



**HAL**  
open science

# MODERATE DEVIATION PRINCIPLES FOR KERNEL ESTIMATOR OF INVARIANT DENSITY IN BIFURCATING MARKOV CHAINS MODELS

Siméon Valère Bitseki Penda

► **To cite this version:**

Siméon Valère Bitseki Penda. MODERATE DEVIATION PRINCIPLES FOR KERNEL ESTIMATOR OF INVARIANT DENSITY IN BIFURCATING MARKOV CHAINS MODELS. 2021. hal-03331911

**HAL Id: hal-03331911**

**<https://hal.science/hal-03331911>**

Preprint submitted on 2 Sep 2021

**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.

# MODERATE DEVIATION PRINCIPLES FOR KERNEL ESTIMATOR OF INVARIANT DENSITY IN BIFURCATING MARKOV CHAINS MODELS.

S. VALÈRE BITSEKI PENDA

ABSTRACT. Bitseki and Delmas (2021) have studied recently the central limit theorem for kernel estimator of invariant density in bifurcating Markov chains models. We complete their work by proving a moderate deviation principle for this estimator. Unlike the work of Bitseki and Gorgui (2021), it is interesting to see that the distinction of the two regimes disappears and that we are able to get moderate deviation principle for large values of the ergodic rate. It is also interesting and surprising to see that for moderate deviation principle, the ergodic rate begins to have an impact on the choice of the bandwidth for values smaller than in the context of central limit theorem studied by Bitseki and Delmas (2021).

**Keywords:** Bifurcating Markov chains, bifurcating auto-regressive process, binary trees, density estimation.

**Mathematics Subject Classification (2020):** 62G05, 62F12, 60F10, 60J80,

## 1. INTRODUCTION

The study of bifurcating Markov chains (BMCs, for short) models has taken a special place in the literature these last years due to their links with the study of the cell dynamics (see for e.g. [6, 10, 13, 16, 17]). The first model of BMC, named “symmetric” bifurcating autoregressive process (BAR, for short) were introduced by Cowan and Staudte [9] in order to understand the cell division mechanisms of *Escherichia Coli* (*E. Coli*, for short). *E. Coli* is a rod shaped bacterium which reproduces by dividing in two, thus producing two new cells. One of type 1 which has the old end of the mother and the other of type 0 which has the new end of the mother. The age of a cell is thus given by the age of its old pole in the sense of the number of divisions from which this pole exists. This cell division mechanism raises several questions, among other that of the symmetry of the division. In order to give a rigorous answer to this question, Guyon [16] has developed and studied the theory of BMCs. We note that to the best of our knowledge, the term BMC appears for the first time in the work of [1]. In particular, Guyon has studied an extension of the model introduced by Cowan and Staudte, named “asymmetric” BAR. In the conclusion of his study, Guyon concludes that aging has an impact on cell reproduction. We note that an extension of the model proposed by Guyon, named nonlinear BAR (NBAR, for short) were studied by Bitseki and Olivier in [6]. Another question of interest related to cell division is estimating the division rate at which cells divide. This question has been tackled recently in the work of Doumic & *al.* [13] and Hoffman & Marguet [17]. In all the previous work, the behaviour and the definition of parameters of interest are associated with the density of the invariant probability of an auxiliary Markov chain (see below for a precise definition). The estimation of this invariant density has recently been the subject of several studies. One can cite [5, 8] where adaptive methods have been proposed for the estimation of this invariant density. More recently, Bitseki and Delmas [2] have studied central limit theorem for kernel estimators of this invariant density. Our main objective

in this paper is to complete the previous study by establishing a moderate deviation principle for these kernel estimators. Before going any further, let us recall the definition of the main concepts that we will use and study.

## 2. THE MODEL OF BIFURCATING MARKOV CHAIN AND DEFINITION OF THE ESTIMATORS

**2.1. The regular binary tree associated to BMC models.** We denote by  $\mathbb{N}$  (resp.  $\mathbb{N}^*$ ) the space of (resp. positive) natural integers. We set  $\mathbb{T}_0 = \mathbb{G}_0 = \{\emptyset\}$ ,  $\mathbb{G}_k = \{0, 1\}^k$  and  $\mathbb{T}_k = \bigcup_{0 \leq r \leq k} \mathbb{G}_r$  for  $k \in \mathbb{N}^*$ , and  $\mathbb{T} = \bigcup_{r \in \mathbb{N}} \mathbb{G}_r$ . The set  $\mathbb{G}_k$  corresponds to the  $k$ -th generation,  $\mathbb{T}_k$  to the tree up to the  $k$ -th generation, and  $\mathbb{T}$  the complete binary tree. One can see that the genealogy of the cells is entirely described by  $\mathbb{T}$  (each vertex of the tree designates an individual). For  $i \in \mathbb{T}$ , we denote by  $|i|$  the generation of  $i$  ( $|i| = k$  if and only if  $i \in \mathbb{G}_k$ ) and  $iA = \{ij; j \in A\}$  for  $A \subset \mathbb{T}$ , where  $ij$  is the concatenation of the two sequences  $i, j \in \mathbb{T}$ , with the convention that  $\emptyset i = i\emptyset = i$ . For  $A \subset \mathbb{T}$ , we denote by  $|A|$  the number of elements of  $A$ . Note that for all  $n \in \mathbb{N}$ ,  $|\mathbb{G}_n| = 2^n$  and  $|\mathbb{T}_n| = 2^{n+1} - 1$ .

## 2.2. The probability kernels associated to BMC models.

For our convenience, we set  $S = \mathbb{R}^d$ ,  $d \geq 1$  and  $S$  is equipped with the Borel sigma-algebra  $\mathcal{S}$ . For any  $q \in \mathbb{N}^*$ , we denote by  $\mathcal{B}(S^q)$  (resp.  $\mathcal{B}_b(S^q)$ , resp.  $\mathcal{C}_b(S^q)$ ) the space of (resp. bounded, resp. bounded continuous)  $\mathbb{R}$ -valued measurable functions defined on  $S^q$ . For all  $q \in \mathbb{N}^*$ , we set  $\mathcal{S}^{\otimes q} = \mathcal{S} \otimes \dots \otimes \mathcal{S}$ . Let  $\mathcal{P}$  be a probability kernel on  $(S, \mathcal{S}^{\otimes 2})$ , that is:  $\mathcal{P}(\cdot, A)$  is measurable for all  $A \in \mathcal{S}^{\otimes 2}$ , and  $\mathcal{P}(x, \cdot)$  is a probability measure on  $(S^2, \mathcal{S}^{\otimes 2})$  for all  $x \in S$ . For any  $g \in \mathcal{B}_b(S^3)$  and  $h \in \mathcal{B}_b(S^2)$ , we set for  $x \in S$ :

$$(1) \quad (\mathcal{P}g)(x) = \int_{S^2} g(x, y, z) \mathcal{P}(x, dy, dz) \quad \text{and} \quad (\mathcal{P}h)(x) = \int_{S^2} h(y, z) \mathcal{P}(x, dy, dz).$$

We define  $(\mathcal{P}g)$  (resp.  $(\mathcal{P}h)$ ), or simply  $Pg$  for  $g \in \mathcal{B}(S^3)$  (resp.  $\mathcal{P}h$  for  $h \in \mathcal{B}(S^2)$ ), as soon as the corresponding integral (1) is well defined, and we have that  $\mathcal{P}g$  and  $\mathcal{P}h$  belong to  $\mathcal{B}(S)$ . We denote by  $\mathcal{P}_0, \mathcal{P}_1$  and  $\mathcal{Q}$  respectively the first and the second marginal of  $\mathcal{P}$ , and the mean of  $\mathcal{P}_0$  and  $\mathcal{P}_1$ , that is, for all  $x \in S$  and  $B \in \mathcal{S}$

$$\mathcal{P}_0(x, B) = \mathcal{P}(x, B \times S), \quad \mathcal{P}_1(x, B) = \mathcal{P}(x, S \times B) \quad \text{and} \quad \mathcal{Q} = \frac{(\mathcal{P}_0 + \mathcal{P}_1)}{2}.$$

Now let us give a precise definition of bifurcating Markov chain.

**Definition 2.1** (Bifurcating Markov Chains, see [16, 2]).

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

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

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

Following [16], we introduce an auxiliary Markov chain  $Y = (Y_n, n \in \mathbb{N})$  on  $(S, \mathcal{S})$  with  $Y_0 = X_1$  and transition probability  $\mathcal{Q}$ . The chain  $(Y_n, n \in \mathbb{N})$  corresponds to a random lineage taken in the population. We shall write  $\mathbb{E}_x$  when  $X_\emptyset = x$  (i.e. the initial distribution  $\nu$  is the Dirac mass at  $x \in S$ ). We will assume that the Markov chain  $Y$  is ergodic and we denote by  $\mu$  its invariant

probability measure. Asymptotic and non-asymptotic behaviour of BMCs are strongly related to the knowledge of  $\mu$ . In particular, Guyon has proved that if  $Y$  is ergodic, then for all  $f \in \mathcal{C}_b(S)$ ,

$$|\mathbb{A}_n|^{-1} \sum_{u \in \mathbb{A}_n} f(X_u) \xrightarrow[n \rightarrow +\infty]{} \langle \mu, f \rangle \quad \text{in probability, where } \mathbb{A}_n \in \{\mathbb{G}_n, \mathbb{T}_n\}.$$

But in most cases, the invariant probability  $\mu$  is unknown, so its estimation from the data is of great interest. For that purpose, we do the following assumption.

**Assumption 2.2.** *The transition kernel  $\mathcal{P}$  has a density, still denoted by  $\mathcal{P}$ , with respect to the Lebesgue measure.*

*Remark 2.3.* Assumption 2.2 implies that the transition kernel  $\mathcal{Q}$  has a density, still denoted by  $\mathcal{Q}$ , with respect to the Lebesgue measure. More precisely, we have  $\mathcal{Q}(x, y) = 2^{-1} \int_S (\mathcal{P}(x, y, z) + \mathcal{P}(x, z, y)) dz$ . This implies in particular that the invariant probability  $\mu$  has a density, still denoted by  $\mu$ , with respect to the Lebesgue measure (for more details, we refer for e.g. to [14], chap 6).

**2.3. Kernel estimator of the invariant density  $\mu$ .** Recall that  $\mathbb{A}_n \in \{\mathbb{G}_n, \mathbb{T}_n\}$  and  $S = \mathbb{R}^d$ ,  $d \geq 1$ . Assume we observe  $\mathbb{X}^n = (X_u, u \in \mathbb{A}_n)$ . Let  $(h_n, n \in \mathbb{N})$  be a sequence of positive numbers which converges to 0 as  $n$  goes to infinity. We will simply write  $h$  for  $h_n$  if there is no ambiguity. Let the kernel function  $K : S \rightarrow \mathbb{R}$  such that  $\int_S K(x) dx = 1$ . Then, for all  $x \in S$ , we propose to estimate  $\mu(x)$  by

$$(2) \quad \hat{\mu}_{\mathbb{A}_n}(x) = |\mathbb{A}_n|^{-1} h_n^{-d/2} \sum_{u \in \mathbb{A}_n} K_{h_n}(x - X_u),$$

where  $K_{h_n}(\cdot) = h_n^{-d/2} K(h_n \cdot)$ . These estimators are strongly inspired from [18, 21, 22]. They have been studied in [13, 8] (non asymptotic studies) and in [2] (central limit theorem).

**2.4. Moderate deviation principle and related topics.** Our aim is to study moderate deviation principles for the estimators defined in (2). Before we proceed, let us introduce the notion of moderate deviation principle. We give the definition in a general setting. Let  $(Z_n)_{n \geq 0}$  be a sequence of random variables with values in  $S$  endowed with its Borel  $\sigma$ -field  $\mathcal{S}$  and let  $(s_n)_{n \geq 0}$  be a positive sequence that converges to  $+\infty$ . We assume that  $Z_n/s_n$  converges in probability to 0 and that  $Z_n/\sqrt{s_n}$  converges in distribution to a centered Gaussian law. Let  $I : S \rightarrow \mathbb{R}^+$  be a lower semicontinuous function, that is for all  $c > 0$  the sub-level set  $\{x \in S, I(x) \leq c\}$  is a closed set. Such a function  $I$  is called *rate function* and it is called *good rate function* if all its sub-level sets are compact sets. Let  $(b_n)_{n \geq 0}$  be a positive sequence such that  $b_n \rightarrow +\infty$  and  $b_n/\sqrt{s_n} \rightarrow 0$  as  $n$  goes to  $+\infty$ .

**Definition 2.4** (Moderate deviation principle, MDP).

*We say that  $Z_n/(b_n \sqrt{s_n})$  satisfies a moderate deviation principle on  $S$  with speed  $b_n^2$  and rate function  $I$  if, for any  $A \in \mathcal{S}$ ,*

$$-\inf_{x \in \overset{\circ}{A}} I(x) \leq \liminf_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{Z_n}{b_n \sqrt{s_n}} \in A\right) \leq \limsup_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{Z_n}{b_n \sqrt{s_n}} \in A\right) \leq -\inf_{x \in \bar{A}} I(x),$$

where  $\overset{\circ}{A}$  and  $\bar{A}$  denote respectively the interior and the closure of  $A$ .

The following two concepts are closely related to the theory of MDP: super-exponential convergence and exponential equivalence. Let  $(Z_n, n \in \mathbb{N})$ ,  $(W_n, n \in \mathbb{N})$  be sequences of random variables and  $Z$  a random variable with value in a metric space  $(S, d)$ .

**Definition 2.5** (Super-exponential convergence). *We say that  $(Z_n)_{n \geq 0}$  converges  $(b_n^2)$ -super-exponentially fast in probability to  $Z$  and we note  $Z_n \xrightarrow[b_n^2]{\text{superexp}} Z$  if, for all  $\delta > 0$ ,*

$$\limsup_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \mathbb{P}(d(Z_n, Z) > \delta) = -\infty.$$

**Definition 2.6** (Exponential equivalence, see [11], Chap 4). *We say that  $(Z_n)_{n \geq 0}$  and  $(W_n)_{n \geq 0}$  are  $(b_n^2)_{n \geq 0}$ -exponentially equivalent and we note  $Z_n \xrightarrow[b_n^2]{\text{superexp}} W_n$  if for any  $\delta > 0$ ,*

$$\limsup_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \mathbb{P}(d(Z_n, W_n) > \delta) = -\infty.$$

*Remark 2.7.* Note that for a deterministic sequence that converges to some limit  $\ell$ , it also converges  $(b_n^2)$ -superexponentially fast to  $\ell$  for any rate  $b_n$ . We also note that if  $(Z_n)_{n \geq 0}$  and  $(W_n)_{n \geq 0}$  are  $(b_n^2)_{n \geq 0}$ -exponentially equivalent and if  $(Z_n)_{n \geq 0}$  satisfies a MDP, then  $(W_n)_{n \geq 0}$  satisfies the same MDP (for more details, see for e.g [11], Chap 4).

The following result give a sufficient condition for super-exponential convergence of a sequence of random variables.

*Remark 2.8.* We assume that  $(S, d)$  is a metric space. Let  $(Z_n)_{n \in \mathbb{N}}$  be a sequence of random variables with values in  $S$ ,  $Z$  a random variable with values in  $S$ . So if  $d(Z_n, Z)$  is upper-bounded by a deterministic sequence which converges to 0, then, for all sequence  $(b_n, n \in \mathbb{N})$  converging to  $+\infty$ ,  $Z_n \xrightarrow[b_n^2]{\text{superexp}} Z$ .

The moderate deviation principle has been proved in the i.i.d. setting for kernel density estimator, see for e.g. Gao [15], Mokkadem & al. [20]. We refer also to [19] where Mokkadem and Pelletier have constructed confidence bands for probability densities based on moderate deviation principles. In this paper, we will establish moderate deviation principle for  $\hat{\mu}_{\mathbb{A}_n}(x)$  following the martingale approach developed in [2]. We will need the following assumption.

**Assumption 2.9.** *There exists a positive real number  $M$  and  $\alpha \in (0, 1)$  such that for all  $f \in \mathcal{B}_b(S)$ :*

$$(3) \quad |\mathcal{Q}^n f - \langle \mu, f \rangle| \leq M \alpha^n \|f\|_\infty \quad \text{for all } n \in \mathbb{N}.$$

*Remark 2.10.* Assumption 2.9 is for example satisfy for nonlinear bifurcating autoregressive process under mild hypotheses on the autoregression functions (see [7] Lemma 9 for more details).

The others assumptions we will need are based on the following bias-variance type decomposition of the estimator  $\hat{\mu}_{\mathbb{A}_n}(x)$ :

$$(4) \quad \hat{\mu}_{\mathbb{A}_n}(x) - \mu(x) = B_{h_n}(x) + V_{\mathbb{A}_n, h_n}(x),$$

where for  $h > 0$  and  $\mathbb{A} \subset \mathbb{T}$  finite:

$$B_h(x) = h^{-d/2} K_h \star \mu(x) - \mu(x) \quad \text{and} \quad V_{\mathbb{A}, h}(x) = |\mathbb{A}|^{-1} h^{-d/2} \sum_{u \in \mathbb{A}} \left( K_h(x - X_u) - K_h \star \mu(x) \right),$$

and for  $h > 0$  and  $u \in \mathbb{T}$ , we set:

$$K_h \star \mu(x) = \mathbb{E}_\mu[K_h(x - X_u)] = \int_S K_h(x - y) \mu(y) dy.$$

To study the variance term  $V_{\mathbb{A}_n, h_n}(x)$ , we will introduce a more general sequence of functions (see Section 3.2).

The following assumptions on the kernel, the bandwidth and the regularity of the unknown density function are usual. Recall  $S = \mathbb{R}^d$  with  $d \geq 1$ .

**Assumption 2.11** (Regularity of the kernel function and the bandwidth).

(i) The kernel function  $K \in \mathcal{B}(S)$  satisfies:

$$\|K\|_\infty < +\infty, \|K\|_1 < +\infty, \|K\|_2 < +\infty, \int_S K(x) dx = 1 \quad \text{and} \quad \lim_{|x| \rightarrow +\infty} |x|K(x) = 0.$$

(ii) There exists  $\gamma \in (0, 1/d)$  such that the bandwidth  $(h_n, n \in \mathbb{N})$  are defined by  $h_n = 2^{-n\gamma}$ .

**Assumption 2.12** (Further regularity on the density  $\mu$ , the kernel function and the bandwidths). Suppose that there exists an invariant probability measure  $\mu$  of  $\mathcal{Q}$  and that Assumptions 2.2 and 2.11 hold. We assume there exists  $s > 0$  such that the following hold:

(i) **The density  $\mu$  belongs to the (isotropic) Hölder class of order  $(s, \dots, s) \in \mathbb{R}^d$ :** The density  $\mu$  admits partial derivatives with respect to  $x_j$ , for all  $j \in \{1, \dots, d\}$ , up to the order  $\lfloor s \rfloor$  and there exists a finite constant  $L > 0$  such that for all  $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ ,  $t \in \mathbb{R}$  and  $j \in \{1, \dots, d\}$ :

$$\left| \frac{\partial^{\lfloor s \rfloor} \mu}{\partial x_j^{\lfloor s \rfloor}}(x_{-j}, t) - \frac{\partial^{\lfloor s \rfloor} \mu}{\partial x_j^{\lfloor s \rfloor}}(x) \right| \leq L|x_j - t|^{\lfloor s \rfloor},$$

where  $(x_{-j}, t)$  denotes the vector  $x$  where we have replaced the  $j^{\text{th}}$  coordinate  $x_j$  by  $t$ , with the convention  $\partial^0 \mu / \partial x_j^0 = \mu$ .

(ii) **The kernel  $K$  is of order  $(\lfloor s \rfloor, \dots, \lfloor s \rfloor) \in \mathbb{N}^d$ :** We have  $\int_{\mathbb{R}^d} |x|^s K(x) dx < \infty$  and  $\int_{\mathbb{R}} x_j^k K(x) dx_j = 0$  for all  $k \in \{1, \dots, \lfloor s \rfloor\}$  and  $j \in \{1, \dots, d\}$ .

For  $\alpha > 1/2$ , we shall also assume the following.

**Assumption 2.13.** Keeping the same notations as in (ii) of Assumption 2.11, we further assume that Assumption 2.9 holds with

$$(5) \quad \lim_{n \rightarrow +\infty} (2^{1-d\gamma} \alpha)^n = 0.$$

*Remark 2.14.* As consequence of Assumption 2.13 and (ii) of Assumption 2.11, for moderate deviation principle, the ergodicity rate  $\alpha$  begins to have an impact on the choice of the bandwidth for  $\alpha > 1/2$ . This is out of step with the central limit theorem where the ergodicity rate  $\alpha$  begins to have an impact on the choice of the bandwidth for  $\alpha > 1/\sqrt{2}$  (see [2] for more details).

In the sequel, we will consider the positive sequence  $(b_n, n \in \mathbb{N})$  such that:

$$(6) \quad \lim_{n \rightarrow +\infty} b_n = +\infty; \quad \lim_{n \rightarrow +\infty} \frac{n^{3/2} b_n}{\sqrt{|\mathbb{G}_n| h_n^d}} = 0; \quad \lim_{n \rightarrow +\infty} \frac{b_n}{\sqrt{|\mathbb{G}_n| h_n^{2s+d}}} = +\infty,$$

where  $s$  is the regularity parameter given in Assumption 2.12.

The paper is organised as follows. In Section 3.1 we state the main result for the moderate deviation principles of the estimators  $\widehat{\mu}_{\mathbb{A}_n}(x)$  for  $x$  in the set continuity of  $\mu$  and  $\mathbb{A}_n \in \{\mathbb{T}_n, \mathbb{G}_n\}$ . In Section 3.2, directly linked to the study of variance term  $V_{\mathbb{A},h}(x)$  defined in (4), we study the moderate deviation principle for general additive functionals of BMCs. Sections 4 and 5 are devoted to the proofs of results. In Section 6, we recall some useful results.

### 3. MAIN RESULT

**3.1. Moderate deviation principle for  $\widehat{\mu}_{\mathbb{A}_n}$ .** First, we state a strong consistency result for the estimators  $\widehat{\mu}_{\mathbb{A}_n}(x)$  for  $x$  in the set of continuity of  $\mu$ . Its proof is given in Section 4.1.

**Lemma 3.1.** *Let  $X$  be a BMC with kernel  $\mathcal{P}$  and initial distribution  $\nu$  such that Assumptions 2.9, 2.11 and 2.12 hold. Furthermore, if  $\alpha > 1/2$  then assume that Assumption 2.13 holds. Let  $(b_n, n \in \mathbb{N})$  be a positive sequence with satisfies (6). Then, for all  $x$  in the set of continuity of  $\mu$  and  $\mathbb{A}_n \in \{\mathbb{T}_n, \mathbb{G}_n\}$  we have  $\widehat{\mu}_{\mathbb{A}_n}(x) \xrightarrow[b_n^2]{\text{superexp}} \mu(x)$ .*

The main result of this Section is the following theorem which state the moderate deviation principle for  $\widehat{\mu}_{\mathbb{A}_n}(x) - \mu(x)$  for  $x$  in the set of continuity of the function  $\mu$ .

**Theorem 3.2.** *Under the hypothesis of Lemma 3.1, for all  $x$  in the set of continuity of  $\mu$  and  $\mathbb{A}_n \in \{\mathbb{T}_n, \mathbb{G}_n\}$ ,  $b_n^{-1} \sqrt{|\mathbb{A}_n| h_n^d} (\widehat{\mu}_{\mathbb{A}_n}(x) - \mu(x))$  satisfies a moderate deviation principle on  $\mathbb{R}$  with speed  $b_n^2$  and rate function  $I$  defined by:  $I(y) = y^2 / (2 \|K\|_2^2 \mu(x))$  for all  $y \in \mathbb{R}$ , that is, for any  $A \subset \mathbb{R}$ ,*

$$\begin{aligned} - \inf_{y \in \overset{\circ}{A}} I(y) &\leq \liminf_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \mathbb{P}(b_n^{-1} \sqrt{|\mathbb{A}_n| h_n^d} (\widehat{\mu}_{\mathbb{A}_n}(x) - \mu(x)) \in A) \\ &\leq \limsup_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \mathbb{P}(b_n^{-1} \sqrt{|\mathbb{A}_n| h_n^d} (\widehat{\mu}_{\mathbb{A}_n}(x) - \mu(x)) \in A) \leq - \inf_{y \in \bar{A}} I(y), \end{aligned}$$

where  $\overset{\circ}{A}$  and  $\bar{A}$  denote respectively the interior and the closure of  $A$ .

In order to obtain confidence intervals for  $\mu(x)$ , it would be interesting to replace  $\mu(x)$  in the expression of the rate function  $I(\cdot)$  by an estimator. In that direction, we have the following. Let  $\mathbb{A}_n^* \in \{\mathbb{G}_n, \mathbb{T}_n\}$ . Obviously,  $\mathbb{A}_n^*$  and  $\mathbb{A}_n$  can be the same. We consider the estimator  $\widehat{\mu}_{\mathbb{A}_n^*}(x)$  of  $\mu(x)$  defined with  $\mathbb{A}_n^*$  instead of  $\mathbb{A}_n$ . Let  $(\varpi_n, n \in \mathbb{N})$  be a sequence of real numbers such that  $\varpi_n \rightarrow 0$  as  $n \rightarrow +\infty$ . Then, we have the following result which the proof is given in Section 4.3.

**Theorem 3.3.** *Under the hypothesis of Lemma 3.1, for all  $x$  in the set of continuity of  $\mu$  and  $\mathbb{A}_n, \mathbb{A}_n^* \in \{\mathbb{T}_n, \mathbb{G}_n\}$ ,  $b_n^{-1} (\|K\|_2 \sqrt{\widehat{\mu}_{\mathbb{A}_n^*}(x)} \vee \varpi_n)^{-1} \sqrt{|\mathbb{A}_n| h_n^d} (\widehat{\mu}_{\mathbb{A}_n}(x) - \mu(x))$  satisfies a moderate deviation principle on  $\mathbb{R}$  with speed  $b_n^2$  and rate function  $I'$  defined by:  $I'(y) = y^2 / 2$  for all  $y \in \mathbb{R}$ .*

In particular, using the contraction principle (see for e.g Dembo and Zeitouni [11], Chap 4), we have the following corollary of Theorem 3.3.

**Corollary 3.4.** *Under the hypothesis of Theorem 3.3, we have the following convergence for  $x$  in the set of continuity of  $\mu$  and  $\mathbb{A}_n, \mathbb{A}_n^* \in \{\mathbb{T}_n, \mathbb{G}_n\}$ :*

$$\lim_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(b_n^{-1} \left( \|K\|_2 \sqrt{\widehat{\mu}_{\mathbb{A}_n^*}(x)} \vee \varpi_n \right)^{-1} \sqrt{|\mathbb{A}_n| h_n^d} \left| \widehat{\mu}_{\mathbb{A}_n}(x) - \mu(x) \right| > \delta\right) = -\frac{\delta^2}{2} \quad \forall \delta > 0.$$

*Remark 3.5.* Corollary 3.4 yields a simple confidence interval for  $\mu(x)$ , of decreasing size  $b_n / \sqrt{|\mathbb{A}_n| h_n^d}$  and with level asymptotically close to  $1 - \exp(-b_n^2 \delta^2 / 2)$ .

Using the structure of the asymptotic variance  $\sigma^2$  in (7), we can prove the following multidimensional result which the proof is given in Section 4.4

**Corollary 3.6.** *Under the hypothesis of Theorem 3.2, we have, for  $x$  in the set of continuity of  $\mu$  and for all  $k \geq 0$ ,  $b_n^{-1} \left( |\mathbb{G}_n|^{1/2} h_n^{1/2} (\widehat{\mu}_{\mathbb{G}_n}(x) - \mu(x)), \dots, |\mathbb{G}_{n-k}|^{1/2} h_{n-k}^{1/2} (\widehat{\mu}_{\mathbb{G}_{n-k}}(x) - \mu(x)) \right)^t$  satisfies a moderate deviation principle on  $\mathbb{R}^{k+1}$  with speed  $b_n^2$  and good rate function  $J_x : \mathbb{R}^{k+1} \rightarrow \mathbb{R}$  defined by*

$$J_x(\mathbf{z}) = (2 \|K\|_2^2 \mu(x))^{-1} \mathbf{z}^t \Gamma^{-1} \mathbf{z}, \quad \mathbf{z} \in \mathbb{R}^{k+1},$$

with  $\Gamma = \text{diag}(2^0, \dots, 2^k)$ , where  $\text{diag}(\cdot)$  denotes the diagonal matrix and  $\mathbf{z}^t$  stands for the transpose of vector  $\mathbf{z}$ .

*Remark 3.7.* We deduce from Corollary 3.6 that the estimators  $|\mathbb{G}_{n-\ell}|^{1/2} h_{n-\ell}^{d/2} (\widehat{\mu}_{\mathbb{G}_{n-\ell}}(x) - \mu(x))$  are asymptotically independent in the sense of moderate deviation for  $\ell \in \{0, \dots, k\}$  and for any  $k \in \mathbb{N}$ .

### 3.2. Moderate deviation principle for additive functionals of BMCs.

In order to study the variance term  $V_{\mathbb{A}_n, h_n}(x)$ , we give here a moderate deviation principle for a general additive functionals of BMCs. For that purpose, we introduce the following assumption.

**Assumption 3.8.** For  $n \in \mathbb{N}$ , let  $\mathfrak{f}_n = (f_{\ell, n}, n \geq \ell \geq 0)$  be a sequence of functions defined on  $S$  such that  $f_{\ell, n} = 0$  if  $\ell > n$  and there exists  $\gamma \in (0, 1/d)$  such that:

- (i)  $\sup_{0 \leq \ell \leq n} \{2^{-d\gamma n/2} \|f_{\ell, n}\|_\infty; 2^{d\gamma n/2} \|\mathcal{Q}f_{\ell, n}\|_\infty; \|\mathcal{Q}(f_{\ell, n}^2)\|_\infty; 2^{d\gamma n} \|\mathcal{P}(f_{\ell, n} \otimes^2)\|_\infty\} < +\infty.$
- (ii)  $\sup_{0 \leq \ell \leq n} \{2^{d\gamma n/2} \langle \mu, |f_{\ell, n}| \rangle; \langle \mu, f_{\ell, n}^2 \rangle\} < +\infty.$
- (iii) The following limit exists and is finite:

$$(7) \quad \sigma^2 = \lim_{n \rightarrow +\infty} \sum_{\ell=0}^n 2^{-\ell} \|f_{\ell, n}\|_{L^2(\mu)}^2 < +\infty.$$

We will use the following notations. For a finite set  $\mathbb{A} \subset \mathbb{T}$  and a function  $f \in \mathcal{B}(S)$ , we set:

$$M_{\mathbb{A}}(f) = \sum_{i \in \mathbb{A}} f(X_i).$$

In this paper, we are interested in the cases  $\mathbb{A} = \mathbb{G}_n$  and  $\mathbb{A} = \mathbb{T}_n$ , that is the  $n$ -th generation and the first  $n$  generation of the tree. Recall  $\mu$  the invariant probability of  $\mathcal{Q}$ , transition probability of the auxiliary Markov chain  $(Y_n, n \in \mathbb{N})$ . For  $f \in L^1(\mu)$ , we set:

$$\tilde{f} = f - \langle \mu, f \rangle.$$

Recall the sequence  $\mathfrak{f}_n$  defined in Assumption 3.8. For  $n \in \mathbb{N}$ , we set:

$$(8) \quad N_{n, \emptyset}(\mathfrak{f}_n) = |\mathbb{G}_n|^{-1/2} \sum_{\ell=0}^n M_{\mathbb{G}_{n-\ell}}(\tilde{f}_{\ell, n}).$$

The notation  $N_{n, \emptyset}$  means that we consider the average from the root  $\emptyset$  to the  $n$ -th generation.

*Remark 3.9.* The definition of  $N_{n, \emptyset}(\mathfrak{f}_n)$  in (8) is mainly motivated by the decomposition (4). It will allow us to treat the variance term of the estimator  $\widehat{\mu}_{\mathbb{A}_n}(x)$  defined in (2). Instead, for  $n \in \mathbb{N}$ , we set  $f_n^x(\cdot) = K_{h_n}(x - \cdot)$ . Then, we consider the sequences of functions  $(f_{\ell, n}^{\text{id}}, n \geq \ell \geq 0)$  and  $(f_{\ell, n}^0, n \geq \ell \geq 0)$  defined by:

$$(9) \quad f_{\ell, n}^{\text{id}} = f_n^x \quad \text{and} \quad f_{\ell, n}^0 = f_n^x \mathbf{1}_{\{\ell=0\}}.$$

It is not difficult to check that under Assumption 2.11, the sequence  $(f_{\ell, n}^{\text{id}}, n \geq \ell \geq 0)$  and  $(f_{\ell, n}^0, n \geq \ell \geq 0)$  defined in (9) satisfy Assumption 3.8. In particular, let  $x$  be in the set of continuity of  $\mu$ . Thanks to Lemma 6.3, we have:

$$(10) \quad \lim_{n \rightarrow +\infty} \|f_n^x\|_{L^2(\mu)}^2 = \lim_{n \rightarrow +\infty} \langle \mu, (f_n^x)^2 \rangle = \mu(x) \|K\|_2^2.$$

If  $\mathbb{A}_n = \mathbb{G}_n$ , it suffices to consider the sequence  $\mathfrak{f}_n = (f_{\ell, n}, 0 \leq \ell \leq n)$  with  $f_{\ell, n} = f_{\ell, n}^0$  and in that case, using (10), the asymptotic variance defined in (7) is given by  $\sigma^2 = \|K\|_2^2 \mu(x)$ . If  $\mathbb{A}_n = \mathbb{T}_n$ , it suffices to consider the sequence  $\mathfrak{f}_n = (f_{\ell, n}, 0 \leq \ell \leq n)$  with  $f_{\ell, n} = f_{\ell, n}^{\text{id}}$  and in that case, using (10), the asymptotic variance defined in (7) is given by  $\sigma^2 = 2\|K\|_2^2 \mu(x)$ .

For our convenience, we assume that the quantity  $\gamma$  which appears in Assumptions 2.11 and 3.8 is the same. The main result of this section is the following.



**Theorem 3.10.** *Let  $X$  be a BMC with kernel  $\mathcal{P}$  and initial distribution  $\nu$  such that Assumptions 2.9, 2.11 and 3.8 hold. Furthermore, if  $\alpha > 1/2$  then assume that Assumption 2.13 holds. Let  $(b_n, n \in \mathbb{N})$  be a positive sequence with satisfies (6). Then  $b_n^{-1}N_{n,\emptyset}(\mathfrak{f}_n)$  satisfies a moderate deviation principle on  $\mathbb{R}$  with speed  $b_n^2$  and rate function  $I$  defined by:  $I(x) = x^2/(2\sigma^2)$  for all  $x \in \mathbb{R}$ , with the finite variance  $\sigma^2$  defined in (7).*

*Remark 3.11.* In particular, using the contraction principle (see for e.g Dembo and Zeitouni [11], Chap 4), Theorem 3.10 implies that

$$\lim_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \mathbb{P}(|b_n^{-1}N_{n,\emptyset}(\mathfrak{f}_n)| > \delta) = -I(\delta) \quad \forall \delta > 0.$$

*Remark 3.12.* Unlike the results of Bitseki and Gorgui [4], one can note that the different regimes disappear in Theorem 3.10. Moreover, we are able here to give the fluctuations if  $2\alpha^2 > 1$  which is not the case in [4].

#### 4. PROOF OF LEMMA 3.1, THEOREMS 3.2 AND 3.3 AND COROLLARY 3.6

We will denote by  $C$  any unimportant finite constant which may vary from line to line (in particular  $C$  does not depend on  $n \in \mathbb{N}$ ).

**4.1. Proof of Lemma 3.1.** We begin the proof with  $\mathbb{A}_n = \mathbb{T}_n$ . Recall the decomposition (4) with  $\mathbb{T}_n$  instead of  $\mathbb{A}$ . Using Lemma 6.3, we have  $\lim_{n \rightarrow +\infty} |B_{h_n}(x)| = 0$ . From Remark 2.7, this implies that  $B_{h_n}(x) \xrightarrow[b_n^2]{\text{superexp}} 0$ . Next, we set  $f_n(\cdot) = K_{h_n}(x - \cdot)$  in such a way that we have

$$|\mathbb{T}_n|^{-1} h_n^{d-/2} \sum_{u \in \mathbb{T}_n} \left( K_h(x - X_u) - K_h \star \mu(x) \right) = |\mathbb{T}_n|^{-1} h_n^{d-/2} \sum_{\ell=0}^n M_{\mathbb{G}_\ell}(\tilde{f}_n).$$

Following line by line the proof of (32) (where we take  $f_{\ell,n} = f_n$  for all  $\ell \leq n$ ), we get

$$\mathbb{P}\left(|\mathbb{T}_n|^{-1} h_n^{d-/2} \left| \sum_{\ell=0}^n M_{\mathbb{G}_\ell}(\tilde{f}_n) \right| > \delta\right) \leq 2 \exp\left(-\frac{3\delta}{c_1 + c_2\delta}\right) \exp\left(-\frac{3\delta^2 |\mathbb{T}_n| h_n^d}{c_1 + c_2\delta}\right).$$

Taking the log, dividing by  $b_n^2$  and letting  $n$  goes to the infinity in the latter inequality, we get

$$|\mathbb{T}_n|^{-1} h_n^{d-/2} \sum_{u \in \mathbb{T}_n} \left( K_h(x - X_u) - K_h \star \mu(x) \right) \xrightarrow[b_n^2]{\text{superexp}} 0.$$

It then follows from the decomposition (4) that  $\hat{\mu}_{\mathbb{T}_n}(x) \xrightarrow[b_n^2]{\text{superexp}} \mu(x)$ . We similarly get the result for  $\mathbb{A}_n = \mathbb{G}_n$  and this ends the proof of the lemma.

**4.2. Proof of Theorem 3.2.** We begin the proof with  $\mathbb{A}_n = \mathbb{T}_n$ . We have the following decomposition:

$$b_n^{-1} \sqrt{|\mathbb{T}_n| h_n^d} (\hat{\mu}_{\mathbb{T}_n}(x) - \mu(x)) = \sqrt{\frac{|\mathbb{G}_n|}{|\mathbb{T}_n|}} b_n^{-1} N_{n,\emptyset}(\mathfrak{f}_n) + \frac{\sqrt{|\mathbb{T}_n| h_n^d}}{b_n} B_{h_n}(x),$$

where  $\mathfrak{f}_n = (f_{\ell,n}, n \geq \ell \geq 0)$  with the functions  $f_{\ell,n} = f_{\ell,n}^{\text{id}}$  defined in (9) for  $n \geq \ell \geq 0$  and  $f_{\ell,n} = 0$  otherwise;  $N_{n,\emptyset}(\mathfrak{f}_n)$  is defined in (8) and the bias term  $B_{h_n}(x)$  is defined in (4). Thanks to Theorem 3.10 applied to the sequence  $(f_{\ell,n}^{\text{id}}, n \geq \ell \geq 0)$  and using that  $\lim_{n \rightarrow +\infty} |\mathbb{G}_n|/|\mathbb{T}_n| = 1/2$ , we get that  $\sqrt{|\mathbb{G}_n|/|\mathbb{T}_n|} b_n^{-1} N_{n,\emptyset}(\mathfrak{f}_n)$  satisfies a moderate deviation principle in  $\mathbb{R}$  with speed  $b_n^2$

and rate function  $I$  defined by:  $I(y) = y^2/(2\|K\|_2^2\mu(x))$  for all  $y \in \mathbb{R}$ . To complete the proof of Theorem 3.2, it suffices to prove that

$$(11) \quad \lim_{n \rightarrow +\infty} \frac{\sqrt{|\mathbb{T}_n|h_n^d}}{b_n} B_{h_n}(x) = 0.$$

Next, using that

$$\begin{aligned} \mu(x - h_n y) - \mu(x) &= \sum_{j=1}^d (\mu(x_1 - h_n y_1, \dots, x_j - h_n y_j, x_{j+1}, \dots, x_d) \\ &\quad - \mu(x_1 - h_n y_1, \dots, x_{j-1} - h_n y_{j-1}, x_j, x_{j+1}, \dots, x_d)), \end{aligned}$$

the Taylor expansion and Assumption 2.12, we get that, for some finite constant  $C > 0$ ,

$$\begin{aligned} |\mathbb{T}_n|^{1/2} h_n^{d/2} B_{h_n}(x) &= \sqrt{|\mathbb{T}_n|h_n^d} \left| \int_{\mathbb{R}^d} h_n^{-d} K(h_n^{-1}(x - y)) \mu(y) dy - \mu(x) \right| \\ &= \sqrt{|\mathbb{T}_n|h_n^d} \left| \int_{\mathbb{R}^d} K(y) (\mu(x - h_n y) - \mu(x)) dy \right| \\ &\leq C \sqrt{|\mathbb{T}_n|h_n^d} \sum_{j=1}^d \int_{\mathbb{R}^d} K(y) \frac{(h_n |y_j|)^s}{[s]!} dy \\ &\leq C \sqrt{|\mathbb{T}_n|h_n^{2s+d}}. \end{aligned}$$

Now, (11) follows using the latter inequality and (6). This ends the proof of Theorem 3.2 for  $\mathbb{A}_n = \mathbb{T}_n$ . The proof is similar for  $\mathbb{A}_n = \mathbb{G}_n$  using  $f_{\ell,n} = f_{\ell,n}^0$ .

**4.3. Proof of Theorem 3.3.** . We begin the proof with  $\mathbb{A}_n = \mathbb{T}_n$ . We have the following decomposition:

$$(12) \quad \frac{b_n^{-1} \sqrt{|\mathbb{T}_n|h_n^d} (\widehat{\mu}_{\mathbb{T}_n}(x) - \mu(x))}{\|K\|_2 \sqrt{\widehat{\mu}_{\mathbb{A}_n^*}(x)} \vee \varpi_n} = T_1(n) + T_2(n)$$

where

$$\begin{aligned} T_1(n) &= (\|K\|_2 \sqrt{\mu(x)} b_n)^{-1} \sqrt{|\mathbb{T}_n|h_n^d} (\widehat{\mu}_{\mathbb{T}_n}(x) - \mu(x)); \\ T_2(n) &= \left( \frac{1}{\|K\|_2 \sqrt{\widehat{\mu}_{\mathbb{A}_n^*}(x)} \vee \varpi_n} - \frac{1}{\|K\|_2 \sqrt{\mu(x)}} \right) b_n^{-1} \sqrt{|\mathbb{T}_n|h_n^d} (\widehat{\mu}_{\mathbb{T}_n}(x) - \mu(x)). \end{aligned}$$

First, we prove that

$$(13) \quad T_2(n) \xrightarrow[b_n^2]{\text{superexp}} 0.$$

Let  $\delta > 0$ . For all  $r > 0$ , we have

$$\begin{aligned} \mathbb{P}(|T_2(n)| > \delta) &\leq \mathbb{P}(|b_n^{-1} \sqrt{|\mathbb{T}_n|h_n^d} (\widehat{\mu}_{\mathbb{T}_n}(x) - \mu(x))| > \delta/r) \\ &\quad + \mathbb{P}\left(\left| \frac{1}{\|K\|_2 \sqrt{\widehat{\mu}_{\mathbb{A}_n^*}(x)} \vee \varpi_n} - \frac{1}{\|K\|_2 \sqrt{\mu(x)}} \right| > r\right). \end{aligned}$$

This implies that (see for e.g [11], Lemma 1.2.15)

$$(14) \quad \limsup_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \mathbb{P}(|T_2(n)| > \delta) \leq \max \left\{ \limsup_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \mathbb{P}(|b_n^{-1} \sqrt{|\mathbb{T}_n| h_n^d} (\widehat{\mu}_{\mathbb{T}_n}(x) - \mu(x))| > \delta/r); \right. \\ \left. \limsup_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \mathbb{P} \left( \left| \frac{1}{\|K\|_2 \sqrt{\widehat{\mu}_{\mathbb{A}_n^*}(x)} \vee \varpi_n} - \frac{1}{\|K\|_2 \sqrt{\mu(x)}} \right| > r \right) \right\}.$$

Using Theorem 3.2 and the contraction principle, we have

$$(15) \quad \limsup_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \mathbb{P}(|b_n^{-1} \sqrt{|\mathbb{T}_n| h_n^d} (\widehat{\mu}_{\mathbb{T}_n}(x) - \mu(x))| > \delta/r) = -\frac{\delta^2}{2\|K\|_2 \mu(x) r^2}.$$

Following the step 1 of the proof of Theorem 6 in [7] and using Lemma 3.1, we can prove that

$$\|K\|_2^2 \widehat{\mu}_{\mathbb{A}_n^*}(x) \vee \varpi_n^2 \xrightarrow[b_n^2]{\text{superexp}} \|K\|_2^2 \mu(x).$$

Using Lemma B.2 in [3], the latter convergence implies that

$$(16) \quad \frac{1}{\|K\|_2 \sqrt{\widehat{\mu}_{\mathbb{A}_n^*}(x)} \vee \varpi_n} \xrightarrow[b_n^2]{\text{superexp}} \frac{1}{\|K\|_2 \sqrt{\mu(x)}}.$$

Using (14), (15) and (16), we get

$$\limsup_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \mathbb{P}(|T_2(n)| > \delta) \leq -\frac{\delta^2}{2\|K\|_2 \mu(x) r^2}.$$

Since  $r$  can be taken arbitrarily close to 0, we get (13) and using (12), this implies that

$$(17) \quad \frac{b_n^{-1} \sqrt{|\mathbb{T}_n| h_n^d} (\widehat{\mu}_{\mathbb{T}_n}(x) - \mu(x))}{\|K\|_2 \sqrt{\widehat{\mu}_{\mathbb{A}_n^*}(x)} \vee \varpi_n} \xrightarrow[b_n^2]{\text{superexp}} T_1(n).$$

Using Theorem 3.2 and the contraction principle, we get that  $T_1(n)$  satisfies a moderate deviation principle on  $\mathbb{R}$  with speed  $b_n^2$  and rate function  $I'$  defined by:  $I'(y) = y^2/2$  for all  $y \in \mathbb{R}$ . Using (17) and Remark 2.7, we get the result of Theorem 3.3.

**4.4. Proof of Corollary 3.6.** Let  $\mathbf{a} = (a_0, \dots, a_k)^t \in \mathbb{R}^{k+1}$ . Let  $n > k$ . We consider the sequence  $\mathbf{f}_n = (f_{\ell,n}, n \geq \ell \geq 0)$  defined by  $f_{\ell,n} = 2^{\ell/2} a_\ell K_{h_{n-\ell}}(x - \cdot)$  for all  $\ell \in \{0, \dots, k\}$  and  $f_{\ell,n} = 0$  otherwise. We easily check that  $\mathbf{f}_n$  satisfies Assumptions 3.8. In particular, the asymptotic variance defined in (7) is given by  $\sigma^2 = (\sum_{\ell=0}^k 2^\ell a_\ell^2) \|K\|_2^2 \mu(x)$ . Observe that the linear combinaison  $M_n(\mathbf{a})$ , with coefficients  $\mathbf{a} = (a_0, \dots, a_k)^t \in \mathbb{R}^{k+1}$ , of the estimators  $|\mathbb{G}_{n-\ell}|^{1/2} h_{n-\ell}^{d/2} (\widehat{\mu}_{\mathbb{G}_{n-\ell}}(x) - \mu(x))$ ,  $\ell \in \{0, \dots, k\}$  has the following decomposition:

$$(18) \quad M_n(\mathbf{a}) = N_{n,\emptyset}(\mathbf{f}_n) + \sum_{\ell=0}^k a_\ell (|\mathbb{G}_{n-\ell}| h_{n-\ell}^d)^{1/2} B_{h_{n-\ell}}(x),$$

where  $N_{n,\emptyset}(\mathbf{f}_n)$  is defined in (8) and the  $B_{h_{n-\ell}}(x)$ ,  $\ell \in \{0, \dots, k\}$ , are defined in (4). Applying Theorem 3.10, we get that  $b_n^{-1} N_{n,\emptyset}(\mathbf{f}_n)$  satisfies a moderate deviation principle on  $\mathbb{R}$  with speed  $b_n^2$  and rate function  $I_{x,\mathbf{a}} : \mathbb{R} \rightarrow \mathbb{R}$  defined by

$$(19) \quad I_{x,\mathbf{a}}(y) = \frac{y^2}{2 (\sum_{\ell=0}^k 2^\ell a_\ell^2) \|K\|_2^2 \mu(x)}, \quad y \in \mathbb{R}.$$

Using (11), we have that

$$\lim_{n \rightarrow +\infty} \frac{1}{b_n} \sum_{\ell=0}^k a_\ell (|\mathbb{G}_{n-\ell}| h_{n-\ell}^d)^{1/2} B_{h_{n-\ell}}(x) = 0.$$

Using Remark 2.7, this implies that

$$(20) \quad \frac{1}{b_n} \sum_{\ell=0}^k a_\ell (|\mathbb{G}_{n-\ell}| h_{n-\ell}^d)^{1/2} B_{h_{n-\ell}}(x) \xrightarrow[b_n^2]{\text{superexp}} 0.$$

Using (18) and (20) we get that  $b_n^{-1}M_n(\mathbf{a})$  and  $b_n^{-1}N_{n,\emptyset}(\mathbf{f}_n)$  satisfy the same moderate deviation principle. We then conclude that  $b_n^{-1}M_n(\mathbf{a})$  satisfies a moderate deviation principle on  $\mathbb{R}$  with speed  $b_n^2$  and rate function  $I_{x,\mathbf{a}}$  defined in (19). Since this is true for all vector  $\mathbf{a} \in \mathbb{R}^{k+1}$ , that is for all the linear combinations of the estimators  $|\mathbb{G}_{n-\ell}|^{1/2} h_{n-\ell}^{d/2} (\widehat{\mu}_{\mathbb{G}_{n-\ell}}(x) - \mu(x))$ ,  $\ell \in \{0, \dots, k\}$ , we get the result of Corollary 3.6.

## 5. PROOF OF THEOREM 3.10

We begin with some notations. We will denote by  $C$  any unimportant finite constant which may vary from line to line (in particular  $C$  does not depend on  $n \in \mathbb{N}$  nor on the considered sequence of functions  $\mathbf{f}_n = (f_{\ell,n}, n \geq \ell \geq 0)$ ). Let  $(p_n, n \in \mathbb{N})$  be a non-decreasing sequence of elements of  $\mathbb{N}^*$  such that

$$\lim_{n \rightarrow +\infty} p_n^3 b_n^2 |\mathbb{G}_{n-p_n}|^{-1} = 0.$$

When there is no ambiguity, we write  $p$  for  $p_n$ .

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

$$\mathcal{F}_i = \{X_u; u \in \mathbb{T} \text{ such that } u \leq i\}.$$

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

We define for  $n \in \mathbb{N}$ ,  $i \in \mathbb{G}_{n-p_n}$  and  $\mathbf{f}_n$  the martingale increments:

$$(21) \quad \Delta_{n,i}(\mathbf{f}_n) = N_{n,i}(\mathbf{f}_n) - \mathbb{E}[N_{n,i}(\mathbf{f}_n) | \mathcal{F}_i] \quad \text{and} \quad \Delta_n(\mathbf{f}_n) = \sum_{i \in \mathbb{G}_{n-p_n}} \Delta_{n,i}(\mathbf{f}_n),$$

where

$$(22) \quad N_{n,i}(\mathbf{f}_n) = |\mathbb{G}_n|^{-1/2} \sum_{\ell=0}^p M_{i\mathbb{G}_{p-\ell}}(\tilde{f}_{\ell,n}) \quad \text{and} \quad i\mathbb{G}_{p-\ell} = \{ij, j \in \mathbb{G}_{p-\ell}\}.$$

We have:

$$\sum_{i \in \mathbb{G}_{n-p_n}} N_{n,i}(\mathbf{f}_n) = |\mathbb{G}_n|^{-1/2} \sum_{\ell=0}^{p_n} M_{\mathbb{G}_{n-\ell}}(\tilde{f}_{\ell,n}) = |\mathbb{G}_n|^{-1/2} \sum_{k=n-p_n}^n M_{\mathbb{G}_k}(\tilde{f}_{n-k,n}).$$

Using the branching Markov property, we get for  $i \in \mathbb{G}_{n-p_n}$ :

$$(23) \quad \mathbb{E}[N_{n,i}(\mathbf{f}_n) | \mathcal{F}_i] = \mathbb{E}[N_{n,i}(\mathbf{f}_n) | X_i] = |\mathbb{G}_n|^{-1/2} \sum_{\ell=0}^{p_n} \mathbb{E}_{X_i} [M_{\mathbb{G}_{p_n-\ell}}(\tilde{f}_{\ell,n})].$$

We have the following decomposition:

$$(24) \quad N_{n,\emptyset}(\mathbf{f}_n) = \Delta_n(\mathbf{f}_n) + R_0(n) + R_1(n),$$

where  $\Delta_n(\mathbf{f})$  is defined in (21) and:

$$R_0(n) = |\mathbb{G}_n|^{-1/2} \sum_{k=0}^{n-p_n-1} M_{\mathbb{G}_k}(\tilde{f}_{n-k,n}) \quad \text{and} \quad R_1(n) = \sum_{i \in \mathbb{G}_{n-p_n}} \mathbb{E}[N_{n,i}(\mathbf{f}_n) | \mathcal{F}_i].$$

From (24), our goals will be achieved if we prove the following:

$$(25) \quad b_n^{-1} R_0(n) \xrightarrow[b_n^2]{\text{superexp}} 0;$$

$$(26) \quad b_n^{-1} R_1(n) \xrightarrow[b_n^2]{\text{superexp}} 0;$$

$$(27) \quad b_n^{-1} \Delta_n(\mathbf{f}) \quad \text{satisfies a MDP on } S \text{ with speed } b_n^2 \text{ and rate function } I.$$

Note that (25) and (26) mean that  $R_0(n)$  and  $R_1(n)$  are negligible in the sense of moderate deviations in such a way that using (24) and Remark 2.7,  $N_{n,\emptyset}(\mathbf{f})$  and  $\Delta_n(\mathbf{f})$  satisfy the same moderate deviation principle. To prove (27), the main method we will use is the moderate deviations for martingale (see [12] for more details).

In the sequel, the sequence  $(2^{-\gamma n}, n \in \mathbb{N})$  which appears in Assumption 3.8 will be denoted  $(h_n, n \in \mathbb{N})$  in such a way that we have  $2^{-d\gamma n/2} = h_n^{d/2}$ . We have the following result.

**Lemma 5.1.** *Under the assumptions of Theorem 3.10, we have  $b_n^{-1} R_0(n) \xrightarrow[b_n^2]{\text{superexp}} 0$ .*

*Proof.* Let  $\delta > 0$ . Using the Chernoff bound, we have, for all  $\lambda > 0$ ,

$$(28) \quad \mathbb{P}\left(b_n^{-1} R_0(n) > \delta\right) \leq \exp\left(-\lambda b_n |\mathbb{G}_n|^{1/2} \delta\right) \mathbb{E}\left[\exp\left(\lambda \sum_{\ell=0}^{n-p-1} M_{\mathbb{G}_\ell}(\tilde{f}_{n-\ell,n})\right)\right].$$

For all  $k \in \{1, \dots, n-p\}$  and for  $u \in \mathbb{T}$ , we set

$$g_{p,k} = \sum_{r=0}^{k-1} 2^r Q^r \tilde{f}_{p+k-r,n} \quad \text{and} \quad Z_{p,k}(u) = g_{p,k}(X_{u0}) + g_{p,k}(X_{u1}) - 2Qg_{p,k}(X_u).$$

Then, using recursively the fact that

$$\sum_{u \in \mathbb{G}_\ell} f(X_u) = \sum_{u \in \mathbb{G}_{\ell-1}} (f(X_{u0}) + f(X_{u1}) - 2Qf(X_u)) + \sum_{u \in \mathbb{G}_{\ell-1}} 2Qf(X_u),$$

for all  $\ell \geq 1$  and for some function  $f$ , we get

$$\mathbb{E}\left[\exp\left(\lambda \sum_{\ell=0}^{n-p-1} M_{\mathbb{G}_\ell}(\tilde{f}_{n-\ell,n})\right)\right] = \mathbb{E}\left[\exp\left(\lambda g_{p,n-p}(X_\emptyset)\right) \prod_{k=1}^{n-p-1} \exp\left(\lambda \sum_{u \in \mathbb{G}_{n-p-k-1}} Z_{p,k}(u)\right)\right].$$

For all  $m \in \{1, \dots, n-p-1\}$ , we set

$$\mathbb{I}_m = \mathbb{E}\left[\exp\left(\lambda g_{p,n-p}(X_\emptyset)\right) \prod_{k=m}^{n-p-1} \exp\left(\lambda \sum_{u \in \mathbb{G}_{n-p-k-1}} Z_{p,k}(u)\right)\right].$$

Using the branching Markov property, we get the following decomposition:

$$\mathbb{I}_m = \mathbb{E} \left[ \exp(\lambda g_{p,n-p}(X_\emptyset)) \mathbb{J}_m \prod_{k=m+1}^{n-p-1} \exp(\lambda \sum_{u \in \mathbb{G}_{n-p-k-1}} Z_{p,k}(u)) \right],$$

with

$$\mathbb{J}_m = \prod_{u \in \mathbb{G}_{n-p-m-1}} \mathbb{E}_{X_u} \left[ \exp(\lambda Z_{p,m}(u)) \right].$$

For all  $u \in \mathbb{G}_{n-p-m-1}$ , we will upper bound the quantity  $\mathbb{E}_{X_u}[\exp(\lambda Z_{p,m}(u))]$  and then  $\mathbb{J}_m$ . We claim that:

$$(29) \quad |Z_{p,m}(u)| \leq M = C h_n^{-d/2};$$

$$(30) \quad \mathbb{E}_{X_u}[Z_{p,m}(u)^2] \leq \sigma_m^2 = C + C h_n^d \left( \sum_{r=0}^{m-1} (2\alpha)^{r-1} \right)^2 \mathbf{1}_{\{m>1\}}.$$

For that purpose, we plan to use the bound

$$(31) \quad \mathbb{E}[\exp(\lambda Z)] \leq \exp\left(\frac{\lambda^2 \sigma^2}{2(1 - \lambda M/3)}\right)$$

valid for any  $\lambda \in (0, 3/M)$ , any random variable  $Z$  such that  $|Z| \leq M$ ,  $\mathbb{E}[Z] = 0$  and  $\mathbb{E}[Z^2] \leq \sigma^2$ .

For all  $u \in \mathbb{G}_{n-p-m-1}$  and for all  $\lambda \in (0, C h_n^{-d/2}/3)$  we get, using (29)-(31),

$$\mathbb{E}_{X_u} \left[ \exp(\lambda Z_{p,m}(u)) \right] \leq \exp\left(\frac{\lambda^2 \sigma_m^2}{2(1 - \lambda M/3)}\right).$$

For all  $m \in \{1, \dots, n-p-1\}$ , the latter inequality implies that

$$\mathbb{J}_m \leq \exp\left(\frac{\lambda^2 \sigma_m^2 |\mathbb{G}_{n-p-m-1}|}{2(1 - \lambda M/3)}\right) \quad \text{and} \quad \mathbb{I}_m \leq \exp\left(\frac{\lambda^2 \sigma_m^2 |\mathbb{G}_{n-p-m-1}|}{2(1 - \lambda M/3)}\right) \mathbb{I}_{m+1}.$$

Recall that  $\mathbb{I}_1 = \mathbb{E}[\exp(\lambda \sum_{\ell=0}^{n-p-1} M_{\mathbb{G}_\ell}(\tilde{f}_{n-\ell,n}))]$ . By recurrence, we get

$$\mathbb{E} \left[ \exp \left( \lambda \sum_{\ell=0}^{n-p-1} M_{\mathbb{G}_\ell}(\tilde{f}_{n-\ell,n}) \right) \right] = \mathbb{I}_1 \leq \exp \left( \frac{\lambda^2 \sum_{m=1}^{n-p-1} \sigma_m^2 |\mathbb{G}_{n-p-m-1}|}{2(1 - \lambda M/3)} \right) \mathbb{E} \left[ \exp \left( \lambda g_{p,n-p}(X_\emptyset) \right) \right].$$

Using (i) and (ii) of Assumption 3.8 and (3), we have

$$|g_{p,n-p}| \leq |\tilde{f}_{n,n}| + \sum_{r=1}^{n-p-1} 2^r |\mathcal{Q}^{r-1}(\mathcal{Q}\tilde{f}_{n-r,n})| \leq C h_n^{-d/2} + C h_n^{d/2} \sum_{r=1}^{n-p-1} (2\alpha)^{r-1}.$$

This implies that

$$\begin{aligned} \mathbb{E} \left[ \exp \left( \lambda \sum_{\ell=0}^{n-p-1} M_{\mathbb{G}_\ell}(\tilde{f}_{n-\ell,n}) \right) \right] &\leq \exp \left( \frac{\lambda^2 \sum_{m=1}^{n-p-1} \sigma_m^2 |\mathbb{G}_{n-p-m-1}|}{2(1 - \lambda M/3)} \right) \\ &\quad \times \exp \left( \lambda C h_n^{-d/2} + \lambda C h_n^{d/2} \sum_{r=0}^{n-p-2} (2\alpha)^r \right). \end{aligned}$$

Distinguishing the cases  $2\alpha \leq 1$ ,  $1/2 < 2\alpha \leq \sqrt{2}$  and  $2\alpha > \sqrt{2}$  and using (5) for  $2\alpha > 1$ , we get

$$\mathbb{E} \left[ \exp \left( \lambda \sum_{\ell=0}^{n-p-1} M_{\mathbb{G}_\ell}(\tilde{f}_{n-\ell,n}) \right) \right] \leq \exp \left( \frac{c_1 \lambda^2 |\mathbb{G}_{n-p}|}{2(1 - c_2 \lambda h_n^{-d/2}/3)} \right) \exp \left( c_3 \lambda h_n^{-d/2} \right),$$

where  $c_1$ ,  $c_2$  and  $c_3$  are some positive constants. The latter inequality and (28) imply that

$$\mathbb{P}\left(b_n^{-1}R_0(n) > \delta\right) \leq \exp\left(-\lambda b_n |\mathbb{G}_n|^{1/2} \delta + \frac{c_1 \lambda^2 |\mathbb{G}_{n-p}|}{2(1 - c_2 \lambda h_n^{-d/2}/3)}\right) \exp\left(c_3 \lambda h_n^{-d/2}\right).$$

Taking<sup>1</sup>

$$\lambda = \frac{3 b_n |\mathbb{G}_n|^{1/2} \delta}{2 c_2 b_n |\mathbb{G}_n|^{1/2} h_n^{-d/2} \delta + 3 c_1 |\mathbb{G}_{n-p}|},$$

we are led to

$$\mathbb{P}\left(b_n^{-1}R_0(n) > \delta\right) \leq C \exp\left(-\frac{3 \delta^2 b_n^2 |\mathbb{G}_n|}{2(c_2 \delta b_n |\mathbb{G}_n|^{1/2} h_n^{-d/2} + 3 c_1 |\mathbb{G}_{n-p}|)}\right).$$

Since we can do the same thing for  $-f_n$  instead of  $f_n$ , we get that

$$(32) \quad \mathbb{P}\left(b_n^{-1}|R_0(n)| > \delta\right) \leq 2C \exp\left(-\frac{3 \delta^2 b_n^2 |\mathbb{G}_n|}{2(c_2 \delta b_n |\mathbb{G}_n|^{1/2} h_n^{-d/2} + 3 c_1 |\mathbb{G}_{n-p}|)}\right).$$

Finally, in the latter inequality, taking the log, dividing by  $b_n^2$  and letting  $n$  goes to infinity, we get the result of Lemma 5.1. Now, to end the proof, we will prove (29) and (30).

*Proof of (29).* Using Assumption 2.9, (i) and (ii) of Assumption 3.8 and Assumption 2.13, we get

$$\begin{aligned} |Z_{p,m}(u)| &\leq C \|\tilde{f}_{p+1,n}\|_\infty + C(1 + 2\alpha) \left( \sum_{r=1}^{m-1} (2\alpha)^{r-1} \|\mathcal{Q}f_{p+m-r,n}\|_\infty \right) \mathbf{1}_{\{m>1\}} \\ &\leq C h_n^{-d/2} + C h_n^{d/2} \sum_{r=0}^{m-1} (2\alpha)^r \leq C h_n^{-d/2}. \end{aligned}$$

*Proof of (30).* Using the branching Markov property for the second inequality, Assumption 2.9 for the fourth inequality and (i) and (ii) of Assumption 3.8 for the last inequality, we get

$$\begin{aligned} \mathbb{E}_{X_u}[Z_{p,m}(u)^2] &\leq \mathbb{E}_{X_u}[(g_{p,m}(X_{u0}) + g_{p,m}(X_{u1}))^2] \leq C \mathcal{Q}(g_{p,m}^2)(X_u) \\ &\leq C \mathcal{Q}(\tilde{f}_{p+1,n}^2)(X_u) + C \mathcal{Q}\left(\left(\sum_{r=1}^{m-1} 2^r \mathcal{Q}^{r-1}(\mathcal{Q}\tilde{f}_{p+m-r,n})\right)^2\right)(X_u) \mathbf{1}_{\{m>1\}} \\ &\leq C \|\mathcal{Q}\tilde{f}_{p+1,n}^2\|_\infty + \left(\sum_{r=1}^{m-1} (2\alpha)^{r-1} \|\mathcal{Q}f_{p+m-r,n}\|_\infty\right)^2 \mathbf{1}_{\{m>1\}} \\ &\leq C + C h_n^d \left(\sum_{r=0}^{m-1} (2\alpha)^r\right)^2 \mathbf{1}_{\{m>1\}}. \end{aligned}$$

□

Next, we have the following result.

**Lemma 5.2.** *Under the assumptions of Theorem 3.10, we have  $b_n^{-1}R_1(n) \xrightarrow{\text{superexp}} \frac{0}{b_n^2}$ .*

<sup>1</sup>In fact, we use the following. For  $\alpha, \beta, \gamma > 0$  and  $h(x) = -\alpha x + \frac{\beta x^2}{2(1-\gamma x)}$  we have  $h(x^*) = \frac{-\alpha^2}{2(\beta+\alpha\gamma)}$  for the choice  $x^* = \frac{\alpha}{2\alpha\gamma+\beta} \in (0, 1/\gamma)$ .

*Proof.* We have, using (23) and (49),

$$(33) \quad R_1(n) = |\mathbb{G}_n|^{-1/2} M_{\mathbb{G}_{n-p}}(g_{p,n}) \quad \text{where} \quad g_{p,n} = \sum_{\ell=0}^p 2^{p-\ell} Q^{p-\ell} \tilde{f}_{\ell,n}.$$

We follow the same arguments that in the proof of Lemma 5.1. For all  $m \in \{1, \dots, n-p\}$  and for all  $u \in \mathbb{T}$ , we set

$$Z_{p,m}(u) = 2^{m-1} Q^{m-1} g_p(X_{u0}) + 2^{m-1} Q^{m-1} g_p(X_{u1}) - 2^m Q^m g_p(X_u).$$

We also consider the following quantities for  $m \in \{1, \dots, n-p\}$  and  $\lambda > 0$ :

$$\begin{aligned} \mathbb{I}_m &= \mathbb{E} \left[ \exp \left( \lambda 2^{n-p} Q^{n-p} g_{p,n}(X_\emptyset) \right) \prod_{k=m}^{n-p} \exp \left( \lambda \sum_{u \in \mathbb{G}_{n-p-k}} Z_{p,k}(u) \right) \right] \quad \text{and} \\ \mathbb{J}_m &= \prod_{u \in \mathbb{G}_{n-p-m}} \mathbb{E}_{X_u} \left[ \exp \left( \lambda Z_{p,m}(u) \right) \right]. \end{aligned}$$

Note that using the branching Markov property, we have

$$(34) \quad \mathbb{I}_m = \mathbb{E} \left[ \exp \left( \lambda 2^{n-p} Q^{n-p} g_{p,n}(X_\emptyset) \right) \prod_{k=m+1}^{n-p} \exp \left( \lambda \sum_{u \in \mathbb{G}_{n-p-k}} Z_{p,k}(u) \right) \mathbb{J}_m \right].$$

As for (29)-(30), for all  $m \in \{1, \dots, n-p\}$  and  $u \in \mathbb{G}_{n-p-m}$ , one can prove that

$$(35) \quad |Z_{p,m}(u)| \leq M = Ch^{-d/2} \quad \text{and} \quad \mathbb{E}_{X_u} \left[ Z_{p,m}(u)^2 \right] \leq \sigma_m^2 = C \mathbf{1}_{\{m=1\}} + Ch_n^d \left( \sum_{\ell=0}^p (2\alpha)^{p+m-\ell-2} \right)^2.$$

Using (31) and (35), we have, for all  $u \in \mathbb{G}_{n-p-m}$  and for all  $\lambda \in (0, Ch^{-d/2}/3)$ ,

$$\mathbb{E}_{X_u} \left[ \exp(\lambda Z_{p,m}(u)) \right] \leq \exp \left( \frac{\lambda^2 \sigma_m^2}{2(1 - \lambda M/3)} \right).$$

The latter inequality and (34) imply that

$$\mathbb{I}_m \leq \exp \left( \frac{\lambda^2 \sigma_m^2 |\mathbb{G}_{n-p-m}|}{2(1 - \lambda M/3)} \right) \mathbb{I}_{m+1}.$$

By recurrence, this implies that

$$(36) \quad \mathbb{I}_1 \leq \exp \left( \frac{\lambda^2 \sum_{m=1}^{n-p} \sigma_m^2 |\mathbb{G}_{n-p-m}|}{2(1 - \lambda M/3)} \right) \mathbb{E} \left[ \exp \left( \lambda 2^{n-p} Q^{n-p} g_{p,n}(X_\emptyset) \right) \right].$$

Using (i) and (ii) of Assumption 3.8 and Assumption 2.9, we get

$$(37) \quad |g_{p,n}| \leq Ch_n^{d/2} \sum_{\ell=0}^p (2\alpha)^{n-\ell}.$$



From (36), (37) and according to the value of  $\alpha$ , we have, for some positive constants  $c_1$ ,  $c_2$  and  $c_3$  (recall the definition of  $M$  and  $\sigma_m^2$  given in (35)):

$$\begin{aligned} \mathbb{I}_1 &\leq C \exp\left(\frac{\lambda^2 c_1 |\mathbb{G}_{n-p}|}{2(1 - \lambda c_2 h_n^{-d/2}/3)}\right) && \text{if } 2\alpha \leq 1; \\ \mathbb{I}_1 &\leq \exp\left(\lambda c_3 (2\alpha)^n h_n^{d/2}\right) \exp\left(\frac{\lambda^2 c_1 |\mathbb{G}_{n-p}| (1 + (2\alpha)^{2p} h_n^d)}{2(1 - \lambda c_2 h_n^{-d/2}/3)}\right) && \text{if } 1 < 2\alpha \leq \sqrt{2}; \\ \mathbb{I}_1 &\leq \exp\left(\lambda c_3 (2\alpha)^n h_n^{d/2}\right) \exp\left(\frac{\lambda^2 c_1 |\mathbb{G}_{n-p}| (1 + (2\alpha)^{2p} h_n^d + 2^p (2\alpha^2)^n h_n^d)}{2(1 - \lambda c_2 h_n^{-d/2}/3)}\right) && \text{if } 2\alpha > \sqrt{2}. \end{aligned}$$

Recall that  $\mathbb{I}_1 = \mathbb{E}[\exp(\lambda M_{\mathbb{G}_{n-p}}(g_{p,n}))]$ . Using the Chernoff bound and (33), we have for all  $\lambda \in (0, C h_n^{-d/2}/3)$  and for all  $\delta > 0$ ,

$$\mathbb{P}\left(b_n^{-1} R_1(n) > \delta\right) \leq \exp\left(-\lambda b_n |\mathbb{G}_n|^{1/2} \delta\right) \mathbb{I}_1.$$

Taking

$$\lambda = \begin{cases} \frac{3b_n |\mathbb{G}_n|^{1/2} \delta}{2c_2 b_n |\mathbb{G}_n|^{1/2} h_n^{-d/2} \delta + 3c_1 |\mathbb{G}_{n-p}|} & \text{if } 2\alpha \leq 1 \\ \frac{3b_n |\mathbb{G}_n|^{1/2} \delta}{2c_2 b_n |\mathbb{G}_n|^{1/2} h_n^{-d/2} \delta + 3c_1 |\mathbb{G}_{n-p}| (1 + (2\alpha)^{2p} h_n^d)} & \text{if } 1 < 2\alpha \leq \sqrt{2} \\ \frac{3b_n |\mathbb{G}_n|^{1/2} \delta}{2c_2 b_n |\mathbb{G}_n|^{1/2} h_n^{-d/2} \delta + 3c_1 |\mathbb{G}_{n-p}| (1 + (2\alpha)^{2p} h_n^d + 2^p (2\alpha^2)^n h_n^d)} & \text{if } 2\alpha > \sqrt{2}, \end{cases}$$

and since we can do the same things for  $-\mathfrak{f}_n$  instead of  $\mathfrak{f}_n$ , we get, if  $2\alpha \leq 1$ :

$$\mathbb{P}\left(b_n^{-1} |R_1(n)| > \delta\right) \leq C \exp\left(-\frac{3b_n^2 |\mathbb{G}_n| \delta^2}{2(c_2 b_n |\mathbb{G}_n|^{1/2} h_n^{-d/2} \delta + 3c_1 |\mathbb{G}_{n-p}|)}\right);$$

if  $1 < 2\alpha \leq \sqrt{2}$ :

$$\begin{aligned} \mathbb{P}\left(b_n^{-1} |R_1(n)| > \delta\right) &\leq 2 \exp\left(\frac{c_3 (2\alpha)^n h_n^{d/2} b_n |\mathbb{G}_n|^{1/2}}{2c_2 b_n |\mathbb{G}_n|^{1/2} h_n^{-d/2} \delta + 3c_1 |\mathbb{G}_{n-p}| (1 + (2\alpha)^{2p} h_n^d)}\right) \\ &\quad \times \exp\left(-\frac{3b_n^2 |\mathbb{G}_n| \delta^2}{2(c_2 b_n |\mathbb{G}_n|^{1/2} h_n^{-d/2} \delta + 3c_1 |\mathbb{G}_{n-p}| (1 + (2\alpha)^{2p} h_n^d))}\right); \end{aligned}$$

if  $2\alpha > \sqrt{2}$ :

$$\begin{aligned} \mathbb{P}\left(b_n^{-1} |R_1(n)| > \delta\right) &\leq 2 \exp\left(\frac{c_3 (2\alpha)^n h_n^{d/2} b_n |\mathbb{G}_n|^{1/2}}{2c_2 b_n |\mathbb{G}_n|^{1/2} h_n^{-d/2} \delta + 3c_1 |\mathbb{G}_{n-p}| (1 + (2\alpha)^{2p} h_n^d + 2^p (2\alpha^2)^n h_n^d)}\right) \\ &\quad \times \exp\left(-\frac{3b_n^2 |\mathbb{G}_n| \delta^2}{2(c_2 b_n |\mathbb{G}_n|^{1/2} h_n^{-d/2} \delta + 3c_1 |\mathbb{G}_{n-p}| (1 + (2\alpha)^{2p} h_n^d + 2^p (2\alpha^2)^n h_n^d))}\right). \end{aligned}$$

Finally, applying the log to each of these last three inequalities, dividing by  $b_n^2$ , letting  $n$  goes to infinity and using (6) and Assumption 2.13, we get the result of Lemma 5.2.  $\square$

From (24), Lemmas 5.1 and 5.2, we have

$$(38) \quad b_n^{-1} N_{n,\emptyset}(\mathfrak{f}_n) \stackrel{\text{superexp}}{\underset{b_n^2}{\sim}} b_n^{-1} \Delta_n(\mathfrak{f}_n).$$

As a consequence, using Remark 2.7,  $b_n^{-1} N_{n,\emptyset}(\mathfrak{f}_n)$  and  $b_n^{-1} \Delta_n(\mathfrak{f}_n)$  satisfy the same moderate deviation principle.

We now study the martingale part  $\Delta_n(\mathbf{f}_n)$  of the decomposition (24). The bracket  $V(n)$  of  $\Delta_n(\mathbf{f}_n)$  is defined by:

$$V(n) = \sum_{i \in \mathbb{G}_{n-p_n}} \mathbb{E} [\Delta_{n,i}(\mathbf{f}_n)^2 | \mathcal{F}_i].$$

Using (22) and (21), we write:

$$(39) \quad V(n) = |\mathbb{G}_n|^{-1} \sum_{i \in \mathbb{G}_{n-p_n}} \mathbb{E}_{X_i} \left[ \left( \sum_{\ell=0}^{p_n} M_{\mathbb{G}_{p_n-\ell}}(\tilde{f}_{\ell,n}) \right)^2 \right] - R_2(n) = V_1(n) + 2V_2(n) - R_2(n),$$

with:

$$\begin{aligned} V_1(n) &= |\mathbb{G}_n|^{-1} \sum_{i \in \mathbb{G}_{n-p_n}} \sum_{\ell=0}^{p_n} \mathbb{E}_{X_i} \left[ M_{\mathbb{G}_{p_n-\ell}}(\tilde{f}_{\ell,n})^2 \right], \\ V_2(n) &= |\mathbb{G}_n|^{-1} \sum_{i \in \mathbb{G}_{n-p_n}} \sum_{0 \leq \ell < k \leq p_n} \mathbb{E}_{X_i} \left[ M_{\mathbb{G}_{p_n-\ell}}(\tilde{f}_{\ell,n}) M_{\mathbb{G}_{p_n-k}}(\tilde{f}_{k,n}) \right], \\ R_2(n) &= \sum_{i \in \mathbb{G}_{n-p_n}} \mathbb{E} [N_{n,i}(\mathbf{f}_n) | X_i]^2. \end{aligned}$$

We have the following result.

**Lemma 5.3.** *Under the Assumptions of Theorem 3.10, we have  $R_2(n) \xrightarrow[b_n^2]{\text{superexp}} 0$ .*

*Proof.* Using the branching Markov property, we have

$$R_2(n) = |\mathbb{G}_n|^{-1} M_{\mathbb{G}_{n-p}}(g_p) \quad \text{with} \quad g_p = \left( \sum_{\ell=0}^p 2^{p-\ell} \mathbb{Q}^{p-\ell} \tilde{f}_{\ell,n} \right)^2.$$

Using Assumption 2.9 and (i) and (ii) of Assumption 3.8, we get

$$\begin{aligned} \|g_p\|_\infty &\leq C \|\tilde{f}_{p,n}\|_\infty^2 + C \left\| \left( \sum_{\ell=0}^{p-1} 2^{p-\ell} \mathbb{Q}^{p-\ell} \tilde{f}_{\ell,n} \right) \right\|_\infty^2 \\ &\leq C h_n^{-d} + C \left( \sum_{\ell=0}^{p-1} (2\alpha)^{p-\ell} h_n^{d/2} \right)^2 \\ &\leq C h_n^{-d} \mathbf{1}_{\{2\alpha \leq 1\}} + C (h_n^{-d} + h_n^d (2\alpha)^{2p}) \mathbf{1}_{\{2\alpha > 1\}}. \end{aligned}$$

This implies that

$$(40) \quad R_2(n) \leq C |\mathbb{G}_n|^{-1} h_n^{-d} \mathbf{1}_{\{2\alpha \leq 1\}} + C (|\mathbb{G}_n|^{-1} h_n^{-d} + (2\alpha^2)^p h_n^d |\mathbb{G}_{n-p}|^{-1}) \mathbf{1}_{\{2\alpha > 1\}}.$$

Recall that  $h_n = 2^{-n\gamma}$  with  $\gamma \in (0, 1/d)$ . Using Assumption 2.13, we conclude from (40) that  $R_2(n)$  is bounded by a deterministic sequence which converge to 0. As a consequence, using Remark 2.8, we get the result of Lemma 5.3.  $\square$

Recall  $\sigma^2$  given in (7). We have the following result.

**Lemma 5.4.** *Under the Assumptions of Theorem 3.10, we have  $V_1(n) \xrightarrow[b_n^2]{\text{superexp}} \sigma^2$ .*

*Proof.* We have the following decomposition which is a consequence of (50):

$$V_1(n) = V_3(n) + V_4(n),$$

with

$$V_3(n) = |\mathbb{G}_n|^{-1} \sum_{i \in \mathbb{G}_{n-p}} \sum_{\ell=0}^p 2^{p-\ell} Q^{p-\ell}(\tilde{f}_{\ell,n}^2)(X_i),$$

$$V_4(n) = |\mathbb{G}_n|^{-1} \sum_{i \in \mathbb{G}_{n-p}} \sum_{\ell=0}^{p-1} \sum_{k=0}^{p-\ell-1} 2^{p-\ell+k} Q^{p-1-(\ell+k)} \left( \mathcal{P} \left( Q^k \tilde{f}_{\ell,n} \otimes^2 \right) \right) (X_i).$$

Now, the result of Lemma 5.4 is a direct consequence of the following:

$$(41) \quad V_3(n) \xrightarrow[b_n^2]{\text{superexp}} \sigma^2;$$

$$(42) \quad V_4(n) \xrightarrow[b_n^2]{\text{superexp}} 0.$$

To end the proof, we will now prove (41) and (42).

*Proof of (41).* Set

$$g_{p,n} = \sum_{\ell=0}^p 2^{-\ell} Q^{p-\ell}(\tilde{f}_{\ell,n}^2 - \langle \mu, \tilde{f}_{\ell,n}^2 \rangle) \quad \text{and} \quad H_3^{[n]}(\mathbf{f}_n) = \sum_{\ell=0}^p 2^{-\ell} \langle \mu, \tilde{f}_{\ell,n}^2 \rangle.$$

Following the same arguments that in the proof of Lemmas 5.1 and 5.2, we get after studious calculations:

if  $2\alpha \leq 1$ ,

$$\begin{aligned} \mathbb{P}\left(|V_3(n) - H_3^{[n]}| > \delta\right) &= \mathbb{P}\left(|\mathbb{G}_{n-p}|^{-1} |M_{\mathbb{G}_{n-p}}(g_{p,n})| > \delta\right) \\ &\leq C \exp\left(\frac{Cp\delta}{C\delta h_n^{-d} + 3(p^2 2^{-p} + 2^{-p} h_n^{-d})}\right) \exp\left(-\frac{3\delta^2 |\mathbb{G}_n|}{2(C\delta h_n^{-d} + 3(p^2 2^{-p} + 2^{-p} h_n^{-d}))}\right); \end{aligned}$$

if  $1 < 2\alpha \leq \sqrt{2}$ ,

$$\begin{aligned} \mathbb{P}\left(|V_3(n) - H_3^{[n]}| > \delta\right) &= \mathbb{P}\left(|\mathbb{G}_{n-p}|^{-1} |M_{\mathbb{G}_{n-p}}(g_{p,n})| > \delta\right) \\ &\leq \exp\left(\frac{c_3(2\alpha)^n h_n^d \delta}{c_2\delta + 3c_1((2\alpha^2)^p h_n^d + 2^{-p})}\right) \exp\left(-\frac{3\delta^2 |\mathbb{G}_n| h_n^d}{2(c_2\delta + 3c_1((2\alpha^2)^p h_n^d + 2^{-p}))}\right); \end{aligned}$$

if  $2\alpha > \sqrt{2}$ ,

$$\begin{aligned} \mathbb{P}\left(|V_3(n) - H_3^{[n]}| > \delta\right) &= \mathbb{P}\left(|\mathbb{G}_{n-p}|^{-1} |M_{\mathbb{G}_{n-p}}(g_{p,n})| > \delta\right) \\ &\leq \exp\left(\frac{c_3(2\alpha)^n h_n^d \delta}{c_2\delta + 3c_1((2\alpha^2)^n h_n^d + 2^{-p})}\right) \exp\left(-\frac{3\delta^2 |\mathbb{G}_n| h_n^d}{2(c_2\delta + 3c_1((2\alpha^2)^n h_n^d + 2^{-p}))}\right); \end{aligned}$$

Taking the log, dividing by  $b_n^2$ , letting  $n$  goes to the infinity and using (6) and Assumption 2.13, we get

$$\limsup_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(|V_3(n) - H_3^{[n]}| > \delta\right) = -\infty.$$

Next, using (iii) of Assumption 3.8, we get  $\lim_{n \rightarrow +\infty} H_3^{[n]}(\mathbf{f}_n) = \sigma^2$ . This ends the proof of (41) since  $(H_3^{[n]}(\mathbf{f}_n))$  is a deterministic sequence.

*Proof of (42).* We set

$$h_{\ell,k}^{(n)} = 2^{k-\ell} \mathcal{Q}^{p-1-(\ell+k)} (\mathcal{P}(\mathcal{Q}^k \tilde{f}_{\ell,n} \otimes^2)) \quad \text{and} \quad H_{4,n} = \sum_{\ell=0}^{p-1} \sum_{k=0}^{p-\ell-1} h_{\ell,k}^{(n)}$$

in such a way that  $V_4(n) = |\mathbb{G}_{n-p}|^{-1} M_{\mathbb{G}_{n-p}}(H_{4,n})$ . Using, (3) and (i) and (ii) of Assumption 3.8, we get

$$|h_{\ell,k}^{(n)}| \leq 2^{k-\ell} \mathcal{P}(|\mathcal{Q}^k \tilde{f}_{\ell,n}| \otimes^2) \leq C 2^{k-\ell} h_n^d \alpha^{2k}.$$

This implies that  $|H_{4,n}| \leq c_n$  and then that  $|V_4(n)| \leq c_n$ , where the sequence  $(c_n, n \in \mathbb{N})$  is defined by

$$c_n = C h_n^d \mathbf{1}_{\{2\alpha^2 \leq 1\}} + C h_n^d (2\alpha^2)^p \mathbf{1}_{\{2\alpha^2 > 1\}}$$

Using (5) and the fact that  $(h_n, n \in \mathbb{N})$  converges to 0, we get that the sequence  $(c_n, n \in \mathbb{N})$  converges to 0. Thus, we have that  $V_4(n)$  is bounded by a deterministic sequence which converges to 0. Then (42) follows using Remark 2.8.  $\square$

**Lemma 5.5.** *Under the Assumptions of Theorem 3.10, we have  $V_2(n) \xrightarrow[b_n^2]{\text{superexp}} 0$ .*

*Proof.* Using (51), we get:

$$V_2(n) = V_5(n) + V_6(n),$$

with

$$V_5(n) = |\mathbb{G}_n|^{-1} \sum_{i \in \mathbb{G}_{n-p}} \sum_{0 \leq \ell < k \leq p} 2^{p-\ell} \mathcal{Q}^{p-k} (\tilde{f}_{k,n} \mathcal{Q}^{k-\ell} \tilde{f}_{\ell,n})(X_i),$$

$$V_6(n) = |\mathbb{G}_n|^{-1} \sum_{i \in \mathbb{G}_{n-p}} \sum_{0 \leq \ell < k < p} \sum_{r=0}^{p-k-1} 2^{p-\ell+r} \mathcal{Q}^{p-1-(r+k)} (\mathcal{P}(\mathcal{Q}^r \tilde{f}_{k,n} \otimes_{\text{sym}} \mathcal{Q}^{k-\ell+r} \tilde{f}_{\ell,n}))(X_i).$$

First, we set

$$h_{k,\ell,r}^{(n)} = 2^{r-\ell} \mathcal{Q}^{p-1-(r+k)} (\mathcal{P}(\mathcal{Q}^r \tilde{f}_{k,n} \otimes_{\text{sym}} \mathcal{Q}^{k-\ell+r} \tilde{f}_{\ell,n})) \quad \text{and} \quad H_{6,n} = \sum_{0 \leq \ell < k < p} \sum_{r=0}^{p-k-1} h_{k,\ell,r}^{(n)}$$

in such a way that  $V_6(n) = |\mathbb{G}_{n-p}|^{-1} M_{\mathbb{G}_{n-p}}(H_{6,n})$ . Using, (3) and (i) and (ii) of Assumption 3.8, we get

$$|h_{k,\ell,r}^{(n)}| \leq C h_n^d (2\alpha^2)^r \alpha^{k-\ell}.$$

This implies that  $|H_{6,n}| \leq c_n$  and then that  $V_6(n) \leq c_n$ , where the sequence  $(c_n, n \in \mathbb{N})$  is defined by

$$c_n = C h_n^d \mathbf{1}_{\{2\alpha^2 \leq 1\}} + C (2\alpha^2)^p h_n^d \mathbf{1}_{\{2\alpha^2 > 1\}}.$$

Since the sequence  $(c_n, n \in \mathbb{N})$  is deterministic and converges to 0, it follows, using Remark 2.8, that

$$V_6(n) \xrightarrow[b_n^2]{\text{superexp}} 0.$$

Next, for the term  $V_5(n)$ , we have for all  $k > \ell$ :

$$\begin{aligned} |2^{-\ell} \mathcal{Q}^{p-k} (\tilde{f}_{k,n} \mathcal{Q}^{k-\ell} \tilde{f}_{\ell,n})| &\leq 2^{-\ell} \mathcal{Q}^{p-k} (|\tilde{f}_{k,n}| |\mathcal{Q}^{k-\ell} \tilde{f}_{\ell,n}|) \leq C 2^{-\ell} h_n^{d/2} \alpha^{k-\ell} \mathcal{Q}^{p-k} (|\tilde{f}_{k,n}|) \\ &\leq C \alpha^p (2\alpha)^{-\ell} \mathbf{1}_{\{k=p\}} + C h_n^d (2\alpha)^{-\ell} \alpha^k \mathbf{1}_{\{k \leq p-1\}}, \end{aligned}$$

where we used (3) for the second inequality and (i) and (ii) of Assumption 3.8 for the second and the last inequality. Using the latter inequality in  $V_5(n)$ , we get

$$|V_5(n)| \leq C(2^{-p}\mathbf{1}_{\{2\alpha < 1\}} + \alpha^p\mathbf{1}_{\{2\alpha \geq 1\}} + h_n^d).$$

We thus have that  $V_5(n)$  is bounded by a deterministic sequence which converges to 0. It then follows from Remark 2.8 that

$$V_5(n) \xrightarrow[b_n^2]{\text{superexp}} 0.$$

From the foregoing, we get the result of Lemma since  $V_2(n) = V_5(n) + V_6(n)$ .  $\square$

As a direct consequence of (39) and the Lemmas 5.3, 5.4 and 5.5, we have the following result.

**Lemma 5.6.** *Under the Assumptions of Theorem 3.10, we have  $V(n) \xrightarrow[b_n^2]{\text{superexp}} \sigma^2$ .*

We now study the 4th-order exponential moment condition. We stress that this condition imply in particular the exponential Lindeberg condition (condition (C3) in Proposition 6.1). We have the following result.

**Lemma 5.7.** *Under the Assumptions of Theorem 3.10, we have*

$$\limsup_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(b_n^2 \sum_{i \in \mathbb{G}_{n-p}} \mathbb{E}[\Delta_{n,i}(f_n)^4 | \mathcal{F}_i] > \delta\right) = -\infty \quad \forall \delta > 0.$$

*Proof.* For all  $i \in \mathbb{G}_{n-p}$ , we have

$$(43) \quad \mathbb{E}[\Delta_{n,i}(f_n)^4 | \mathcal{F}_i] \leq 16(p+1)^3 2^{-2n} \sum_{\ell=0}^p \mathbb{E}_{X_i} \left[ M_{\mathbb{G}_{p-\ell}}(\tilde{f}_{\ell,n})^4 \right],$$

where we have used the definition of  $\Delta_{n,i}(f_n)$ , the inequality  $(\sum_{k=0}^r a_k)^4 \leq (r+1)^3 \sum_{k=0}^r a_k^4$  and the branching Markov property. Using (43), we get

$$(44) \quad b_n^2 \sum_{i \in \mathbb{G}_{n-p}} \mathbb{E}[\Delta_{n,i}(f_n)^4 | \mathcal{F}_i] \leq C b_n^2 p^3 2^{-2n} \sum_{\ell=0}^p \sum_{i \in \mathbb{G}_{n-p}} h_{n,\ell}(X_i),$$

where  $h_{n,\ell}(x) = \mathbb{E}_x[M_{\mathbb{G}_{p-\ell}}(\tilde{f}_{\ell,n})^4]$ . We will now prove that the right hand side of (44) converges superexponentially to 0 at the speed  $b_n^2$ , that is

$$\limsup_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(C b_n^2 p^3 2^{-2n} \left| \sum_{\ell=0}^p \sum_{i \in \mathbb{G}_{n-p}} h_{n,\ell}(X_i) \right| > \delta\right) = -\infty.$$

For that purpose, we will treat the case  $\ell = p$ ,  $\ell = p-1$  and finally the case  $\ell \in \{0, \dots, p-2\}$ . First, we treat the case  $\ell = p$ . Set  $g_{p,n} = \tilde{f}_{p,n}^4$ . We have

$$(45) \quad b_n^2 p^3 2^{-2n} \sum_{i \in \mathbb{G}_{n-p}} h_{n,p}(X_i) = b_n^2 p^3 2^{-2n} \sum_{i \in \mathbb{G}_{n-p}} \tilde{g}_{p,n}(X_i) + b_n^2 p^3 2^{-2n} |\mathbb{G}_{n-p}| \langle \mu, g_{p,n} \rangle.$$

Since  $b_n^2 p^3 2^{-2n} |\mathbb{G}_{n-p}| \langle \mu, g_{p,n} \rangle \leq p^3 2^{-p} b_n^2 (|\mathbb{G}_n| h_n^d)^{-1} \rightarrow 0$  as  $n \rightarrow \infty$ , it suffices to prove that the first term of the right hand side in (45) converges superexponentially to 0 at the speed  $b_n^2$ , that is, for all  $\delta > 0$ ,

$$(46) \quad \limsup_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(b_n^2 p^3 2^{-2n} \left| \sum_{i \in \mathbb{G}_{n-p}} \tilde{g}_{p,n}(X_i) \right| > \delta\right) = -\infty.$$

As in the proof of Lemma 5.2, we can prove that

$$\mathbb{P}\left(b_n^2 p^3 2^{-2n} \left| \sum_{i \in \mathbb{G}_{n-p}} \tilde{g}_{p,n}(X_i) \right| > \delta\right) \leq C \exp\left(-\frac{\delta^2 |\mathbb{G}_n|^2 h_n^{2d}}{C p^3 b_n^2 (\delta + C p^3 b_n^2 (|\mathbb{G}_{n+p}| h_n^d)^{-1})}\right).$$

Taking the log and dividing by  $b_n^2$ , we get (46).

Next, for  $\ell \in \{0, \dots, p-1\}$ , we plan to prove that the quantity  $b_n^2 p^3 2^{-2n} \sum_{\ell=0}^{p-1} \sum_{i \in \mathbb{G}_{n-p}} h_{n,\ell}(X_i)$  is bounded by a deterministic sequence which converges to 0. First, for  $\ell = p-1$ , using the branching Markov property, (i) and (ii) of Assumption 3.8, we have, for all  $i \in \mathbb{G}_{n-p}$ ,

$$h_{n,p-1}(X_i) = \mathbb{E}_{X_i}[M_{\mathbb{G}_1}(\tilde{f}_{p-1,n})^4] \leq C \mathcal{Q}(\tilde{f}_{p-1,n}^4) \leq C h_n^{-d}.$$

Using (6), this implies that

$$b_n^2 p^3 2^{-2n} \sum_{i \in \mathbb{G}_{n-p}} h_{n,p-1}(X_i) \leq C b_n^2 2^{-p} p^3 (|\mathbb{G}_n| h_n^d)^{-1} \rightarrow 0 \quad \text{as } n \rightarrow +\infty.$$

Now we consider the case  $\ell \in \{0, \dots, p-2\}$ . From Lemma 6.4 with  $f$  replaced by  $\tilde{f}_{\ell,n}$  and  $\nu$  by the Dirac mass at  $X_i$  ( $\delta_{X_i}$ ), we have

$$(47) \quad b_n^2 p^3 2^{-2n} \sum_{\ell=0}^{p-2} \sum_{i \in \mathbb{G}_{n-p}} h_{n,\ell}(X_i) \leq b_n^2 |\mathbb{G}_n|^{-2} p^3 \sum_{\ell=0}^{p-2} \sum_{i \in \mathbb{G}_{n-p}} \sum_{j=1}^9 |\psi_{j,p-\ell}|(X_i).$$

For all  $j \in \{1, \dots, 9\}$ , we will upper bound each term of the right hand side in (47) by a deterministic sequence which converges to 0.

**Upper bound of  $b_n^2 |\mathbb{G}_n|^{-2} p^3 \sum_{\ell=0}^{p-2} \sum_{i \in \mathbb{G}_{n-p}} |\psi_{1,p-\ell}|(X_i)$ .** Using (i) of Assumption 3.8, we have

$$|\psi_{1,p-\ell}| \leq C 2^{p-\ell} \mathcal{Q}^{p-\ell}(f_{\ell,n}^4) \leq C 2^{p-\ell} h_n^{-d}.$$

Using (6), this implies that

$$b_n^2 |\mathbb{G}_n|^{-2} p^3 \sum_{\ell=0}^{p-2} \sum_{i \in \mathbb{G}_{n-p}} |\psi_{1,p-\ell}|(X_i) \leq C b_n^2 p^3 (|\mathbb{G}_n| h_n^d)^{-1} \rightarrow 0 \quad \text{as } n \rightarrow +\infty.$$

**Upper bound of  $b_n^2 |\mathbb{G}_n|^{-3} p^3 \sum_{\ell=0}^{p-2} \sum_{i \in \mathbb{G}_{n-p}} |\psi_{2,p-\ell}|(X_i)$ .** Using Assumption 2.9 and (i) and (ii) of Assumption 3.8 for the second inequality, we get

$$\begin{aligned} |\psi_{2,p-\ell}| &\leq C 2^{2(p-\ell)} \sum_{k=0}^{p-\ell-1} 2^{-k} \mathcal{Q}^k \mathcal{P}(|\mathcal{Q}^{p-k-1-\ell}(\tilde{f}_{\ell,n}^3)| \otimes_{\text{sym}} |\mathcal{Q}^{p-\ell-k-2}(\mathcal{Q}\tilde{f}_{\ell,n})|) \\ &\leq C 2^{2(p-\ell)} \sum_{k=0}^{p-\ell-1} 2^{-k} \alpha^{p-\ell-k} \leq C 2^{p-\ell} (\mathbf{1}_{\{2\alpha < 1\}} + (p-\ell)\mathbf{1}_{\{2\alpha = 1\}} + (2\alpha)^{p-\ell} \mathbf{1}_{\{2\alpha > 1\}}). \end{aligned}$$

Using (6) and (5), this implies that

$$b_n^2 |\mathbb{G}_n|^{-2} p^3 \sum_{\ell=0}^{p-2} \sum_{i \in \mathbb{G}_{n-p}} |\psi_{2,p-\ell}|(X_i) \leq C b_n^2 |\mathbb{G}_n|^{-1} (p^4 \mathbf{1}_{\{2\alpha \leq 1\}} + (2\alpha)^p \mathbf{1}_{\{2\alpha > 1\}}) \rightarrow 0 \quad \text{as } n \rightarrow +\infty.$$

**Upper bound of  $b_n^2 |\mathbb{G}_n|^{-2} p^3 \sum_{\ell=0}^p \sum_{i \in \mathbb{G}_{n-p}} |\psi_{3,p-\ell}|(X_i)$ .** Using (i) and (ii) of Assumption 3.8 for the second inequality, we get

$$|\psi_{3,p-\ell}| \leq 2^{2(p-\ell)} \sum_{k=0}^{p-\ell-1} 2^{-k} \mathcal{Q}^k \mathcal{P}(\mathcal{Q}^{p-\ell-k-1}(\tilde{f}_{\ell,n}^2) \otimes^2) \leq C 2^{2(p-\ell)} \sum_{k=0}^{p-\ell-1} 2^{-k} \leq C 2^{2(p-\ell)}.$$

Using (6), this implies that

$$b_n^2 |\mathbb{G}_n|^{-2} p^3 \sum_{\ell=0}^p \sum_{i \in \mathbb{G}_{n-p}} |\psi_{3,p-\ell}|(X_i) \leq C b_n^2 p^3 2^{-n+p} \rightarrow 0 \quad \text{as } n \rightarrow +\infty.$$

**Upper bound of  $b_n^2 |\mathbb{G}_n|^{-2} p^3 \sum_{\ell=0}^p \sum_{i \in \mathbb{G}_{n-p}} |\psi_{4,p-\ell}|(X_i)$ .** Using Assumption (2.9) and (i) and (ii) of Assumption 3.8 for the second inequality, we get

$$|\psi_{4,p-\ell}| \leq C 2^{4(p-\ell)} \mathcal{P}(\mathcal{P}(|\mathcal{Q}^{p-\ell-2} \tilde{f}_{\ell,n}| \otimes^2) \otimes^2) \leq C 2^{4(p-\ell)} \alpha^{4(p-\ell-2)} h_n^{2d}.$$

Using (6) and (5), this implies that

$$\begin{aligned} & b_n^2 |\mathbb{G}_n|^{-2} p^3 \sum_{\ell=0}^p \sum_{i \in \mathbb{G}_{n-p}} |\psi_{4,p-\ell}|(X_i) \\ & \leq C (b_n^2 p^4 2^{-n-p} h_n^{2d} \mathbf{1}_{\{2\alpha^2 \leq 1\}} + b_n^2 p^3 2^{-n+p} (2\alpha^2)^{2p} h_n^{2d} \mathbf{1}_{\{2\alpha^2 > 1\}}) \rightarrow 0 \quad \text{as } n \rightarrow +\infty. \end{aligned}$$

**Upper bound of  $b_n^2 |\mathbb{G}_n|^{-3} p^3 \sum_{\ell=0}^p \sum_{i \in \mathbb{G}_{n-p}} |\psi_{5,p-\ell}|(X_i)$ .** Using Assumption (2.9) and (i) and (ii) of Assumption 3.8 for the second inequality, we get

$$\begin{aligned} |\psi_{5,p-\ell}| & \leq C 2^{4(p-\ell)} \sum_{k=2}^{p-\ell-1} \sum_{r=0}^{k-1} 2^{-2k-r} \mathcal{Q}^r \mathcal{P}(\mathcal{Q}^{k-r-1}(\mathcal{P}(|\mathcal{Q}^{p-\ell-k-1} \tilde{f}_{\ell,n}| \otimes^2)) \otimes^2) \\ & \leq C 2^{4(p-\ell)} \sum_{k=2}^{p-\ell-1} \sum_{r=0}^{k-1} 2^{-2k-r} h_n^{2d} \alpha^{4(p-\ell-k)} \\ & \leq C h_n^{2d} 2^{2(p-\ell)} (\mathbf{1}_{\{2\alpha^2 < 1\}} + (p-\ell) \mathbf{1}_{\{2\alpha^2 = 1\}} + (2\alpha^2)^{2(p-\ell)} \mathbf{1}_{\{2\alpha^2 > 1\}}). \end{aligned}$$

Using (6) and (5), this implies that

$$\begin{aligned} & b_n^2 |\mathbb{G}_n|^{-2} p^3 \sum_{\ell=0}^p \sum_{i \in \mathbb{G}_{n-p}} |\psi_{5,p-\ell}|(X_i) \\ & \leq C (b_n^2 p^4 2^{-n+p} h_n^{2d} \mathbf{1}_{\{2\alpha^2 \leq 1\}} + b_n^2 p^3 2^{-n+p} (2\alpha^2)^{2p} h_n^{2d} \mathbf{1}_{\{2\alpha^2 > 1\}}) \rightarrow 0 \quad \text{as } n \rightarrow +\infty. \end{aligned}$$

**Upper bound of  $b_n^2 |\mathbb{G}_n|^{-2} p^3 \sum_{\ell=0}^p \sum_{i \in \mathbb{G}_{n-p}} |\psi_{6,p-\ell}|(X_i)$ .** Using Assumption (2.9) and (i) and (ii) of Assumption 3.8 for the second inequality, we get

$$\begin{aligned} |\psi_{6,n}| & \leq C 2^{3(p-\ell)} \sum_{k=1}^{p-\ell-1} \sum_{r=0}^{k-1} 2^{-k-r} \mathcal{Q}^r \mathcal{P}(\mathcal{Q}^{k-r-1} \mathcal{P}(|\mathcal{Q}^{p-\ell-k-1} \tilde{f}_{\ell,n}| \otimes^2) \otimes_{\text{sym}} \mathcal{Q}^{p-\ell-r-1}(\tilde{f}_{\ell,n}^2)) \\ & \leq C 2^{3(p-\ell)} \sum_{k=1}^{p-\ell-1} \sum_{r=0}^{k-1} 2^{-k-r} h_n^d \alpha^{2(p-\ell-k)} \\ & \leq C h_n^d 2^{2(p-\ell)} (\mathbf{1}_{\{2\alpha^2 < 1\}} + (p-\ell) \mathbf{1}_{\{2\alpha^2 = 1\}} + (2\alpha^2)^{p-\ell} \mathbf{1}_{\{2\alpha^2 > 1\}}). \end{aligned}$$

Using (6) and (5), this implies that

$$\begin{aligned} b_n^2 |\mathbb{G}_n|^{-2} p^3 \sum_{\ell=0}^p \sum_{i \in \mathbb{G}_{n-p}} |\psi_{6,p-\ell}|(X_i) \\ \leq C (b_n^2 p^4 2^{-n+p} h_n^d \mathbf{1}_{\{2\alpha^2 \leq 1\}} + b_n^2 p^3 2^{-n+p} (2\alpha^2)^{2p} h_n^d \mathbf{1}_{\{2\alpha^2 > 1\}}) \rightarrow 0 \quad \text{as } n \rightarrow +\infty. \end{aligned}$$

**Upper bound of  $b_n^2 |\mathbb{G}_n|^{-2} p^3 \sum_{\ell=0}^p \sum_{i \in \mathbb{G}_{n-p}} |\psi_{7,p-\ell}|(X_i)$ .** In the same way as for  $\psi_{6,p-\ell}$ , we have

$$\begin{aligned} b_n^2 |\mathbb{G}_n|^{-2} p^3 \sum_{\ell=0}^p \sum_{i \in \mathbb{G}_{n-p}} |\psi_{7,p-\ell}|(X_i) \\ \leq C (b_n^2 p^4 2^{-n} h_n^d \mathbf{1}_{\{2\alpha^2 \leq 1\}} + b_n^2 p^3 2^{-n+p} (2\alpha^2)^{2p} h_n^d \mathbf{1}_{\{2\alpha^2 > 1\}}) \rightarrow 0 \quad \text{as } n \rightarrow +\infty. \end{aligned}$$

**Upper bound of  $b_n^2 |\mathbb{G}_n|^{-2} p^3 \sum_{\ell=0}^p \sum_{i \in \mathbb{G}_{n-p}} |\psi_{8,p-\ell}|(X_i)$ .** Using Assumption (2.9) and (i) and (ii) of Assumption 3.8 for the second inequality, we get

$$\begin{aligned} |\psi_{8,p-\ell}| &\leq C 2^{4(p-\ell)} \sum_{k=2}^{p-\ell-1} \sum_{r=1}^{k-1} \sum_{j=0}^{r-1} 2^{-k-r-j} Q^j \mathcal{P}(Q^{r-j-1} \mathcal{P}(|Q^{p-\ell-1-r} \tilde{f}_{\ell,n}| \otimes^2) \\ &\quad \otimes_{\text{sym}} Q^{k-j-1} \mathcal{P}(|Q^{p-\ell-1-k} \tilde{f}_{\ell,n}| \otimes^2)) \\ &\leq C 2^{4(p-\ell)} \sum_{k=2}^{p-\ell-1} \sum_{r=1}^{k-1} \sum_{j=0}^{r-1} 2^{-k-r-j} h_n^{2d} \alpha^{4(p-\ell)-2r-2k} \\ &\leq C h_n^{2d} 2^{2(p-\ell)} (\mathbf{1}_{\{2\alpha^2 < 1\}} + (p-\ell)^2 \mathbf{1}_{\{2\alpha^2 = 1\}} + (2\alpha^2)^{2(p-\ell)} \mathbf{1}_{\{2\alpha^2 > 1\}}). \end{aligned}$$

Using (6) and (5), this implies that

$$\begin{aligned} b_n^2 |\mathbb{G}_n|^{-2} p^3 \sum_{\ell=0}^p \sum_{i \in \mathbb{G}_{n-p}} |\psi_{8,p-\ell}|(X_i) \\ \leq C (b_n^2 p^5 2^{-n+p} h_n^{2d} \mathbf{1}_{\{2\alpha^2 \leq 1\}} + b_n^2 p^3 2^{-n+p} (2\alpha^2)^{2p} h_n^{2d} \mathbf{1}_{\{2\alpha^2 > 1\}}) \rightarrow 0 \quad \text{as } n \rightarrow +\infty. \end{aligned}$$

**Upper bound of  $b_n^2 |\mathbb{G}_n|^{-2} p^3 \sum_{\ell=0}^p \sum_{i \in \mathbb{G}_{n-p}} |\psi_{9,p-\ell}|(X_i)$ .** In the same way as for  $\psi_{8,p-\ell}$ , we have

$$\begin{aligned} b_n^2 |\mathbb{G}_n|^{-2} p^3 \sum_{\ell=0}^p \sum_{i \in \mathbb{G}_{n-p}} |\psi_{9,p-\ell}|(X_i) \\ \leq C (b_n^2 p^5 2^{-n+p} h_n^{2d} \mathbf{1}_{\{2\alpha^2 \leq 1\}} + b_n^2 p^3 2^{-n+p} (2\alpha^2)^{2p} h_n^{2d} \mathbf{1}_{\{2\alpha^2 > 1\}}) \rightarrow 0 \quad \text{as } n \rightarrow +\infty. \end{aligned}$$

Putting together all the upper bounds for  $\ell \in \{0, \dots, p-1\}$  and using (43) and (47), we deduce that  $b_n^2 p^3 2^{-2n} \sum_{\ell=0}^{p-1} \sum_{i \in \mathbb{G}_{n-p}} h_{n,\ell}(X_i)$  is bounded by a deterministic sequence which converges to 0. As a consequence, it follows, using Remark 2.8, that

$$\limsup_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \mathbb{P} \left( b_n^2 p^3 2^{-2n} \sum_{\ell=0}^{p-1} \sum_{i \in \mathbb{G}_{n-p}} h_{n,\ell}(X_i) > \delta \right) = -\infty.$$

Finally, using (43), (44), (46), we get

$$\limsup_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \mathbb{P} \left( \sum_{i \in \mathbb{G}_{n-p}} \mathbb{E}[\Delta_{n,i}(f_n)^4 | \mathcal{F}_i] > \frac{\delta}{b_n^2} \right) = -\infty \quad \forall \delta > 0.$$



□

For Chen-Ledoux type condition, we have the following result.

**Lemma 5.8.** *Under the assumptions of Theorem 3.10, we have*

$$\limsup_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \left( |\mathbb{G}_n| \sup_{i \in \mathbb{G}_{n-p}} \mathbb{P}_{\mathcal{F}_i} \left( |\Delta_{n,i}(\mathbf{f}_n)| > b_n \right) \right) = -\infty.$$

*Proof.* For all  $i \in \mathbb{G}_{n-p}$ , using (21) we have

$$(48) \quad \mathbb{P}_{\mathcal{F}_i} \left( |\Delta_{n,i}(\mathbf{f})| > b_n \sqrt{n} \right) \leq \mathbb{P}_{\mathcal{F}_i} \left( |N_{n,i}(\mathbf{f})| > b_n \sqrt{n}/2 \right) + \mathbb{P}_{\mathcal{F}_i} \left( |\mathbb{E}_{X_i} [N_{n,i}(\mathbf{f})]| > b_n \sqrt{n}/2 \right),$$

with  $N_{n,i}(\mathbf{f})$  defined in (22). Following the proof of (32), we get

$$\begin{aligned} \mathbb{P}_{\mathcal{F}_i} \left( |N_{n,i}(\mathbf{f})| > \frac{b_n \sqrt{n}}{2} \right) &= \mathbb{P}_{X_i} \left( \left| \sum_{\ell=0}^p M_{\mathbb{G}_{p-\ell}}(\tilde{f}_\ell, n) \right| > \frac{b_n \sqrt{|\mathbb{G}_n|}}{2} \right) \\ &\leq C \exp \left( - \frac{b_n^2 |\mathbb{G}_n|}{2(Cb_n |\mathbb{G}_n|^{1/2} h_n^{-d/2} + 6C |\mathbb{G}_p|)} \right). \end{aligned}$$

Next, for

$$\lambda = \frac{b_n \sqrt{|\mathbb{G}_n|}}{2(c_2 h_n^{-d/2} b_n \sqrt{n} + 3c_1 |\mathbb{G}_p|)},$$

we have

$$\begin{aligned} \mathbb{P}_{\mathcal{F}_i} \left( |\mathbb{E}_{X_i} [N_{n,i}(\mathbf{f})]| > \frac{b_n \sqrt{n}}{2} \right) &= \mathbb{P}_{X_i} \left( \sum_{\ell=0}^p 2^{p-\ell} Q^{p-\ell}(\tilde{f}_\ell)(X_i) > \frac{b_n \sqrt{n |\mathbb{G}_n|}}{2} \right) \\ &\leq \exp \left( - \frac{\lambda b_n \sqrt{n |\mathbb{G}_n|}}{2} \right) \mathbb{E}_{X_i} \left[ \exp \left( \lambda \sum_{\ell=0}^p 2^{p-\ell} Q^{p-\ell}(\tilde{f}_\ell)(X_i) \right) \right] \\ &\leq C \exp \left( - \frac{b_n^2 |\mathbb{G}_n|}{2(Cb_n |\mathbb{G}_n|^{1/2} h_n^{-d/2} + 6C |\mathbb{G}_p|)} \right), \end{aligned}$$

where we used (49) and the branching Markov property for the first equality, Chernoff bound for the first inequality and (3) for the last inequality. Doing the same thing for  $-\mathbf{f}$  instead of  $\mathbf{f}$ , we get

$$\mathbb{P}_{\mathcal{F}_i} \left( |\mathbb{E}_{X_i} [N_{n,i}(\mathbf{f})]| > \frac{b_n \sqrt{n}}{2} \right) \leq 2C \exp \left( - \frac{b_n^2 |\mathbb{G}_n|}{2(Cb_n |\mathbb{G}_n|^{1/2} h_n^{-d/2} + 6C |\mathbb{G}_p|)} \right).$$

From the foregoing, we get, using (48),

$$|\mathbb{G}_n| \sup_{i \in \mathbb{G}_{n-p}} \mathbb{P}_{\mathcal{F}_i} \left( |\Delta_{n,i}(\mathbf{f})| > b_n \sqrt{n} \right) \leq C |\mathbb{G}_n| \exp \left( - \frac{b_n^2 |\mathbb{G}_n|}{2(Cb_n |\mathbb{G}_n|^{1/2} h_n^{-d/2} + 6C |\mathbb{G}_p|)} \right).$$

Finally, taking the log and dividing by  $b_n^2$  in the latter inequality, we get the result of Lemma 5.8. □

We can now use Proposition 6.1 to deduce from Lemmas 5.6, 5.7 and 5.8 that  $\Delta_n(\mathbf{f}_n)$  satisfies a moderate deviation principle with speed  $b_n^2$  and rate function  $I$  defined by:  $I(x) = x^2/(2\sigma^2)$  for all  $x \in \mathbb{R}$ , with the finite variance  $\sigma^2$  defined in (7). Using (38) and Remark 2.7, we then deduce Theorem 3.10.

## 6. APPENDIX

We recall here a simplified version of Theorem 1 in [12]. We consider the real martingale  $(M_n, n \in \mathbb{N})$  with respect to the filtration  $(\mathcal{H}_n, n \in \mathbb{N})$  and we denote  $(\langle M \rangle_n, n \in \mathbb{N})$  its bracket.

**Proposition 6.1.** *Let  $(b_n)$  a sequence satisfying*

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

such that  $c(n) := \sqrt{n}/b_n$  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 for all  $\delta > 0$ ,*

$$\limsup_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \left( \mathbb{P} \left( \left| \frac{\langle M \rangle_n}{n} - Q \right| > \delta \right) \right) = -\infty,$$

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

(C3) *for all  $a > 0$  and for all  $\delta > 0$ ,*

$$\limsup_{n \rightarrow +\infty} \frac{1}{b_n^2} \log \left( \mathbb{P} \left( \frac{1}{n} \sum_{k=1}^n \mathbb{E} \left( |M_k - M_{k-1}|^2 \mathbf{1}_{\{|M_k - M_{k-1}| \geq a \frac{\sqrt{n}}{b_n}\}} | \mathcal{H}_{k-1} \right) > \delta \right) \right) = -\infty,$$

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

We have the following many-to-one formulas. Ideas of the proofs can be found in [16] and [3].

**Lemma 6.2.** *Let  $f, g \in \mathcal{B}(S)$ ,  $x \in S$  and  $n \geq m \geq 0$ . Assuming that all the quantities below are well defined, we have:*

$$(49) \quad \mathbb{E}_x [M_{\mathbb{G}_n}(f)] = |\mathbb{G}_n| \mathcal{Q}^n f(x) = 2^n \mathcal{Q}^n f(x),$$

$$(50) \quad \mathbb{E}_x [M_{\mathbb{G}_n}(f)^2] = 2^n \mathcal{Q}^n (f^2)(x) + \sum_{k=0}^{n-1} 2^{n+k} \mathcal{Q}^{n-k-1} (\mathcal{P}(\mathcal{Q}^k f \otimes \mathcal{Q}^k f))(x),$$

$$(51) \quad \mathbb{E}_x [M_{\mathbb{G}_n}(f) M_{\mathbb{G}_m}(g)] = 2^n \mathcal{Q}^m (g \mathcal{Q}^{n-m} f)(x) \\ + \sum_{k=0}^{m-1} 2^{n+k} \mathcal{Q}^{m-k-1} (\mathcal{P}(\mathcal{Q}^k g \otimes_{\text{sym}} \mathcal{Q}^{n-m+k} f))(x).$$

We recall the following result due to Bochner (see [21, Theorem 1A] which can be easily extended to any dimension  $d \geq 1$ ).

**Lemma 6.3.** *Let  $(h_n, n \in \mathbb{N})$  be a sequence of positive numbers converging to 0 as  $n$  goes to infinity. Let  $g : \mathbb{R}^d \rightarrow \mathbb{R}$  be a measurable function such that  $\int_{\mathbb{R}^d} |g(x)| dx < +\infty$ . Let  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  be a measurable function such that  $\|f\|_\infty < +\infty$ ,  $\int_{\mathbb{R}^d} |f(y)| dy < +\infty$  and  $\lim_{|x| \rightarrow +\infty} |x|f(x) = 0$ . Define*

$$g_n(x) = h_n^{-d} \int_{\mathbb{R}^d} f(h_n^{-1}(x-y))g(y)dy.$$

Then, we have at every point  $x$  of continuity of  $g$ ,

$$\lim_{n \rightarrow +\infty} g_n(x) = g(x) \int_{\mathbb{R}^d} f(y)dy.$$

We also give some bounds on  $\mathbb{E}_x [M_{\mathbb{G}_n}(f)^4]$ , see the proof of Theorem 2.1 in [3]. We will use the notation:

$$g \otimes^2 = g \otimes g.$$

**Lemma 6.4.** *There exists a finite constant  $C$  such that for all  $f \in \mathcal{B}(S)$ ,  $n \in \mathbb{N}$  and  $\nu$  a probability measure on  $S$ , assuming that all the quantities below are well defined, there exist functions  $\psi_j$ , for  $1 \leq j \leq 9$  such that:*

$$\mathbb{E}_\nu [M_{\mathbb{G}_n}(f)^4] = \sum_{j=1}^9 \langle \nu, \psi_j \rangle,$$

and, with  $h_k = \mathcal{Q}^{k-1}(f)$  and (notice that either  $|\psi_j|$  or  $|\langle \nu, \psi_j \rangle|$  is bounded), writing  $\nu g = \langle \nu, g \rangle$ :

$$\begin{aligned} |\psi_{1,n}| &\leq C 2^n \mathcal{Q}^n(f^4), \\ |\nu \psi_{2,n}| &\leq C 2^{2n} \sum_{k=0}^{n-1} 2^{-k} |\nu \mathcal{Q}^k \mathcal{P}(\mathcal{Q}^{n-k-1}(f^3) \otimes_{\text{sym}} h_{n-k})|, \\ |\psi_{3,n}| &\leq C 2^{2n} \sum_{k=0}^{n-1} 2^{-k} \mathcal{Q}^k \mathcal{P}(\mathcal{Q}^{n-k-1}(f^2) \otimes^2), \\ |\psi_{4,n}| &\leq C 2^{4n} \mathcal{P}(|\mathcal{P}(h_{n-1} \otimes^2) \otimes^2|), \\ |\psi_{5,n}| &\leq C 2^{4n} \sum_{k=2}^{n-1} \sum_{r=0}^{k-1} 2^{-2k-r} \mathcal{Q}^r \mathcal{P}(\mathcal{Q}^{k-r-1} |\mathcal{P}(h_{n-k} \otimes^2)| \otimes^2), \\ |\psi_{6,n}| &\leq C 2^{3n} \sum_{k=1}^{n-1} \sum_{r=0}^{k-1} 2^{-k-r} \mathcal{Q}^r |\mathcal{P}(\mathcal{Q}^{k-r-1} \mathcal{P}(h_{n-k} \otimes^2) \otimes_{\text{sym}} \mathcal{Q}^{n-r-1}(f^2))|, \\ |\nu \psi_{7,n}| &\leq C 2^{3n} \sum_{k=1}^{n-1} \sum_{r=0}^{k-1} 2^{-k-r} |\nu \mathcal{Q}^r \mathcal{P}(\mathcal{Q}^{k-r-1} \mathcal{P}(h_{n-k} \otimes_{\text{sym}} \mathcal{Q}^{n-k-1}(f^2)) \otimes_{\text{sym}} h_{n-r})|, \\ |\psi_{8,n}| &\leq C 2^{4n} \sum_{k=2}^{n-1} \sum_{r=1}^{k-1} \sum_{j=0}^{r-1} 2^{-k-r-j} \mathcal{Q}^j \mathcal{P}(|\mathcal{Q}^{r-j-1} \mathcal{P}(h_{n-r} \otimes^2)| \otimes_{\text{sym}} |\mathcal{Q}^{k-j-1} \mathcal{P}(h_{n-k} \otimes^2)|), \\ |\psi_{9,n}| &\leq C 2^{4n} \sum_{k=2}^{n-1} \sum_{r=1}^{k-1} \sum_{j=0}^{r-1} 2^{-k-r-j} \mathcal{Q}^j |\mathcal{P}(\mathcal{Q}^{r-j-1} |\mathcal{P}(h_{n-r} \otimes_{\text{sym}} \mathcal{Q}^{k-r-1} \mathcal{P}(h_{n-k} \otimes^2)) \otimes_{\text{sym}} h_{n-j})|. \end{aligned}$$

## REFERENCES

- [1] I. V. Basawa and J. Zhou. Non-Gaussian bifurcating models and quasi-likelihood estimation. *Adv. in Appl. Probab.*, 41(A):55–64, 2004.
- [2] S. V. Bitseki Penda and J.-F. Delmas. Central limit theorem for kernel estimator of invariant density in bifurcating markov chains models. *arXiv preprint arXiv:2106.08626*, 2021.
- [3] S. V. Bitseki Penda, H. Djellout, and A. Guillin. Deviation inequalities, moderate deviations and some limit theorems for bifurcating Markov chains with application. *Ann. Appl. Probab.*, 24(1):235–291, 2014.
- [4] S. V. Bitseki Penda and G. Gackou. Moderate deviation principles for bifurcating markov chains: case of functions dependent of one variable. *arXiv e-prints*, pages arXiv–2105, 2021.
- [5] S. V. Bitseki Penda, M. Hoffmann, and A. Olivier. Adaptive estimation for bifurcating Markov chains. *Bernoulli*, 23(4B):3598–3637, 2017.
- [6] S. V. Bitseki Penda and A. Olivier. Autoregressive functions estimation in nonlinear bifurcating autoregressive models. *Stat. Inference Stoch. Process.*, 20(2):179–210, 2017.

- [7] S. V. Bitseki Penda and A. Olivier. Moderate deviation principle in nonlinear bifurcating autoregressive models. *Statistics & Probability Letters*, 138:20–26, 2018.
- [8] S. V. Bitseki Penda and A. Roche. Local bandwidth selection for kernel density estimation in a bifurcating Markov chain model. *Journal of Nonparametric Statistics*, 0(0):1–28, 2020.
- [9] R. Cowan and R. Staudte. The bifurcating autoregression model in cell lineage studies. *Biometrics*, 42(4):769–783, December 1986.
- [10] J.-F. Delmas and L. Marsalle. Detection of cellular aging in a Galton-Watson process. *Stochastic Process. Appl.*, 120(12):2495–2519, 2010.
- [11] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Applications of mathematics. Springer, 1998.
- [12] H. Djellout. Moderate deviations for martingale differences and applications to  $\alpha$ -mixing sequences. *Stochastics and stochastics reports*, 73(1):37–64, 2002.
- [13] M. Dounic, M. Hoffmann, N. Krell, and L. Robert. Statistical estimation of a growth-fragmentation model observed on a genealogical tree. *Bernoulli*, 21(3):1760–1799, 2015.
- [14] M. Duflo. *Random iterative models*, volume 34. Springer Science & Business Media, 2013.
- [15] F. Gao. Moderate deviations and large deviations for kernel density estimators. *Journal of Theoretical Probability*, 16(2):401–418, 2003.
- [16] J. Guyon. Limit theorems for bifurcating Markov chains. Application to the detection of cellular aging. *Ann. Appl. Probab.*, 17(5-6):1538–1569, 2007.
- [17] M. Hoffmann and A. Marguet. Statistical estimation in a randomly structured branching population. *Stochastic Process. Appl.*, 129(12):5236–5277, 2019.
- [18] E. Masry. Recursive probability density estimation for weakly dependent stationary processes. *IEEE Transactions on Information Theory*, 32(2):254–267, 1986.
- [19] A. Mokkadem and M. Pelletier. Confidence bands for densities, logarithmic point of view. *Alea*, 2:231–266, 2006.
- [20] A. Mokkadem, M. Pelletier, and J. Worms. Large and moderate deviations principles for kernel estimation of a multivariate density and its partial derivatives. *Australian & New Zealand Journal of Statistics*, 47(4):489–502, 2005.
- [21] E. Parzen. On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, 33(3):1065–1076, 1962.
- [22] G. G. Roussas. Nonparametric estimation in Markov processes. *Annals of the Institute of Statistical Mathematics*, 21(1):73–87, 1969.

S. VALÈRE BITSEKI PENDA, IMB, CNRS-UMR 5584, UNIVERSITÉ BOURGOGNE FRANCHE-COMTÉ, 9 AVENUE ALAIN SAVARY, 21078 DIJON CEDEX, FRANCE.

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