}{2^{r_n+1}}.$$

Now,

$$\begin{aligned}
 \mathbb{E} \left[ \left( \frac{1}{n} \sum_{i=2^{r_n}}^n f(X_{\pi(i)}) \right)^2 \right] &= \left\| \frac{1}{n} \sum_{i=2^{r_n}}^n f(X_{\pi(i)}) \right\|_2^2 \\
 &\leq \frac{n-2^{r_n}+1}{n^2} \nu Q^{r_n} f^2 \\
 &\quad + \frac{(n-2^{r_n})(n-2^{r_n}+1)}{n^2(2^{r_n}-1)} \sum_{p=0}^{r_n-1} 2^{-p-1} \nu Q^p P(Q^{r_n-p-1} f \otimes Q^{r_n-p-1} f) \\
 &\leq \frac{c}{n} + c \sum_{p=0}^{r_n-1} 2^{-p-1} \alpha^{2r_n-2p-2} \\
 &= \frac{c}{n} + c \frac{\alpha^{2r_n-2-r_n}}{2\alpha^2-1}.
 \end{aligned}$$

- If  $\alpha^2 < \frac{1}{2}$  then

$$\mathbb{E} \left[ \left( \frac{1}{n} \sum_{i=2^{r_n}}^n f(X_{\pi(i)}) \right)^2 \right] \leq c \left( \frac{1}{2} \right)^{r_n+1}.$$

- If  $\alpha^2 > \frac{1}{2}$  then

$$\mathbb{E} \left[ \left( \frac{1}{n} \sum_{i=2^{r_n}}^n f(X_{\pi(i)}) \right)^2 \right] \leq c \alpha^{2(r_n+1)}.$$

- If  $\alpha^2 = \frac{1}{2}$  then

$$\mathbb{E} \left[ \left( \frac{1}{n} \sum_{i=2^{r_n}}^n f(X_{\pi(i)}) \right)^2 \right] \leq c \frac{r_n^2}{2^{r_n+1}}.$$

Inequality (2.9) then follows.

**Part 3.** The case of  $\overline{M}_{\mathbb{T}_r}(f)$  can be deduced from the previous by taking  $n = |\mathbb{T}_r|$ . □

We shall also need an extension of theorem 2.10 to the case when  $f$  does not only depend on an individual  $X_i$ , but on the mother-daughters triangle  $(\Delta_i)$ .

**Theorem 2.11.** *Let  $F$  satisfy conditions (i)-(vi). We assume that **(H1)** is fulfilled. Let  $f \in \mathcal{B}(S^3)$  such that  $Pf$  and  $Pf^2$  exists and belong to  $F$  and  $(\mu, Pf) = 0$ . Then we have the same conclusion as the theorem 2.10 for the three empirical averages given in (1.3):  $\overline{M}_{\mathbb{G}_r}(f)$ ,  $\overline{M}_{\mathbb{T}_r}(f)$  and  $\overline{M}_n^\pi(f)$ .*

*Proof.* Let  $f \in \mathcal{B}(S^3)$  such that  $Pf$  and  $Pf^2$  exist and belong to  $F$  and  $(\mu, Pf) = 0$ . We shall study the three empirical averages  $\overline{M}_{\mathbb{G}_r}(f)$ ,  $\overline{M}_n^\pi(f)$  and  $\overline{M}_{\mathbb{T}_r}(f)$  successively.

**Part 1.** Let us first deal with  $\overline{M}_{\mathbb{G}_r}(f)$ . By the Markov inequality, we get for all  $\delta > 0$

$$\begin{aligned}
\mathbb{P}\left(|\overline{M}_{\mathbb{G}_r}(f)| > \delta\right) &\leq \frac{1}{\delta^2} \mathbb{E}\left[(\overline{M}_{\mathbb{G}_r}(f))^2\right] \\
&= \frac{1}{\delta^2} \mathbb{E}\left[(\overline{M}_{\mathbb{G}_r}(Pf))^2\right] + \frac{1}{\delta^2} \frac{1}{|\mathbb{G}_r|} \mathbb{E}\left[\overline{M}_{\mathbb{G}_r}(Pf^2 - (Pf)^2)\right] \\
&\leq \frac{1}{\delta^2} \mathbb{E}\left[(\overline{M}_{\mathbb{G}_r}(Pf))^2\right] + \frac{c}{\delta^2} \left(\frac{1}{2}\right)^r.
\end{aligned}$$

The last inequality follows from the convergence of the sequence  $\left(\mathbb{E}\left[\overline{M}_{\mathbb{G}_r}(Pf^2 - (Pf)^2)\right]\right)_r$  (see [12]).

Now, using the **Part 1** of the proof of the theorem 2.10 with  $Pf$  instead of  $f$  leads us to a similar inequality (2.8) in theorem 2.11 for  $f \in \mathcal{B}(S^3)$ .

**Part 2.** Let us now treat  $\overline{M}_n^\pi(f)$ . Using the two equalities

$$\overline{M}_n^\pi(f) = \sum_{q=0}^{r_n-1} \frac{|\mathbb{G}_q|}{n} \overline{M}_{\mathbb{G}_q}(f) + \frac{1}{n} \sum_{i=2^{r_n}}^n f(\Delta_{\pi(i)}),$$

$$\mathbb{E}\left[\left(\frac{1}{n} \sum_{i=2^{r_n}}^n f(\Delta_{\pi(i)})\right)^2\right] = \mathbb{E}\left[(\overline{M}_n^\pi(Pf))^2\right] + \frac{1}{n} \mathbb{E}\left[\overline{M}_n^\pi(Pf^2 - (Pf)^2)\right],$$

and the **Part 1** of the proof of the theorem 2.10 with  $Pf$  instead of  $f$  leads us to a similar inequality (2.9) in theorem 2.11 for  $f \in \mathcal{B}$ .

**Part 3.** The case of  $\overline{M}_{\mathbb{T}_r}(f)$  can be deduced from the previous by taking  $n = |\mathbb{T}_r|$ . □

We thus have the following first result on the superexponential convergence in probability, whose definition we present now

**Definition 2.12.** Let  $(E, d)$  a metric space. Let  $(Z_n)$  be a sequence of random variables values in  $E$ ,  $Z$  be a random variable value in  $E$  and  $(v_n)$  be a rate. We say that  $Z_n$  converges  $v_n$ -superexponentially fast in probability to  $Z$  if for all  $\delta > 0$

$$\limsup_{n \rightarrow \infty} \frac{1}{v_n} \log \mathbb{P}(d(Z_n, Z) > \delta) = -\infty.$$

This "exponential convergence" with speed  $v_n$  will be shortened as

$$Z_n \xrightarrow[v_n]{\text{superexp}} Z.$$

We may now set

**Proposition 2.13.** Let  $F$  satisfy conditions (i)-(vi). Let  $f \in \mathcal{B}(S^3)$  such that  $Pf$  and  $Pf^2$  exists and belong to  $F$  and  $(\mu, Pf) = 0$ . We assume that **(H1)** is fulfilled. Let  $(b_n)$  a sequence of positive real numbers satisfying

$$(b_n) \text{ is increasing, } \frac{b_n}{\sqrt{n}} \rightarrow +\infty, \quad \frac{b_n}{\sqrt{n \log n}} \rightarrow 0.$$

Then

$$\overline{M}_n^\pi(f) \xrightarrow[\frac{b_n^2}{n}]{\text{superexp}} 0.$$

*Proof.* The proof is a direct consequence of Theorem 2.11.  $\square$

## 2.6. Moderate deviations for BMC.

Now, using the MDP for martingale (see e.g [9], [18]), we are going to prove a MDP for BMC. We will use Proposition 5.4, in the Appendix.

**Theorem 2.14.** *Let  $f \in \mathcal{B}(S^3)$  such that  $Pf^2$  and  $Pf^4$  exist and belong to  $F$ . Assume that  $Pf = 0$ . Let  $(b_n)$  be a sequence of positive real numbers satisfying*

$$(b_n) \text{ is increasing, } \frac{b_n}{\sqrt{n}} \rightarrow +\infty, \quad \frac{b_n}{\sqrt{n \log n}} \rightarrow 0.$$

If

$$\limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \log \left( n \operatorname{ess\,sup}_{1 \leq k \leq c^{-1}(b_{n+1})} \mathbb{P} \left( |f(\Delta_{\pi(k)})| > b_n / \mathcal{H}_{k-1} \right) \right) = -\infty, \quad (2.11)$$

where  $c^{-1}(b_{n+1}) := \inf \{k \in \mathbb{N} : \frac{k}{b_k} \geq b_{n+1}\}$ ; then  $\left(\frac{1}{b_n} M_n^\pi(f)\right)$  satisfies a MDP in  $\mathbb{R}$  with the speed  $\frac{b_n^2}{n}$  and the rate function  $I(x) = \frac{x^2}{2(\mu, Pf^2)}$ .

*Proof.* Firstly, note that under the hypothesis  $M_n^\pi(f)$  is a  $\mathcal{H}_n$ -martingale, with  $\mathcal{H}_0 = \sigma(X_1)$  and  $\mathcal{H}_n = \sigma(\Delta_{\pi(i)}, 1 \leq i \leq n)$ . From proposition 5.4 in the Appendix, it then suffices to check conditions **(C1)** and **(C3)**.

On the one hand, (2.9) applied to  $Pf^4 - (\mu, Pf^4)$  implies that for all  $\delta > 0$ ,

$$\limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \log \mathbb{P} \left( \frac{1}{n} \sum_{i=1}^n Pf^4(X_{\pi(i)}) > (\mu, Pf^4) + \delta \right) = -\infty$$

and this implies the exponential Lindeberg condition (see for e.g [18]), that is condition **(C3)**.

On the other hand, we have  $\langle M^\pi(f) \rangle_n = M_n^\pi(Pf^2)$  and (2.9) applied to  $Pf^2 - (\mu, Pf^2)$  implies that

$$\overline{M}_n^\pi(Pf^2 - (\mu, Pf^2)) \xrightarrow[\frac{b_n^2}{n}]{\text{superexp}} 0,$$

that is condition **(C1)**.  $\square$

**Remark 2.15.** One of the main difficulty in the application of this Theorem lies in the verification of (2.11). Note however that in the range of speed considered it is sufficient to have some uniform control in  $X_i$  of some moment of  $f(X_i, X_{2i}, X_{2i+1})$  conditionnally on  $X_i$ , which leads to condition of the type  $P|f|^k$  bounded for some  $k \geq 2$ . It is of course the case if  $f$  is bounded.

**Remark 2.16.** In the special case of model (1.1), we have (see section 4), for all  $k$

$$\mathbb{E} \left[ \exp \left( \lambda \frac{b_n}{n} f(\Delta_{\pi(k)}) \right) / \mathcal{H}_{k-1} \right] = \exp \left( \frac{b_n^2}{n} \left( \frac{\lambda Pf^2}{2n} \right) (X_{\pi(k)}) \right).$$



This condition implies that a MDP is satisfied for  $\frac{1}{b_n} M_n^\pi(f)$ . Indeed, if this relation is satisfied, we then have that for  $\lambda \in \mathbb{R}$  the quantity

$$G_n(\lambda) = \frac{\lambda^2}{2} \overline{M}_n^\pi(Pf^2) = \frac{\lambda^2}{2n} \sum_{k=1}^n Pf^2(X_{\pi(k)})$$

is an upper and lower cumulant (see e.g [18]), and we may apply some Gärtner-Ellis type methodology. In addition, due to (2.9) applied to  $Pf^2 - (\mu, Pf^2)$ , we have for  $\lambda \in \mathbb{R}$

$$G_n(\lambda) \xrightarrow[\frac{b_n^2}{n}]{\text{superexp}} \frac{\lambda^2(\mu, Pf^2)}{2},$$

which implies that  $\left(\frac{1}{b_n} M_n^\pi(f)\right)$  satisfies a MDP in  $\mathbb{R}$  with the speed  $\frac{b_n^2}{n}$  and the rate function  $I(x) = \frac{x^2}{2(\mu, Pf^2)}$ .

### 3. EXPONENTIAL PROBABILITY INEQUALITIES FOR BMC AND CONSEQUENCES

We give here stronger deviation inequalities than the one obtained in the previous section, namely exponential deviation inequalities. Of course, it requires strongest assumptions.

#### 3.1. Exponential probability inequalities.

We assume that for all  $f \in \mathcal{B}_b(S)$  such that  $(\mu, f) = 0$  the hypothesis **(H1)** is satisfied uniformly for all  $x \in S$ , that is there exists a positive constant  $c$  such that

$$\textbf{(H2)} \quad |Q^r f(x)| \leq c\alpha^r \quad \text{for some } \alpha \in (0, 1) \text{ and for all } x \in S,$$

then,  $\mathcal{B}_b(S)$  fulfills hypothesis (i)-(vi).

Under this assumption, we will examine exponential probability inequalities for  $\overline{M}_{\mathbb{G}_r}(f)$ ,  $\overline{M}_{\mathbb{T}_r}(f)$  and  $\overline{M}_n^\pi(f)$  when  $f \in \mathcal{B}_b(S)$  with  $(\mu, f) = 0$  (resp.  $f \in \mathcal{B}_b(S^3)$  with  $(\mu, Pf) = 0$ ).

**Theorem 3.1.** *Let  $f \in \mathcal{B}_b(S)$  such that  $(\mu, f) = 0$ . Assume that **(H2)** is satisfied. Then we have for  $r$  and  $n$  large enough and for all  $\delta > 0$*

$$\mathbb{P}\left(\overline{M}_{\mathbb{G}_r}(f) > \delta\right) \leq \begin{cases} \exp(-c'\delta^2|\mathbb{G}_r|) & \text{if } \alpha^2 < \frac{1}{2} \\ \exp\left(-c'\frac{|\mathbb{G}_r|}{r}\right) & \text{if } \alpha^2 = \frac{1}{2}, \\ \exp\left(-c'\delta^2\left(\frac{1}{\alpha}\right)^{2r}\right) & \text{if } \alpha^2 > \frac{1}{2} \end{cases} \quad (3.1)$$

$$\mathbb{P}\left(\overline{M}_{\mathbb{T}_r}(f) > \delta\right) \leq \begin{cases} c'' \exp(-c' \delta^2 |\mathbb{T}_r|) & \text{if } \alpha < \frac{1}{2} \\ \exp(-c' \delta^2 |\mathbb{T}_r|) \exp(2c' \delta(r+1)) & \text{if } \alpha = \frac{1}{2} \\ 2 \exp(-c' \delta^2 |\mathbb{T}_r|) & \text{if } \frac{1}{2} < \alpha < \frac{\sqrt{2}}{2} \\ \exp\left(-c' \delta^2 \frac{|\mathbb{T}_r|}{r+1}\right) & \text{if } \alpha = \frac{\sqrt{2}}{2}, \\ \exp\left(-c' \delta^2 \left(\frac{1}{\alpha^2}\right)^{r+1}\right) & \text{if } \alpha > \frac{\sqrt{2}}{2} \end{cases} \quad (3.2)$$

and

$$\mathbb{P}\left(\overline{M}_n^\pi(f) > \delta\right) \leq \begin{cases} c'' \exp(-c' \delta^2 n) & \text{if } \alpha < \frac{1}{2} \\ \exp(-c' \delta^2 n) \exp(2c' \delta(r_n+1)) & \text{if } \alpha = \frac{1}{2} \\ 2 \exp(-c' \delta^2 n) & \text{if } \frac{1}{2} < \alpha < \frac{\sqrt{2}}{2} \\ \exp\left(-c' \delta^2 \frac{n}{r_n+1}\right) & \text{if } \alpha = \frac{\sqrt{2}}{2}, \\ \exp\left(-c' \delta^2 \left(\frac{1}{\alpha^2}\right)^{r_n+1}\right) & \text{if } \alpha > \frac{\sqrt{2}}{2} \end{cases} \quad (3.3)$$

where  $c'$  and  $c''$  are positive constants which depend on  $\delta$ ,  $\alpha$ ,  $\|f\|_\infty$  and  $c'$ ,  $c''$  differs line by line (see the proofs for the dependence).

*Proof.* Let  $f \in \mathcal{B}_b(S)$  such that  $(\mu, f) = 0$ . We shall study the three empirical averages  $\overline{M}_{\mathbb{G}_r}(f)$ ,  $\overline{M}_n^\pi(f)$  and  $\overline{M}_{\mathbb{T}_r}(f)$  successively.

**Part 1.** Let us first deal with  $\overline{M}_{\mathbb{G}_r}(f)$ . We have for all  $\lambda > 0$  and for all  $\delta > 0$

$$\mathbb{P}\left(\overline{M}_{\mathbb{G}_r}(f) > \delta\right) \leq \exp(-\lambda \delta |\mathbb{G}_r|) \mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathbb{G}_r} f(X_i) \right) \right], \quad (3.4)$$

By subtracting and adding terms, we get

$$\begin{aligned} & \mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathbb{G}_r} f(X_i) \right) \right] \\ &= \mathbb{E} \left[ \mathbb{E} \left[ \prod_{i \in \mathbb{G}_{r-1}} \exp(\lambda(f(X_{2i}) + f(X_{2i+1}) - 2Qf(X_i))) \prod_{i \in \mathbb{G}_{r-1}} \exp(2\lambda Qf(X_i)) / \mathcal{F}_{r-1} \right] \right]. \end{aligned}$$

Now using the fact that conditionally to the  $(r-1)$  first generations the sequence  $\{\Delta_i, i \in \mathbb{G}_{r-1}\}$  is a sequence of independent random variables, we have that

$$\begin{aligned} & \mathbb{E} \left[ \mathbb{E} \left[ \prod_{i \in \mathbb{G}_{r-1}} \exp(\lambda(f(X_{2i}) + f(X_{2i+1}) - 2Qf(X_i))) \prod_{i \in \mathbb{G}_{r-1}} \exp(2\lambda Qf(X_i)) / \mathcal{F}_{r-1} \right] \right] \\ &= \mathbb{E} \left[ \prod_{i \in \mathbb{G}_{r-1}} \exp(2\lambda Qf(X_i)) \prod_{i \in \mathbb{G}_{r-1}} \mathbb{E} [\exp(\lambda(f(X_{2i}) + f(X_{2i+1}) - 2Qf(X_i))) / \mathcal{F}_{r-1}] \right]. \end{aligned}$$

Using the Azuma-Bennet-Hoeffding inequalities [1], [2], [14], we get according to **(H2)**, for all  $i \in \mathbb{G}_{r-1}$

$$\mathbb{E} \left[ \exp \left( \lambda(f(X_{2i}) + f(X_{2i+1}) - 2Qf(X_i)) \right) / \mathcal{F}_{r-1} \right] \leq \exp \left( 2\lambda^2 c^2 (1 + \alpha)^2 \right).$$

This leads us to

$$\mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathbb{G}_r} f(X_i) \right) \right] \leq \exp \left( \lambda^2 c^2 (1 + \alpha)^2 |\mathbb{G}_r| \right) \mathbb{E} \left[ \prod_{i \in \mathbb{G}_{r-1}} \exp(2\lambda Qf(X_i)) \right].$$

Doing the same thing for  $\mathbb{E} \left[ \prod_{i \in \mathbb{G}_{r-1}} \exp(2\lambda Qf(X_i)) \right]$  with  $Qf$  replacing  $f$ , we get

$$\mathbb{E} \left[ \prod_{i \in \mathbb{G}_{r-1}} \exp(2\lambda Qf(X_i)) \right] \leq \exp \left( 2\lambda^2 c^2 (\alpha + \alpha^2)^2 |\mathbb{G}_r| \right) \mathbb{E} \left[ \prod_{i \in \mathbb{G}_{r-2}} \exp(2^2 \lambda Q^2 f(X_i)) \right].$$

Iterating this procedure, we get

$$\begin{aligned} \mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathbb{G}_r} f(X_i) \right) \right] &\leq \exp \left( \lambda^2 c^2 (1 + \alpha)^2 |\mathbb{G}_r| \right) \times \exp \left( 2\lambda^2 c^2 (\alpha + \alpha^2)^2 |\mathbb{G}_r| \right) \\ &\quad \times \exp \left( 2^2 \lambda^2 c^2 (\alpha^2 + \alpha^3)^2 |\mathbb{G}_r| \right) \times \cdots \times \exp \left( 2^{r-1} \lambda^2 c^2 (\alpha^{r-1} + \alpha^r)^2 |\mathbb{G}_r| \right) \\ &\quad \times \mathbb{E} \left[ \exp \left( 2^r \lambda Q^r f(X_1) \right) \right]. \end{aligned}$$

Once again, according to **(H2)**, we have

$$\begin{aligned} \mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathbb{G}_r} f(X_i) \right) \right] &\leq \exp \left( \lambda^2 c^2 (1 + \alpha)^2 |\mathbb{G}_r| \right) \times \exp \left( 2\lambda^2 c^2 (\alpha + \alpha^2)^2 |\mathbb{G}_r| \right) \\ &\quad \times \exp \left( 2^2 \lambda^2 c^2 (\alpha^2 + \alpha^3)^2 |\mathbb{G}_r| \right) \times \cdots \times \exp \left( 2^{r-1} \lambda^2 c^2 (\alpha^{r-1} + \alpha^r)^2 |\mathbb{G}_r| \right) \\ &\quad \times \exp \left( \lambda \alpha^r |\mathbb{G}_r| \right). \end{aligned}$$

Hence

- if  $\alpha^2 \neq \frac{1}{2}$  then  $\mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathbb{G}_r} f(X_i) \right) \right] \leq \exp \left( \lambda^2 c^2 (1 + \alpha)^2 \frac{1 - (2\alpha^2)^r}{1 - 2\alpha^2} |\mathbb{G}_r| \right)$   
 $\times \exp \left( \lambda c \alpha^r |\mathbb{G}_r| \right);$
- if  $\alpha^2 = \frac{1}{2}$  then  $\mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathbb{G}_r} f(X_i) \right) \right] \leq \exp \left( \lambda^2 c^2 (1 + \alpha)^2 r |\mathbb{G}_r| \right)$   
 $\times \exp \left( \lambda c \left( \frac{\sqrt{2}}{2} \right)^r |\mathbb{G}_r| \right).$

We then consider three cases.

**(a)** If  $\alpha^2 < \frac{1}{2}$ . Then, there exists a positive constant  $c'$  such that  $\frac{1 - (2\alpha^2)^r}{1 - 2\alpha^2} < c'$  for all  $r$ .

Taking  $\lambda = \frac{\delta}{2c^2 c' (1 + \alpha)^2}$  in (3.4) leads us to

$$\mathbb{P} \left( \overline{M}_{\mathbb{G}_r}(f) > \delta \right) \leq \exp \left( - \left( \frac{\delta^2}{4c^2 c' (1 + \alpha)^2} - \alpha^r \frac{\delta}{4cc' (1 + \alpha)^2} \right) |\mathbb{G}_r| \right).$$

Since

$$\alpha^r \frac{\delta}{4cc' (1 + \alpha)^2} \rightarrow 0 \quad \text{as } r \rightarrow \infty,$$

there exists  $r_0 \in \mathbb{N}$  such that for  $r > r_0$ ,

$$\alpha^r \frac{\delta}{4cc' (1 + \alpha)^2} < \frac{\delta^2}{8c^2 c' (1 + \alpha)^2},$$

and hence for  $r > r_0$ , we get

$$\mathbb{P} \left( \overline{M}_{\mathbb{G}_r}(f) > \delta \right) \leq \exp \left( - \frac{\delta^2}{8c^2 c' (1 + \alpha)^2} |\mathbb{G}_r| \right).$$

**(b)** If  $\alpha^2 = \frac{1}{2}$ . Then as in the previous case, there exists  $r_0 \in \mathbb{N}$  such that for  $r > r_0$

$$\mathbb{P} \left( \overline{M}_{\mathbb{G}_r}(f) > \delta \right) \leq \exp \left( - \frac{\delta^2}{8c^2 (1 + \alpha)^2} \frac{|\mathbb{G}_r|}{r} \right).$$

**(c)** If  $\alpha^2 > \frac{1}{2}$ . Then for all  $\lambda > 0$

$$\begin{aligned} \mathbb{P} \left( \overline{M}_{\mathbb{G}_r}(f) > \delta \right) &\leq \exp \left( - \lambda \delta |\mathbb{G}_r| \right) \times \exp \left( \lambda^2 c^2 (1 + \alpha)^2 \frac{(2\alpha^2)^r - 1}{2\alpha^2 - 1} |\mathbb{G}_r| \right) \times \exp \left( \lambda c \alpha^r |\mathbb{G}_r| \right) \\ &\leq \exp \left( - |\mathbb{G}_r| \left( \lambda \delta - \frac{\lambda^2 c^2 (1 + \alpha)^2}{2\alpha^2 - 1} (2\alpha^2)^r \right) \right) \times \exp \left( \lambda c \alpha^r |\mathbb{G}_r| \right). \end{aligned}$$

Taking  $\lambda = \frac{(2\alpha^2 - 1)\delta}{2c^2 (1 + \alpha)^2 (2\alpha^2)^r}$  leads us to

$$\mathbb{P} \left( \overline{M}_{\mathbb{G}_r}(f) > \delta \right) \leq \exp \left( - \left( \frac{(2\alpha^2 - 1)\delta^2}{4c^2 (1 + \alpha)^2} - \frac{(2\alpha^2 - 1)\delta}{2c(1 + \alpha)^2} \alpha^r \right) \left( \frac{1}{\alpha} \right)^{2r} \right).$$

Since  $\alpha^r \rightarrow 0$  as  $r \rightarrow \infty$ , there exists  $r_0 \in \mathbb{N}$  such that for all  $r > r_0$ ,

$$\mathbb{P} \left( \overline{M}_{\mathbb{G}_r}(f) > \delta \right) \leq \exp \left( - \frac{(2\alpha^2 - 1)\delta^2}{8c^2 (1 + \alpha)^2} \left( \frac{1}{\alpha} \right)^{2r} \right).$$

**Part 2.** Let us now deal with  $\overline{M}_{\mathbb{T}_r}(f)$ . We have for all  $\lambda > 0$  and all  $\delta > 0$

$$\mathbb{P}\left(\overline{M}_{\mathbb{T}_r}(f) > \delta\right) \leq \exp\left(-\lambda\delta|\mathbb{T}_r|\right) \mathbb{E}\left[\exp\left(\lambda \sum_{i \in \mathbb{T}_r} f(X_i)\right)\right]. \quad (3.5)$$

By subtracting and adding terms, we get

$$\begin{aligned} & \mathbb{E}\left[\exp\left(\lambda \sum_{i \in \mathbb{T}_r} f(X_i)\right)\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\prod_{i \in \mathbb{G}_{r-1}} \exp\left(\lambda(f(X_{2i}) + f(X_{2i+1}) - 2Qf(X_i))\right) \times \prod_{i \in \mathbb{G}_{r-1}} \exp\left(2\lambda Qf(X_i)\right) \right. \right. \\ & \qquad \qquad \qquad \left. \left. \times \prod_{i \in \mathbb{T}_{r-1}} \exp\left(\lambda f(X_i)\right) / \mathcal{F}_{r-1}\right)\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\prod_{i \in \mathbb{G}_{r-1}} \exp\left(\lambda(f(X_{2i}) + f(X_{2i+1}) - 2Qf(X_i))\right) \times \prod_{i \in \mathbb{G}_{r-1}} \exp\left(\lambda(f + 2Qf)(X_i)\right) \right. \right. \\ & \qquad \qquad \qquad \left. \left. \times \prod_{i \in \mathbb{T}_{r-2}} \exp\left(\lambda f(X_i)\right) / \mathcal{F}_{r-1}\right)\right] \end{aligned}$$

The fact that conditionally to the  $(r-1)$  first generations the sequence  $\{\Delta_i, i \in \mathbb{G}_{r-1}\}$  is a sequence of independent random variables and Azuma-Bennet-Hoeffding inequalities leads us according to **(H2)** to

$$\begin{aligned} \mathbb{E}\left[\exp\left(\lambda \sum_{i \in \mathbb{T}_r} f(X_i)\right)\right] &\leq \exp\left(2\lambda^2 c^2 (1 + \alpha)^2 |\mathbb{G}_{r-1}|\right) \\ &\times \mathbb{E}\left[\prod_{i \in \mathbb{G}_{r-1}} \exp\left(\lambda(f + 2Qf)(X_i)\right) \prod_{i \in \mathbb{T}_{r-2}} \exp\left(\lambda f(X_i)\right)\right]. \end{aligned}$$

Doing the same things for

$$\mathbb{E}\left[\prod_{i \in \mathbb{G}_{r-1}} \exp\left(\lambda(f + 2Qf)(X_i)\right) \prod_{i \in \mathbb{T}_{r-2}} \exp\left(\lambda f(X_i)\right)\right]$$

with  $f + 2Qf$  replacing  $f$  we get

$$\begin{aligned} \mathbb{E}\left[\exp\left(\lambda \sum_{i \in \mathbb{T}_r} f(X_i)\right)\right] &\leq \exp\left(2\lambda^2 c^2 (1 + \alpha)^2 |\mathbb{G}_{r-1}|\right) \times \exp\left(2\lambda^2 c^2 (1 + 3\alpha + 2\alpha^2)^2 |\mathbb{G}_{r-2}|\right) \\ &\times \mathbb{E}\left[\prod_{i \in \mathbb{G}_{r-2}} \exp\left(\lambda(f + 2Qf + 2^2 Q^2 f)(X_i)\right) \prod_{i \in \mathbb{T}_{r-3}} \exp\left(\lambda f(X_i)\right)\right]. \end{aligned}$$

Iterating this procedure leads us to

$$\begin{aligned} \mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathbb{T}_r} f(X_i) \right) \right] &\leq \exp \left( 2\lambda^2 c^2 (1 + \alpha)^2 \sum_{q=1}^r \left( \sum_{k=0}^{q-1} (2\alpha)^{2k} \right)^2 |\mathbb{G}_{r-q}| \right) \\ &\quad \times \mathbb{E} \left[ \exp \left( \lambda (f + 2Qf + 2^2f + \dots + 2^r Q^r f)(X_1) \right) \right]. \end{aligned}$$

Now for  $\alpha \neq \frac{1}{2}$  and  $\alpha^2 \neq \frac{1}{2}$  we have

$$\begin{aligned} &\mathbb{P}(\overline{M}_{\mathbb{T}_r}(f) > \delta) \\ &\leq \exp \left( -\lambda \delta |\mathbb{T}_r| \right) \exp \left( 2\lambda^2 c^2 (1 + \alpha)^2 \left( \frac{2^r - 1}{(1 - 2\alpha)^2} - \frac{\alpha(1 - \alpha^r)2^{r+1}}{(1 - 2\alpha)^2(1 - \alpha)} + \frac{2\alpha^2(1 - (2\alpha^2)^r)2^r}{(1 - 2\alpha)^2(1 - 2\alpha^2)} \right) \right) \\ &\quad \times \exp \left( \lambda c \frac{1 - (2\alpha)^{r+1}}{1 - 2\alpha} \right) \\ &\leq \exp \left( -|\mathbb{T}_r| \left( \lambda \delta - \frac{\lambda^2 c^2 (1 + \alpha)^2}{(1 - 2\alpha)^2} \left( 1 + \frac{4\alpha^2(1 - (2\alpha^2)^r)}{1 - 2\alpha^2} \right) \right) \right) \exp \left( \lambda c \frac{1 - (2\alpha)^{r+1}}{1 - 2\alpha} \right). \end{aligned}$$

Taking  $\lambda = \frac{\delta}{\frac{2c^2(1+\alpha)^2}{(1-2\alpha)^2} \left( 1 + \frac{4\alpha^2(1-(2\alpha^2)^r)}{1-2\alpha^2} \right)}$  leads us to

$$\begin{aligned} \mathbb{P}(\overline{M}_{\mathbb{T}_r}(f) > \delta) &\leq \exp \left( -|\mathbb{T}_r| \frac{(1-\alpha)^2 \delta^2}{4c^2(1+\alpha)^2 \left( 1 + \frac{4\alpha^2(1-(2\alpha^2)^r)}{1-2\alpha^2} \right)} \right) \\ &\quad \times \exp \left( \frac{(1-\alpha)^2 \delta}{2c(1+\alpha)^2 \left( 1 + \frac{4\alpha^2(1-(2\alpha^2)^r)}{1-2\alpha^2} \right)} \frac{1 - (2\alpha)^{r+1}}{1 - 2\alpha} \right). \end{aligned}$$

- If  $\alpha^2 < \frac{1}{2}$  and  $\alpha < \frac{1}{2}$  then for all  $r \in \mathbb{N}$

$$\mathbb{P}(\overline{M}_{\mathbb{T}_r}(f) > \delta) \leq \exp \left( \frac{1 - 2\alpha}{2c(1 + \alpha)^2} \delta \right) \times \exp \left( -\frac{(1 - 2\alpha)^2 \delta^2}{4c^2(1 + \alpha)^2(1 + 4\alpha^2 c')} |\mathbb{T}_r| \right)$$

such that  $\frac{1 - (2\alpha^2)^r}{1 - 2\alpha^2} < c'$ .

- If  $\alpha^2 < \frac{1}{2}$  and  $\frac{1}{2} < \alpha < \frac{\sqrt{2}}{2}$  then there exists  $r_0 \in \mathbb{N}$  such that for all  $r > r_0$

$$\mathbb{P}(\overline{M}_{\mathbb{T}_r}(f) > \delta) \leq \exp \left( -\frac{(1 - 2\alpha)^2 \delta^2}{8c^2(1 + \alpha)^2(1 + 4\alpha^2 c')} |\mathbb{T}_r| \right).$$

- If  $\alpha^2 > \frac{1}{2}$  then there exists  $r_0 \in \mathbb{N}$  such that for all  $r > r_0$

$$\mathbb{P}(\overline{M}_{\mathbb{T}_r}(f) > \delta) \leq \exp \left( -\frac{(1 - 2\alpha)^2 (2\alpha^2 - 1) \delta^2}{64c^2(1 + \alpha)^2} \left( \frac{1}{\alpha^2} \right)^{r+1} \right).$$

Now if  $\alpha = \frac{1}{2}$  then for all  $\lambda > 0$

$$\mathbb{P}(\overline{M}_{\mathbb{T}_r}(f) > \delta) \leq \exp \left( -\left( \lambda \delta - \frac{9c^2 c' \lambda^2}{2} \right) |\mathbb{T}_r| \right) \times \exp(\lambda(r + 1)) \quad \text{where } c' = \sum_{q=1}^{\infty} \frac{q^2}{2^q}.$$

Taking  $\lambda = \frac{\delta}{9c^2c'}$  leads us to

$$\mathbb{P}\left(\overline{M}_{\mathbb{T}_r}(f) > \delta\right) \leq \exp\left(-\frac{\delta^2}{18c^2c'}|\mathbb{T}_r|\right) \times \exp\left(\frac{\delta}{9c^2c'}(r+1)\right).$$

Finally, if  $\alpha^2 = \frac{1}{2}$  then there exists  $r_0 \in \mathbb{N}$  such that for all  $r > r_0$

$$\mathbb{P}\left(\overline{M}_{\mathbb{T}_r}(f) > \delta\right) \leq \exp\left(-\frac{(1-2\alpha)^2\delta^2}{16c^2(1+\alpha)^2} \frac{|\mathbb{T}_r|}{r+1}\right).$$

**Part 3.** Eventually, let us look at  $\overline{M}_n^\pi(f)$ . We have for all  $\delta > 0$

$$\mathbb{P}\left(\frac{1}{n}M_n^\pi(f) > \delta\right) \leq \mathbb{P}\left(\frac{1}{n}\sum_{q=0}^{r_{n-1}} f(X_i) > \frac{\delta}{2}\right) + \mathbb{P}\left(\frac{1}{n}\sum_{i=2^{r_n}}^n f(X_{\pi(i)}) > \frac{\delta}{2}\right).$$

On the one hand, (3.2) leads us for  $n$  large enough to

$$\mathbb{P}\left(\frac{1}{n}\sum_{q=0}^{r_{n-1}} f(X_i) > \frac{\delta}{2}\right) \leq \begin{cases} c_1 \exp(-c_2\delta^2n) & \text{if } \alpha < \frac{1}{2} \\ 2 \exp(-c_1\delta^2n) & \text{if } \frac{1}{2} < \alpha < \frac{\sqrt{2}}{2} \\ \exp(-c_1\delta^2(\frac{1}{\alpha^2})^{r_n}) & \text{if } \alpha^2 > \frac{1}{2} \\ \exp(-c_1\delta^2n) \exp(2c_2\delta r_n) & \text{if } \alpha = \frac{1}{2} \\ \exp(-c_1\delta^2\frac{n}{r_n}) & \text{if } \alpha^2 = \frac{1}{2}, \end{cases} \quad (3.6)$$

where  $c_1$  and  $c_2$  are positive constants which depend on  $\delta$ ,  $\alpha$ ,  $\|f\|_\infty$  and  $c$ ;  $c_1$  differs line by line. On the other hand, for all  $\lambda > 0$ ,

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=2^{r_n}}^n f(X_{\pi(i)}) > \frac{\delta}{2}\right) \leq \exp\left(-\frac{\lambda\delta}{2}n\right) \mathbb{E}\left[\exp\left(\lambda\sum_{i=2^{r_n}}^n f(X_{\pi(i)})\right)\right].$$

Now let

- $\mathcal{O}_{r_n} = \{\pi(2^{r_n}), \pi(2^{r_n} + 1), \dots, \pi(n)\}$ ,
- $\mathcal{O}_{r_{n-1}}^1$  the set of individuals of generation  $\mathbb{G}_{r_{n-1}}$  which are ancestors of one individual in  $\mathcal{O}_{r_n}$ ,
- $\mathcal{O}_{r_{n-1}}^2$  the set of individuals of generation  $\mathbb{G}_{r_{n-1}}$  which are ancestors of two individuals in  $\mathcal{O}_{r_n}$ ,
- $\mathcal{O}'_{r_n}$  the set of individuals of  $\mathcal{O}_{r_n}$  whose parents belong to  $\mathcal{O}_{r_{n-1}}^1$ .
- $\mathcal{O}_{r_{n-1}} = \mathcal{O}_{r_{n-1}}^1 \cup \mathcal{O}_{r_{n-1}}^2$ .

We have

$$\begin{aligned} \mathbb{E} \left[ \exp \left( \lambda \sum_{i=2^{r_n}}^n f(X_{\pi(i)}) \right) \right] &= \mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathcal{O}_{r_n-1}^2} 2Qf(X_i) + \lambda \sum_{i \in \mathcal{O}_{r_n-1}^1} Qf(X_i) \right) \right] \\ &\quad \times \mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathcal{O}'_{r_n}} f(X_i) - Qf(X_{\lfloor \frac{i}{2} \rfloor}) \right) / \mathcal{F}_{r_n-1} \right] \\ &\quad \times \mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathcal{O}_{r_n-1}^2} f(X_{2i}) + f(X_{2i+1}) - 2Qf(X_{\lfloor \frac{i}{2} \rfloor}) \right) / \mathcal{F}_{r_n-1} \right] \end{aligned}$$

Using Azuma-Bennett-Hoeffding inequality, as in **Part 1**, we get

$$\mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathcal{O}'_{r_n}} f(X_i) - Qf(X_{\lfloor \frac{i}{2} \rfloor}) \right) / \mathcal{F}_{r_n-1} \right] \leq \exp \left( \frac{\lambda^2 c^2 (1+\alpha)^2}{2} |\mathcal{O}'_{r_n}| \right),$$

and

$$\mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathcal{O}_{r_n-1}^2} f(X_{2i}) + f(X_{2i+1}) - 2Qf(X_{\lfloor \frac{i}{2} \rfloor}) \right) / \mathcal{F}_{r_n-1} \right] \leq \exp \left( 2\lambda^2 c^2 (1+\alpha)^2 |\mathcal{O}_{r_n-1}^2| \right).$$

Now, we have

$$\begin{aligned} &\exp \left( \frac{\lambda^2 c^2 (1+\alpha)^2}{2} |\mathcal{O}'_{r_n}| \right) + \exp \left( 2\lambda^2 c^2 (1+\alpha)^2 |\mathcal{O}_{r_n-1}^2| \right) \\ &= \exp \left( \lambda^2 c^2 (1+\alpha)^2 \left( 2|\mathcal{O}_{r_n-1}^2| + \frac{|\mathcal{O}'_{r_n}|}{2} \right) \right) \\ &\leq \exp \left( \lambda^2 c^2 (1+\alpha)^2 n \right). \end{aligned}$$

This leads us to

$$\begin{aligned} &\mathbb{E} \left[ \exp \left( \lambda \sum_{i=2^{r_n}}^n f(X_{\pi(i)}) \right) \right] \\ &\leq \exp \left( \lambda^2 c^2 (1+\alpha)^2 n \right) \mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathcal{O}_{r_n-1}^2} 2Qf(X_i) + \lambda \sum_{i \in \mathcal{O}_{r_n-1}^1} Qf(X_i) \right) \right]. \end{aligned}$$

Now let

- $\mathcal{O}_{r_n-2}^{1,1}$  the set of individuals of  $\mathbb{G}_{r_n-2}$  which are ancestors of one individual in  $\mathcal{O}_{r_n-1}$  and one individual in  $\mathcal{O}_{r_n}$ ,
- $\mathcal{O}_{r_n-2}^{1,2}$  the set of individuals of  $\mathbb{G}_{r_n-2}$  which are ancestors of one individual in  $\mathcal{O}_{r_n-1}$  and two individuals in  $\mathcal{O}_{r_n}$ ,
- $\mathcal{O}_{r_n-2}^{2,2}$  the set of individuals of  $\mathbb{G}_{r_n-2}$  which are ancestors of two individuals in  $\mathcal{O}_{r_n-1}$  and two individuals in  $\mathcal{O}_{r_n}$ ,



- $\mathcal{O}_{r_n-2}^{2,3}$  the set of individuals of  $\mathbb{G}_{r_n-2}$  which are ancestors of two individuals in  $\mathcal{O}_{r_n-1}$  and three individuals in  $\mathcal{O}_{r_n}$ ,
- $\mathcal{O}_{r_n-2}^{2,4}$  the set of individuals of  $\mathbb{G}_{r_n-2}$  which are ancestors of two individuals in  $\mathcal{O}_{r_n-1}$  and four individuals in  $\mathcal{O}_{r_n}$ ,
- $\mathcal{O}'_{r_n-1}$  the set of individuals of  $\mathcal{O}_{r_n-1}$  whose parents belong to  $\mathcal{O}_{r_n-2}^{1,1}$ ,
- $\mathcal{O}''_{r_n-1}$  the set of individuals of  $\mathcal{O}_{r_n-1}$  whose parents belong to  $\mathcal{O}_{r_n-2}^{1,2}$ .

Then we have

$$\mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathcal{O}_{r_n-1}^2} 2Qf(X_i) + \lambda \sum_{i \in \mathcal{O}_{r_n-1}^1} Qf(X_i) \right) \right] = \mathbb{E} [I_1 \times I_2 \times I_3 \times I_4 \times I_5 \times I_6 \times I_7],$$

where

$$\begin{aligned} I_1 &= \exp \left( \lambda \sum_{i \in \mathcal{O}_{r_n-2}^{1,1}} Q^2f(X_i) + \lambda \sum_{i \in \mathcal{O}_{r_n-2}^{1,2}} 2Q^2f(X_i) + \lambda \sum_{i \in \mathcal{O}_{r_n-2}^{2,2}} 2Q^2f(X_i) \right. \\ &\quad \left. + \lambda \sum_{i \in \mathcal{O}_{r_n-2}^{2,3}} 3Q^2f(X_i) + \lambda \sum_{i \in \mathcal{O}_{r_n-2}^{2,4}} 4Q^2f(X_i) \right), \\ I_2 &= \mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathcal{O}'_{r_n-1}} Qf(X_i) - Q^2f(X_{[\frac{i}{2}]}) \right) / \mathcal{F}_{r_n-2} \right], \\ I_3 &= \mathbb{E} \left[ \exp \left( 2\lambda \sum_{i \in \mathcal{O}''_{r_n-1}} Qf(X_i) - Q^2f(X_{[\frac{i}{2}]}) \right) / \mathcal{F}_{r_n-2} \right], \\ I_4 &= \mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathcal{O}_{r_n-1}^{2,2}} Qf(X_{2i}) + Qf(X_{2i+1}) - 2Q^2f(X_i) \right) / \mathcal{F}_{r_n-2} \right], \\ I_5 &= \frac{1}{2} \mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathcal{O}_{r_n-1}^{2,3}} 2Qf(X_{2i}) + Qf(X_{2i+1}) - 3Q^2f(X_i) \right) / \mathcal{F}_{r_n-2} \right], \\ I_6 &= \frac{1}{2} \mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathcal{O}_{r_n-1}^{2,3}} Qf(X_{2i}) + 2Qf(X_{2i+1}) - 3Q^2f(X_i) \right) / \mathcal{F}_{r_n-2} \right], \\ I_7 &= \mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathcal{O}_{r_n-1}^{2,4}} 2Qf(X_{2i}) + 2Qf(X_{2i+1}) - 4Q^2f(X_i) \right) / \mathcal{F}_{r_n-2} \right]. \end{aligned}$$

Using Azuma-Bennett-Hoeffding inequality, we get

$$\begin{aligned} &I_2 \times I_3 \times I_4 \times I_5 \times I_6 \times I_7 \\ &\leq \exp \left( \lambda^2 c^2 (\alpha + \alpha^2)^2 \left( \frac{|\mathcal{O}'_{r_n-1}|}{2} + 2|\mathcal{O}''_{r_n-1}| + 2|\mathcal{O}_{r_n-1}^{2,2}| + \frac{9|\mathcal{O}_{r_n-1}^{2,3}|}{2} + 8|\mathcal{O}_{r_n-1}^{2,4}| \right) \right) \\ &\leq \exp \left( 2\lambda^2 c^2 (\alpha + \alpha^2)^2 n \right), \end{aligned}$$

hence

$$\mathbb{E} \left[ \exp \left( \lambda \sum_{i=2^{rn}}^n f(X_{\pi(i)}) \right) \right] \leq \exp \left( \lambda^2 c^2 (1 + \alpha)^2 n \right) \exp \left( 2\lambda^2 c^2 (\alpha + \alpha^2)^2 n \right) \mathbb{E} [I_1].$$

Now, iterating this procedure we get

$$\mathbb{E} \left[ \exp \left( \lambda \sum_{i=2^{rn}}^n f(X_{\pi(i)}) \right) \right] \leq \exp \left( \lambda^2 c^2 (1 + \alpha)^2 n \sum_{p=0}^{r_n} (2\alpha^2)^p \right) \exp \left( \lambda c \alpha^{r_n} n \right).$$

Then it follows as in **Part 1** that for  $n$  large enough

$$\mathbb{P} \left( \frac{1}{n} \sum_{i=2^{rn}}^n f(X_{\pi(i)}) > \frac{\delta}{2} \right) \leq \begin{cases} \exp(-c_1 \delta^2 n) & \text{if } \alpha^2 < \frac{1}{2} \\ \exp(-c_1 \delta^2 (\frac{1}{\alpha^2})^{r_n}) & \text{if } \alpha^2 > \frac{1}{2} \\ \exp(-c_1 \frac{n}{r_n}) & \text{if } \alpha^2 = \frac{1}{2}, \end{cases} \quad (3.7)$$

where the positive constant  $c_1$  depends on  $\alpha$ ,  $\delta$ ,  $c$  and differs line to line. Finally (3.6) and (3.7) lead us to (3.3).  $\square$

**Theorem 3.2.** *Let  $f \in \mathcal{B}_b(S^3)$  such that  $(\mu, Pf) = 0$ . Assume that **(H2)** is satisfied. Then we have the same conclusions, for the three empirical averages  $\overline{M}_{\mathbb{G}_r}(f)$ ,  $\overline{M}_n^n(f)$  and  $\overline{M}_{\mathbb{T}_r}(f)$ , as in the theorem 3.1.*

*Proof.* Let  $f \in \mathcal{B}_b(S^3)$  such that  $(\mu, Pf) = 0$ .

**Part 1.** Let us first deal with  $\overline{M}_{\mathbb{G}_r}(f)$ . We have for all  $\delta > 0$  and  $\lambda > 0$ ,

$$\mathbb{P} \left( \overline{M}_{\mathbb{G}_r}(f) > \delta \right) \leq \exp \left( -\lambda \delta |\mathbb{G}_r| \right) \mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathbb{G}_r} f(\Delta_i) \right) \right].$$

Conditioning and using Bennet-Hoeffding inequality give us

$$\mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathbb{G}_r} f(\Delta_i) \right) \right] \leq \exp \left( 2\lambda^2 \|f\|_{\infty} |\mathbb{G}_r| \right) \mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathbb{G}_r} Pf(X_i) \right) \right].$$

Now, applying the **Part 1** of the proof of the theorem 3.1 to  $Pf$ , we get (3.1) for  $f \in \mathcal{B}_b(S^3)$ .

**Part 2.** Let us now treat  $\overline{M}_{\mathbb{T}_r}(f)$ . We have for all  $\delta > 0$

$$\mathbb{P} \left( \overline{M}_{\mathbb{T}_r}(f) > \delta \right) \leq \mathbb{P} \left( \overline{M}_{\mathbb{T}_r}(f - Pf) > \frac{\delta}{2} \right) + \mathbb{P} \left( \overline{M}_{\mathbb{T}_r}(Pf) > \frac{\delta}{2} \right) \quad (3.8)$$

Now, since  $\left(\overline{M}_n^\pi(f - Pf)\right)_{n \geq 1}$  is a  $\mathcal{H}_n$ -martingale with bounded jumps, Azuma inequality [1], give us

$$\mathbb{P}\left(\overline{M}_{\mathbb{T}_r}(f - Pf) > \frac{\delta}{2}\right) \leq \exp\left(-c'\delta^2|\mathbb{T}_r|\right), \quad \text{for some positive constant } c'.$$

For the second term in the right hand side of (3.8), we use the inequalities (3.2) with  $Pf$  instead of  $f$ . Gathering this inequalities, we get (3.2) for all  $r$  large enough.

**Part 3.** The proof for the case  $\overline{M}_n^\pi(f)$  follows the same lines as the proof of **Part 2**.  $\square$

Now, using Borel Cantelli theorem and (3.3), we state easily the following

**Corollary 3.3.** *Let  $f \in \mathcal{B}_b(S)$  such that  $(\mu, f) = 0$  (resp.  $f \in \mathcal{B}_b(S^3)$  and  $(\mu, Pf) = 0$ ). Assume that **(H2)** is satisfied. Then  $\overline{M}_n^\pi(f)$  almost surely converges to 0 as  $n$  goes to  $\infty$ .*

**Remark 3.4.** Of course uniform ergodicity, and bounded test functions are surely a very strong set of assumptions but it is not so difficult to verify if the Markov chains daughters lie in a compact set. We are convinced that it is possible to consider the geometric ergodic case and bounded test functions but to the price of tedious calculations that we will pursue in an other work. We will also investigate the use of transportation inequalities, leading to deviation inequality for Lipschitz test functions under some Wasserstein contraction property for the kernel  $P$ , in the spirit of the Theorems 2.5 or 2.11 in [10].

### 3.2. Moderate deviation principle for BMC.

We introduce the following assumption on the speed of the MDP.

**Assumption 1.** *Let  $(b_n)$  be an increasing sequence of positive real numbers such that*

$$\frac{b_n}{\sqrt{n}} \longrightarrow +\infty,$$

and

- if  $\alpha^2 < \frac{1}{2}$ , the sequence  $(b_n)$  is such that  $\frac{b_n}{n} \longrightarrow 0$ ,
- if  $\alpha^2 = \frac{1}{2}$ , the sequence  $(b_n)$  is such that  $\frac{b_n \log n}{n} \longrightarrow 0$ ,
- if  $\alpha^2 > \frac{1}{2}$ , the sequence  $(b_n)$  is such that  $\frac{b_n \alpha^{r_n+1}}{\sqrt{n}} \longrightarrow 0$ .

Using the MDP for martingale with bounded jumps (see e.g [7], [9]), we can now state the following

**Theorem 3.5.** *Let  $f \in \mathcal{B}_b(S^3)$  such that  $Pf = 0$ . Assume that **(H2)** is satisfied. Let  $(b_n)$  be a sequence of real numbers satisfying the Assumption 1, then  $\left(\frac{1}{b_n} M_n^\pi(f)\right)$  satisfies a MDP in  $\mathbb{R}$  with the speed  $\frac{b_n^2}{n}$  and rate function  $I(x) = \frac{x^2}{2(\mu, Pf^2)}$ .*

*Proof.* The proof easily follows from the previous exponential probability inequalities and the MDP for martingale with bounded jumps (see e.g [7], [9], [18]).  $\square$

**Remark 3.6.** Taking particularly  $n = |\mathbb{T}_r|$ , and  $(b_n)$  be a sequence of real numbers satisfying the Assumption 1 we get that  $\left(\frac{M_{\mathbb{T}_r}}{b_{|\mathbb{T}_r|}}\right)$  satisfies a MDP in  $\mathbb{R}$  with the speed  $\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}$  and the rate function  $I(x) = \frac{x^2}{2(\mu, Pf^2)}$ .

#### 4. APPLICATION: FIRST ORDER BIFURCATING AUTOREGRESSIVE PROCESSES

In this section, we seek to apply the results of the previous sections to the following bifurcating autoregressive process with memory 1 defined by

$$\mathcal{L}(X_1) = \nu, \quad \text{and} \quad \forall n \geq 1, \quad \begin{cases} X_{2n} = \alpha_0 X_n + \beta_0 + \varepsilon_{2n} \\ X_{2n+1} = \alpha_1 X_n + \beta_1 + \varepsilon_{2n+1}, \end{cases} \quad (4.1)$$

where  $\alpha_0, \alpha_1 \in (-1, 1)$ ;  $\beta_0, \beta_1 \in \mathbb{R}$ ,  $\left((\varepsilon_{2n}, \varepsilon_{2n+1}), n \geq 1\right)$  forms a sequence of i.i.d bivariate random variables and  $\nu$  a probability measure on  $\mathbb{R}$ . In all this section, we assume that the distribution  $\nu$  has finite moments of all orders.

In the sequel, we will study (4.1) in two settings:

- the gaussian setting which corresponds to the case where  $\left((\varepsilon_{2n}, \varepsilon_{2n+1}), n \geq 1\right)$  forms a sequence of i.i.d bivariate random variables with law  $\mathcal{N}_2(0, \Gamma)$  with

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

- the bounded setting which corresponds to the case where  $\left((\varepsilon_{2n}, \varepsilon_{2n+1}), n \geq 1\right)$  forms a sequence of i.i.d bivariate random variables values in a compact set. Let us note that in this case,  $(X_n, n \in \mathbb{T})$  takes its values in a compact set.

Our main goal is to give deviation inequalities and MDP for the estimator of the 4-dimensional unknown parameter  $\theta = (\alpha_0, \beta_0, \alpha_1, \beta_1)$  and for the statistical test defined in [12].

To estimate the 4-parameter  $\theta = (\alpha_0, \beta_0, \alpha_1, \beta_1)$ , as well as  $\sigma^2$  and  $\rho$ , one assume observe a complete subtree  $\mathbb{T}_{r+1}$ . The maximum likelihood estimator  $\hat{\theta}^r = (\hat{\alpha}_0^r, \hat{\beta}_0^r, \hat{\alpha}_1^r, \hat{\beta}_1^r)$  of  $\theta$  is given by (see [12]), for  $\eta \in \{0, 1\}$

$$\begin{cases} \hat{\alpha}_\eta^r = \frac{|\mathbb{T}_r|^{-1} \sum_{i \in \mathbb{T}_r} X_i X_{2i+\eta} - \left(|\mathbb{T}_r|^{-1} \sum_{i \in \mathbb{T}_r} X_i\right) \left(|\mathbb{T}_r|^{-1} \sum_{i \in \mathbb{T}_r} X_{2i+\eta}\right)}{|\mathbb{T}_r|^{-1} \sum_{i \in \mathbb{T}_r} X_i^2 - \left(|\mathbb{T}_r|^{-1} \sum_{i \in \mathbb{T}_r} X_i\right)^2} \\ \hat{\beta}_\eta^r = |\mathbb{T}_r|^{-1} \sum_{i \in \mathbb{T}_r} X_{2i+\eta} - \hat{\alpha}_\eta^r |\mathbb{T}_r|^{-1} \sum_{i \in \mathbb{T}_r} X_i. \end{cases}$$

We also need to introduce the estimators of the conditional variance  $\sigma^2$  and the conditional sister-sister correlation  $\rho$ . These estimators are naturally given by

$$\begin{cases} \hat{\sigma}_r^2 = \frac{1}{2|\mathbb{T}_r|} \sum_{i \in \mathbb{T}_r} (\hat{\varepsilon}_{2i}^2 + \hat{\varepsilon}_{2i+1}^2) \\ \hat{\rho}_r = \frac{1}{\hat{\sigma}_r^2} \sum_{i \in \mathbb{T}_r} \hat{\varepsilon}_{2i} \hat{\varepsilon}_{2i+1} \end{cases}$$

where the residues are defined by  $\hat{\varepsilon}_{2i+\eta} = X_{2i+\eta} - \hat{\alpha}_\eta^r X_i - \hat{\beta}_\eta^r$ , with  $\eta \in \{0, 1\}$ .

We will take  $F = \mathcal{C}_{pol}(\mathbb{R})$  the set of continuous and polynomially growing functions. Then  $F$  fulfills hypothesis (i)-(vi). Moreover, for all  $f \in F$ , hypothesis **(H1)** holds with  $\alpha = \max(|\alpha_0|, |\alpha_1|)$ . Let  $\mu$  be the unique stationary distribution of the induced Markov chain  $(Y_r, r \in \mathbb{N})$ , see [12] for more details.

We shall denote by  $\mathbf{x}$  (resp.  $\mathbf{x}^2, \mathbf{xy}, \mathbf{y} \dots$ ) the element of  $\mathcal{C}_{pol}(\mathbb{R}^3)$  defined by  $(x, y, z) \mapsto x$  (resp.  $x^2, xy, y, \dots$ ).

We define two continuous functions  $\mu_1 : \Theta \rightarrow \mathbb{R}$  and  $\mu_2 : \Theta \times \mathbb{R}_+^* \rightarrow \mathbb{R}$  by writing

$$(\mu, \mathbf{x}) = \mu_1(\theta) \quad \text{and} \quad (\mu, \mathbf{x}^2) = \mu_2(\theta, \sigma^2),$$

where  $\theta = (\alpha_0, \beta_0, \alpha_1, \beta_1) \in \Theta = (-1, 1) \times \mathbb{R} \times (-1, 1) \times \mathbb{R}$ .

To segregate between  $H_0 = \{(\alpha_0, \beta_0) = (\alpha_1, \beta_1)\}$  and its alternative  $H_1 = \{(\alpha_0, \beta_0) \neq (\alpha_1, \beta_1)\}$ , we shall use the test statistic

$$\chi_r^{(1)} = \frac{|\mathbb{T}_r|}{2\hat{\sigma}_r^2} \left\{ (\hat{\alpha}_0^r - \hat{\alpha}_1^r)^2 (\hat{\mu}_{2,r}^2 - \hat{\mu}_{1,r}^2) + \left( (\hat{\alpha}_0^r - \hat{\alpha}_1^r) \hat{\mu}_{1,r} + \hat{\beta}_0^r - \hat{\beta}_1^r \right)^2 \right\},$$

where we write  $\hat{\mu}_{1,r} = \mu_1(\hat{\theta}^r)$  and  $\hat{\mu}_{2,r} = \mu_2(\hat{\theta}^r, \hat{\sigma}_r)$ .

#### 4.1. The gaussian setting.

We introduce the following assumption on the speed of the MDP.

**Assumption 2.** Let  $(b_n)$  be an increasing sequence of positive real numbers such that

$$\frac{b_n}{\sqrt{n}} \rightarrow +\infty \quad \text{and} \quad \frac{b_n}{\sqrt{n \log n}} \rightarrow 0.$$

**Proposition 4.1.** Let  $(b_n)$  be a sequence of real numbers satisfying Assumption 2. Then

$$\hat{\theta}^r \xrightarrow[\frac{b_n^2}{|\mathbb{T}_r|}]{\text{superexp}} \theta.$$

*Proof.* We will treat the case of  $\hat{\alpha}_0^r$ . The others  $\hat{\beta}_0^r, \hat{\alpha}_1^r$  and  $\hat{\beta}_1^r$  may be treated in a similar way. Note that  $\hat{\alpha}_0^r = \frac{C_r}{B_r}$ , where

$$C_r = \overline{M}_{\mathbb{T}_r}(\mathbf{xy}) - \overline{M}_{\mathbb{T}_r}(\mathbf{x}) \overline{M}_{\mathbb{T}_r}(\mathbf{y}) \quad \text{and} \quad B_r = \overline{M}_{\mathbb{T}_r}(\mathbf{x}^2) - \overline{M}_{\mathbb{T}_r}(\mathbf{x})^2.$$

Now, using lemma (5.1) and proposition (2.13), it follows that

$$\hat{\alpha}_0^r \xrightarrow[\frac{b_n^2}{|\mathbb{T}_r|}]{\text{superexp}} \alpha_0.$$

□

We recall that in the BAR model (4.1), we use  $\alpha = \max\{|\alpha_0|, |\alpha_1|\}$ , so we have the following deviation inequality

**Proposition 4.2.** *For all  $\delta > 0$  and all  $r \in \mathbb{N}$*

$$\mathbb{P}\left(|\hat{\theta}^r - \theta| > \delta\right) \leq \begin{cases} c\left(\frac{1}{4}\right)^{r+1} & \text{if } \alpha^2 < \frac{1}{2} \\ cr^2\left(\frac{1}{4}\right)^{r+1} & \text{if } \alpha^2 = \frac{1}{2} \\ c\alpha^{4(r+1)} & \text{if } \alpha^2 > \frac{1}{2}, \end{cases} \quad (4.2)$$

where the constant  $c$  depends on  $\alpha$ ,  $\delta$ ,  $\mu_1$ ,  $\mu_2$  and differs line by line.

*Proof.* We will prove the deviation inequality for  $|\hat{\alpha}_0^r - \alpha_0|$ . The other deviation inequalities for  $|\hat{\beta}_0^r - \beta_0|$ ,  $|\hat{\alpha}_1^r - \alpha_1|$  and  $|\hat{\beta}_1^r - \beta_1|$  may be treated in a similar way. We have for all  $\delta > 0$

$$\mathbb{P}\left(|\hat{\alpha}_0^r - \alpha_0| > \delta\right) \leq \mathbb{P}\left(\frac{|\overline{M}_{\mathbb{T}_r}(\mathbf{xy} - P(\mathbf{xy}))|}{B_r} > \frac{\delta}{2}\right) + \mathbb{P}\left(\frac{|\overline{M}_{\mathbb{T}_r}(\mathbf{x})| |\overline{M}_{\mathbb{T}_r}(\mathbf{y} - P(\mathbf{y}))|}{B_r} > \frac{\delta}{2}\right).$$

Let  $b := \mu_2(\theta, \sigma^2) - \mu_1(\theta)^2$ . On the one hand,

$$\mathbb{P}\left(\frac{|\overline{M}_{\mathbb{T}_r}(\mathbf{xy} - P(\mathbf{xy}))|}{B_r} > \frac{\delta}{2}\right) \leq \mathbb{P}(B_r < b/2) + \mathbb{P}\left(|\overline{M}_{\mathbb{T}_r}(\mathbf{xy} - P(\mathbf{xy}))| > \frac{\delta b}{4}\right).$$

Now,

$$\begin{aligned} \mathbb{P}(B_r < b/2) &\leq \mathbb{P}\left(-\overline{M}_{\mathbb{T}_r}(\mathbf{x}^2 - \mu_2) > \frac{b}{2}\right) + \mathbb{P}\left(|\overline{M}_{\mathbb{T}_r}(\mathbf{x} - \mu_1)| > \frac{\sqrt{b}}{2\sqrt{2}}\right) \\ &\quad + \mathbb{P}\left(\overline{M}_{\mathbb{T}_r}(\mathbf{x} - \mu_1) > \frac{b}{16\mu_1}\right). \end{aligned}$$

On the other hand, we have

$$\begin{aligned} \mathbb{P}\left(\frac{|\overline{M}_{\mathbb{T}_r}(\mathbf{x})| |\overline{M}_{\mathbb{T}_r}(\mathbf{y} - P(\mathbf{y}))|}{B_r} > \frac{\delta}{2}\right) &\leq \mathbb{P}(B_r < b/2) + \mathbb{P}\left(|\overline{M}_{\mathbb{T}_r}(\mathbf{y} - P(\mathbf{y}))| > \frac{\delta b}{4\mu_1}\right) \\ &\quad + \mathbb{P}\left(|\overline{M}_{\mathbb{T}_r}(\mathbf{y} - P(\mathbf{y}))| > \frac{\delta}{4}\right) + \mathbb{P}\left(|\overline{M}_{\mathbb{T}_r}(\mathbf{x} - \mu_1)| > b\right). \end{aligned}$$

Now, using (2.6) we get

$$\begin{aligned} \mathbb{P}\left(|\overline{M}_{\mathbb{T}_r}(\mathbf{xy} - P(\mathbf{xy}))| > \frac{\delta b}{4}\right) &\leq \frac{c}{\delta^4 b^4} \left(\frac{1}{4}\right)^{r+1}, \\ \mathbb{P}\left(|\overline{M}_{\mathbb{T}_r}(\mathbf{y} - P(\mathbf{y}))| > \frac{\delta b}{4\mu_1}\right) &\leq \frac{c\mu_1^4}{\delta^4 b^4} \left(\frac{1}{4}\right)^{r+1}, \end{aligned}$$

and

$$\mathbb{P} \left( \left| \overline{M}_{\mathbb{T}_r}(\mathbf{y} - P(\mathbf{y})) \right| > \frac{\delta}{4} \right) \leq \frac{c}{\delta^4} \left( \frac{1}{4} \right)^{r+1}$$

where the constant  $c$  can be found as in remark (2.3) and **Part 1** of proof of theorem 2.10. Finally,  $\mathbb{P} \left( -\overline{M}_{\mathbb{T}_r}(\mathbf{x}^2 - \mu_2) > \frac{b}{4} \right)$ ,  $\mathbb{P} \left( \left| \overline{M}_{\mathbb{T}_r}(\mathbf{x} - \mu_1) \right| > \frac{\sqrt{b}}{2} \right)$ ,  $\mathbb{P} \left( \left| \overline{M}_{\mathbb{T}_r}(\mathbf{x} - \mu_1) \right| > b \right)$  and  $\mathbb{P} \left( \overline{M}_{\mathbb{T}_r}(\mathbf{x} - \mu_1) > \frac{b}{16\mu_1} \right)$  can be bounded as in corollary 2.2 and this ends the proof.  $\square$

**Remark 4.3.** The proposition 4.2 can be improved by calculating the  $2k$ -th order moments, with  $k > 2$ , as in the proof of theorem 2.1. But, as we have said, this comes at the price of enormous computation.

**Proposition 4.4.** *Let  $(b_n)$  be a sequence of real numbers satisfying Assumption 2. Then*

$$(\hat{\sigma}_r^2, \hat{\rho}_r) \xrightarrow[\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}]{\text{superexp}} (\sigma^2, \rho).$$

*Proof.* Let us first deal with  $\sigma_r^2$ . We have (see e.g [12])

$$\hat{\sigma}_r^2 - \sigma^2 = \frac{1}{2} \overline{M}_{\mathbb{T}_r}(f(\cdot, \theta)) + D_r$$

where

$$f(x, y, z, \theta) = (y - \alpha_0 x - \beta_0)^2 + (z - \alpha_1 x - \beta_1)^2,$$

and

$$D_r = \frac{1}{2|\mathbb{T}_r|} \sum_{i \in \mathbb{T}_r} (f(\Delta_i, \hat{\theta}^r) - f(\Delta_i, \theta)).$$

By Taylor-Lagrange formula, we can find  $g \in \mathcal{C}_{pol}(\mathbb{R}^3)$  such that (see [12])

$$|D_r| \leq \frac{1}{2} \|\hat{\theta}^r - \theta\| \left( 1 + \|\theta\| + \|\hat{\theta}^r - \theta\| \right) \overline{M}_{\mathbb{T}_r}(g).$$

Now, proposition (2.13) and proposition (4.1) leads us to

$$\hat{\sigma}_r^2 \xrightarrow[\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}]{\text{superexp}} \sigma^2.$$

The proof for  $\hat{\rho}_r$  is very similar.  $\square$

**Proposition 4.5.** *Let  $(b_n)$  be a sequence of real numbers satisfying Assumption 2. Then  $\frac{|\mathbb{T}_r|}{b_{|\mathbb{T}_r|}} (\hat{\theta}^r - \theta)$  satisfies the MDP on  $\mathbb{R}^4$  with the speed  $\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}$  and the rate function  $I$  given by*

$$I(x) = \frac{1}{2} x^t (\Sigma')^{-1} x,$$

where

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

*Proof.* We first observe that

$$\frac{|\mathbb{T}_r|}{b_{|\mathbb{T}_r|}} \left( \hat{\theta}^r - \theta \right) = M(A_r, B_r) \cdot \frac{U^r(f)}{b_{\mathbb{T}_r}}$$

where  $U^r(f) = M_{\mathbb{T}_r}(f - Pf)$ ,  $f = (\mathbf{x}\mathbf{y}, \mathbf{y}, \mathbf{x}\mathbf{z}, \mathbf{z})$ ,  $A_r = \overline{M}_{\mathbb{T}_r}(\mathbf{x})$ ,  $B_r = \overline{M}_{\mathbb{T}_r}(\mathbf{x}^2) - \overline{M}_{\mathbb{T}_r}(\mathbf{x})^2$  and

$$M(A_r, B_r) = \begin{pmatrix} \frac{1}{B_r} & \frac{-A_r}{B_r} & 0 & 0 \\ \frac{-A_r}{B_r} & \frac{B_r + A_r^2}{B_r} & 0 & 0 \\ 0 & 0 & \frac{1}{B_r} & \frac{-A_r}{B_r} \\ 0 & 0 & \frac{-A_r}{B_r} & \frac{B_r + A_r^2}{B_r} \end{pmatrix}.$$

On the one hand, from proposition (2.13)

$$A_r \xrightarrow[\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}]{\text{superexp}} a := \mu_1(\theta), \quad \text{and} \quad B_r \xrightarrow[\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}]{\text{superexp}} b := \mu_2(\theta, \sigma^2) - \mu_1(\theta)^2.$$

So that by lemma (5.1), we obtain that

$$M(A_r, B_r) \xrightarrow[\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}]{\text{superexp}} M(a, b) := \begin{pmatrix} K & 0 \\ 0 & K \end{pmatrix}.$$

On the other hand, let  $\lambda \in \mathbb{R}^4$ . Then studious calculations shows us that

$$G_n(\lambda) = \frac{1}{2n} \lambda^t \langle M^n(f - Pf) \rangle_{n\lambda}$$

is an upper and lower cumulant. Moreover, from proposition (2.13) and lemma (5.1)

$$G_n(\lambda) \xrightarrow[\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}]{\text{superexp}} \frac{1}{2} \lambda^t \Sigma \lambda \quad \text{where} \quad \Sigma = \sigma^2 \begin{pmatrix} K^{-1} & \rho K^{-1} \\ \rho K^{-1} & K^{-1} \end{pmatrix}.$$

We thus deduce that (see e.g [18])  $(\frac{M_n^\pi(f)}{b_n})$  satisfies a MDP on  $\mathbb{R}^4$  with speed  $\frac{b_n^2}{n}$  and the rate function

$$J(x) = x^t \Sigma^{-1} x. \quad (4.3)$$

Taking  $n = |\mathbb{T}_r|$ , it follows that  $\frac{U^r(f)}{b_{|\mathbb{T}_r|}}$  satisfies a MDP with speed  $\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}$  and the rate function  $J$  given in (4.3). Finally, using the contraction principle(see e.g [8]) as in ([19]), we get the result.  $\square$

**Proposition 4.6.** *Let  $(b_n)$  a sequence of real numbers satisfying the Assumption 2. Then under the null hypothesis  $H_0 = \{(\alpha_0, \beta_0) = (\alpha_1, \beta_1)\}$ ,  $\frac{|\mathbb{T}_r|^{1/2}}{b_{|\mathbb{T}_r|}} (\chi_r^{(1)})^{1/2}$  satisfies a MDP on  $\mathbb{R}$  with speed  $\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}$  and the rate function*

$$I'(y) = \begin{cases} \frac{y^2}{2} & \text{if } y \in \mathbb{R}_+ \\ +\infty & \text{otherwise.} \end{cases}$$

Under the alternative hypothesis  $H_1$  of  $H_0$ , we have for all  $A > 0$

$$\limsup_{r \rightarrow \infty} \frac{|\mathbb{T}_r|}{b_{\mathbb{T}_r}^2} \log \mathbb{P} \left( \chi_r^{(1)} < A \right) = -\infty.$$



*Proof.* We have

$$H_0 = \{g(\theta) = 0\} \quad \text{where} \quad g(\theta) = (\alpha_0 - \alpha_1, \beta_0 - \beta_1)^t.$$

From proposition (4.5),  $\frac{|\mathbb{T}_r|}{b_{|\mathbb{T}_r|}}(\hat{\theta}^r - \theta)$  satisfies a MDP on  $\mathbb{R}^4$  with speed  $\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}$  and the rate function  $I(x) = \frac{1}{2}x^t(\Sigma')^{-1}x$ . So that, using the delta method for the MDP (see e.g [11]) and the theory of optimisation, we conclude that  $\frac{|\mathbb{T}_r|}{b_{|\mathbb{T}_r|}}(g(\hat{\theta}^r) - g(\theta))$  satisfies a MDP on  $\mathbb{R}^2$  with speed  $\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}$  and the rate function

$$J(x) = \frac{1}{2}x^t(\Sigma'')^{-1}x, \quad \text{where} \quad \Sigma'' = 2\sigma^2(1 - \rho)K. \quad (4.4)$$

Under the null hypothesis  $H_0$ , we have  $g(\theta) = 0$ , so that  $\frac{|\mathbb{T}_r|}{b_{|\mathbb{T}_r|}}g(\hat{\theta}^r)$  satisfies a MDP on  $\mathbb{R}^2$  with speed  $\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}$  and rate function  $J$  given in (4.4).

Now, since  $K = K(\theta, \sigma)$  is a continuous function of  $(\theta, \sigma)$  (see [12]), so that, letting  $\hat{K}_r = K(\hat{\theta}^r, \hat{\sigma}_r)$ , lemma (5.1), proposition (2.13) and (4.4) entail that

$$\hat{\Sigma}_r'' = 2\hat{\sigma}_r^2(1 - \hat{\rho}_r)\hat{K}_r \xrightarrow[\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}]{\text{superexp}} \Sigma''.$$

It follows using the contraction principle (see e.g [19]) that  $\frac{|\mathbb{T}_r|}{b_{|\mathbb{T}_r|}}\hat{\Sigma}_r''^{-1/2}g(\hat{\theta}^r)$  satisfies a MDP on  $\mathbb{R}^2$  with speed  $\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}$  and the rate function  $J'(y) = \frac{\|y\|^2}{2}$ .

In particular,

$$\left\| \frac{|\mathbb{T}_r|}{b_{|\mathbb{T}_r|}}\hat{\Sigma}_r''^{-1/2}g(\hat{\theta}^r) \right\| = \frac{|\mathbb{T}_r|^{1/2}}{b_{|\mathbb{T}_r|}}\sqrt{\chi_r^{(1)}}$$

satisfies a MDP with speed  $\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}$  and the rate function  $I'$  given in the proposition 4.6.

Now, under the alternative hypothesis  $H_1$ ,

$$\frac{\chi_r^{(1)}}{|\mathbb{T}_r|} = g(\hat{\theta}^r)^t \hat{\Sigma}_r''^{-1} g(\hat{\theta}^r) \xrightarrow[\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}]{\text{superexp}} g(\theta)^t (\Sigma'')^{-1} g(\theta) > 0,$$

so that  $\chi_r^{(1)}$  converges  $\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}$ -superexponentially fast to  $+\infty$ . This conclude the proof of the proposition 4.6.  $\square$

## 4.2. The uniformly ergodic setting: compact case.

We recall that the model under study in this section is the model (4.1) where we assume that the noise values in a compact set. The results will be given without proofs, since the proofs are similar to those done in the previous section. The advantage of this section is that the range of speed is improved in comparison to the previous section. However, we suppose that the process takes its values in a compact set, which is not the case in the previous section.

We take  $F \in \mathcal{C}_b(\mathbb{R})$  and therefore (i)-(vi) and **(H2)** are automatically satisfied with

$\alpha = \max(|\alpha_0|, |\alpha_1|)$ . We use the same notations as in the previous section.

**Proposition 4.7.** *Let  $(b_n)$  a sequence of real numbers satisfying the Assumption 1. Then we have*

$$\hat{\theta}^r \xrightarrow[\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}]{\text{superexp}} \theta.$$

**Proposition 4.8.** *We have for all  $\delta > 0$ , and for  $r$  large enough*

$$\mathbb{P}\left(\left|\hat{\theta}^r - \theta\right| > \delta\right) \leq \begin{cases} c'' \exp(-c'|\mathbb{T}_r|) & \text{if } \alpha^2 < \frac{1}{2} \text{ and } \alpha < \frac{1}{2} \\ c'' \exp\left(-c' \left(\frac{1}{\alpha^2}\right)^{r+1}\right) & \text{if } \alpha^2 > \frac{1}{2} \\ c'' \exp(-c'|\mathbb{T}_r|) \exp(2c'\delta(r+1)) & \text{if } \alpha = \frac{1}{2} \\ c'' \exp\left(-c' \frac{|\mathbb{T}_r|}{r+1}\right) & \text{if } \alpha^2 = \frac{1}{2}, \end{cases} \quad (4.5)$$

where the constants  $c'$  and  $c''$  depends on  $\alpha$ ,  $\delta$ ,  $\mu_1$ ,  $\mu_2$ , and may differ line by line.

**Proposition 4.9.** *Let  $(b_n)$  a sequence of real numbers satisfying the Assumption 1. Then we have*

$$(\hat{\sigma}_r^2, \hat{\rho}_r) \xrightarrow[\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}]{\text{superexp}} (\sigma^2, \rho).$$

**Proposition 4.10.** *Let  $(b_n)$  a sequence of real numbers satisfying the the Assumption 1. Then  $\frac{|\mathbb{T}_r|}{b_{|\mathbb{T}_r|}} (\hat{\theta}^r - \theta)$  satisfies the MDP on  $\mathbb{R}^4$  with the speed  $\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}$  and rate function*

$$I(x) = \frac{1}{2} x^t (\Sigma')^{-1} x,$$

where

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

**Proposition 4.11.** *Let  $(b_n)$  a sequence of real numbers satisfying the Assumption 1. Then under the null hypothesis  $H_0 = \{(\alpha_0, \beta_0) = (\alpha_1, \beta_1)\}$ ,  $\frac{|\mathbb{T}_r|^{1/2}}{b_{|\mathbb{T}_r|}} (\chi_r^{(1)})^{1/2}$  satisfies a MDP on  $\mathbb{R}$  with speed  $\frac{b_{|\mathbb{T}_r|}^2}{|\mathbb{T}_r|}$  and the rate function*

$$I'(y) = \begin{cases} \frac{y^2}{2} & \text{if } y \in \mathbb{R}_+ \\ +\infty & \text{otherwise.} \end{cases}$$

Under the alternative hypothesis  $H_1$  of  $H_0$ , we have for all  $A > 0$

$$\limsup_{r \rightarrow \infty} \frac{|\mathbb{T}_r|}{b_{|\mathbb{T}_r|}^2} \log \mathbb{P}\left(\chi_r^{(1)} < A\right) = -\infty.$$

## 5. APPENDIX

Let us gather here for the convenience of the readers various Theorems useful to establish LIL, AFSCLT or MDP.

**Lemma 5.1.** *Let  $(E, d)$  a metric space. Let  $(Z_n)$  a sequence of random variables values in  $E$ ,  $(v_n)$  a rate and  $g : \mathcal{D}_E \subset E \rightarrow \mathbb{R}$ . If*

$$Z_n \xrightarrow[v_n]{\text{superexp}} Z,$$

then

$$g(Z_n) \xrightarrow[v_n]{\text{superexp}} g(Z).$$

*Proof.* For all  $\delta > 0$ , there exists (see e.g [17])  $\alpha_0(\delta) > 0$

$$\mathbb{P}\left(|g(Z_n) - g(Z)| > \delta\right) \leq \mathbb{P}\left(d(Z_n, Z) > \alpha_0(\delta)\right).$$

□

Let  $M = (M_n, \mathcal{H}_n, n \geq 0)$  be a centered square integrable martingale defined on a probability space  $(\Omega, \mathcal{H}, \mathbb{P})$  and  $(\langle M \rangle_n)$  its bracket. We recall some limit theorems for martingale used intensively in this paper.

We recall the following result due to W. F. Stout (Theorem 3 in [16]).

**Theorem 5.2.** *Let  $(M_n)$  such that  $M_0 = 0$ . If  $\langle M \rangle_n \rightarrow \infty$  a.s. and*

$$\sum_{n=1}^{\infty} \frac{2 \log \log \langle M \rangle_n}{K_n^2 \langle M \rangle_n} \mathbb{E} \left[ (M_n - M_{n-1})^2 \mathbf{1}_{\left\{ (M_n - M_{n-1})^2 > \frac{K_n^2 \langle M \rangle_n}{2 \log \log \langle M \rangle_n} \right\}} / \mathcal{H}_{n-1} \right] < \infty \quad \text{a.s.}$$

where  $K_n$  are  $\mathcal{H}_{n-1}$  measurable and  $K_n \rightarrow 0$  a.s., then  $\limsup \frac{M_n}{\sqrt{2 \langle M \rangle_n \log \log \langle M \rangle_n}} = 1$  a.s..

We recall the following result due to F. Chaabane (Corollary 2.2, see [4]).

**Theorem 5.3.** *Let  $(V_n)$  be a  $(\mathcal{H}_n)$ -predictable increasing such that*

$$\text{H-1 } V_n^{-2} \langle M \rangle_n \xrightarrow[n \rightarrow \infty]{} 1, \text{ a.s.}$$

$$\text{H-2 for all } \varepsilon > 0, \sum_{n \geq 1} V_n^{-2} \mathbb{E} \left[ (M_n - M_{n-1})^2 \mathbf{1}_{|M_n - M_{n-1}| > \varepsilon V_n} / \mathcal{H}_{n-1} \right] < \infty; \quad \text{a.s.}$$

$$\text{H-3 for some } a > 1, \sum_{n \geq 1} V_n^{-2a} \mathbb{E} \left[ (M_n - M_{n-1})^{2a} \mathbf{1}_{|M_n - M_{n-1}| \leq \varepsilon V_n} / \mathcal{H}_{n-1} \right] < \infty, \quad \text{a.s.}$$

Then  $M_n$  satisfies an ASFCLT, that is, for almost all  $\omega$ , the weighted random measures

$$W_N(\omega, \bullet) = (\log V_N^2)^{-1} \sum_{n=1}^N \left( 1 - \frac{V_n^2}{V_{n+1}^2} \right) \delta_{\{\psi_n(\omega) \in \bullet\}}$$

associated to continuous processes  $\Psi_n(\omega) = \{\Psi_n(\omega, t), 0 \leq t \leq 1\}$  defined by

$$\Psi_n(\omega, t) = V_n^{-1} \{M_k + (V_{k+1}^2 - V_k^2)^{-1} (tV_n^2 - V_k^2)(M_{k+1} - M_k)\},$$

when  $V_k^2 \leq tV_n^2 < V_{k+1}^2$ ,  $0 \leq k \leq n-1$ , weakly converge to the Wiener measure on  $\mathcal{C}([0, 1], \mathbb{R})$ .

Let us enunciate the following which correspond to the unidimensional case of theorem 1 in [9].

**Proposition 5.4.** *Let  $(b_n)$  a sequence satisfying*

$$b_n \text{ is increasing, } \frac{b_n}{\sqrt{n}} \rightarrow +\infty, \quad \frac{b_n}{n} \rightarrow 0,$$

such that  $c(n) := \frac{n}{b_n}$  is non-decreasing, and define the reciprocal function  $c^{-1}(t)$  by

$$c^{-1}(t) := \inf\{n \in \mathbb{N} : c(n) \geq t\}.$$

Under the following conditions:

(C1) *there exists  $Q \in \mathbb{R}_+^*$  such that  $\frac{\langle M \rangle_n}{n} \xrightarrow[\frac{b_n^2}{n}]{\text{superexp}} Q$ ;*

(C2)  $\limsup_{n \rightarrow +\infty} \frac{n}{b_n^2} \log \left( n \operatorname{ess\,sup}_{1 \leq k \leq c^{-1}(b(n+1))} \mathbb{P}(|M_k - M_{k-1}| > b_n/\mathcal{H}_{k-1}) \right) = -\infty$ ;

(C3) *for all  $a > 0$   $\frac{1}{n} \sum_{k=1}^n \mathbb{E} \left( |M_k - M_{k-1}|^2 \mathbf{1}_{\{|M_k - M_{k-1}| \geq a \frac{n}{b_n}\}} / \mathcal{H}_{k-1} \right) \xrightarrow[\frac{b_n^2}{n}]{\text{superexp}} 0$ ;*

$\left( \frac{M_n}{b_n} \right)_{n \in \mathbb{N}}$  *satisfies the MDP in  $\mathbb{R}$  with the speed  $\frac{b_n^2}{n}$  and the rate function  $I(x) = \frac{x^2}{2Q}$ .*

#### REFERENCES

- [1] Azuma, K. *Weighted sums of certain dependent random variables.* Tôhoku Math. J. (1967), Vol. 19, No. 3, pp. 357-367.
- [2] Bennett, G. *Probability inequalities for sum of independant random variables.* Journal of the American Statistical Association (Mar. 1962), Vol. 57, No. 297, pp. 33-45.
- [3] Bercu, B. De Saporta, B. and Gégout-Petit, A. *Asymptotic analysis for bifurcating autoregressive processes via a martingale approach.* Electronic. J. Probab. (2009), Vol. 14, pp. 2492-2526.
- [4] Chaabane, F. *Version forte du théorème de la limite centrale fonctionnel pour les martingales.* C. R. Acad. Sci. Paris Série 1, (1996), Vol. 323, pp. 195-198.
- [5] De Saporta, B., Gégout-Petit, A., Marsalle, L. *Parameters estimation for asymmetric bifurcating autoregressive processes with missing data.* Electronic Journal of Statistics (2011), Vol. 5, pp. 1313-1353.
- [6] Delmas, J.F. and Marsalle, L. *Detection of cellular aging in Galton-Watson process,* arxiv, (2008) 0807.0749.
- [7] Dembo, A. *Moderate deviations for martingales with bounded jumps,* Elect. Comm. Probab. (1996), Vol. 1, pp. 11-17.
- [8] Dembo, A. and Zeitouni, O. (1998) *Large Deviations Techniques and Applications,* 2nd Ed. (Springer, New York).
- [9] Djellout, H. *Moderate deviations for martingale differences and applications to  $\phi$ -mixing sequences.* Stochastics and stochastics reports, (2002), Vol. 73, No. 1-2, pp. 37-63.
- [10] Djellout, H. and Guillin, A. and Wu, L. (2004) *Transportation cost-information inequalities and applications to random dynamical systems and diffusions.* Annals of Probability, Vol. 32, No. 3B, 2702-2732.
- [11] Gao, F. and Zhao, X. *Delta method in large deviations and moderate deviations for estimators.* The Annals of Statistics, (2011), Vol. 39, No. 2, pp. 1211-1240.
- [12] Guyon, J. *Limit theorems for bifurcating markov chains. Application to the detection of cellular aging.* Ann. Appl. Probab., (2007), Vol. 17, No. 5-6, pp. 1538-1569.
- [13] Guyon, J. Bize, A. Paul, G. Stewart, E.J. Delmas, J.F. Taddéi, F. *Statistical study of cellular aging.* CEMRACS 2004 Proceedings, ESAIM Proceedings, (2005), 14, pp. 100-114.

- [14] Hoeffding, W. *Probability inequalities for sums of bounded random variables*. Journal of the American Statistical Association, Vol. 58, No. 301 (Mar. 1963), pp. 13-30.
- [15] E. J. Stewart, R. Madden, G. Paul and F. Taddéi. *Aging and death in an organism that reproduces by morphologically symmetric division*. PLoS Biol, 2005, 3(2): e45.
- [16] Stout, W. F. *A martingale analogue of kolmogorov's law of the iterated logarithm*. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete (1970), Vol. 15, pp. 279-290.
- [17] Van Der Vaart, A. W. (1998) *Asymptotic statistics*. New York. Cambridge University Press.
- [18] Worms, J. *Principes de déviations modérées pour des martingales et applications statistiques*, Thèse, Université de Marne-la-Vallée, 2000.
- [19] Worms, J. *Moderate deviations for stable Markov chains and regression models*. Electronic. J. Probab. (1999), Vol. 4, pp. 1-28.

**Valère BITSEKI PENDA**,, LABORATOIRE DE MATHÉMATIQUES, CNRS UMR 6620, UNIVERSITÉ BLAISE PASCAL, AVENUE DES LANDAIS 63177 AUBIÈRE.

*E-mail address:* Valere.Bitsekipenda@math.univ-bpclermont.fr

**Hacène DJELLOUT**,, LABORATOIRE DE MATHÉMATIQUES, CNRS UMR 6620, UNIVERSITÉ BLAISE PASCAL, AVENUE DES LANDAIS 63177 AUBIÈRE.

*E-mail address:* Hacene.Djellout@math.univ-bpclermont.fr

**Arnaud GUILLIN**,, INSTITUT UNIVERSITAIRE DE FRANCE ET LABORATOIRE DE MATHÉMATIQUES, CNRS UMR 6620, UNIVERSITÉ BLAISE PASCAL, AVENUE DES LANDAIS 63177 AUBIÈRE.

*E-mail address:* Arnaud.Guillin@math.univ-bpclermont.fr