Moderate deviations for stationary sequences of bounded random variables

Jérôme Dedecker [a], Florence Merlevède [b], Magda Peligrad [c] [1] *and* Sergey Utev [d]

[a] Université Paris 6, LSTA, 175 rue du Chevaleret, 75013 Paris, FRANCE

[b] Université Paris 6, LPMA and C.N.R.S UMR 7599, 175 rue du Chevaleret, 75013 Paris, FRANCE

[c] Department of Mathematical Sciences, University of Cincinnati, PO Box 210025, Cincinnati, Oh 45221-0025,

[d] School of Mathematical Sciences, University of Nottingham, Nottingham, NG7 2RD, UK

Key words: moderate deviations, martingale approximation, stationary processes.
Mathematical Subject Classification (2000): 60F10, 60G10.

Abstract

In this paper we derive the moderate deviation principle for stationary sequences of bounded random variables under martingale-type conditions. Applications to functions of ϕ-mixing sequences, contracting Markov chains, expanding maps of the interval, and symmetric random walks on the circle are given.

Résumé

Dans cet article, nous établissons un principe de déviation modérée pour des suites stationnaires de variables aléatoires bornées sous différentes conditions projectives. Nous appliquons ces résultats aux suites ϕ-mélangeantes, à certaines chaînes de Markov contractantes, aux transformations uniformément dilatantes de l'intervalle, ainsi qu'à la marche aléatoire symétrique sur le cercle.

1 Introduction

For the stationary sequence $(X_i)_{i\in\mathbf{Z}}$ of centered random variables, define the partial sums and the normalized partial sums process by

$$S_n = \sum_{j=1}^{n} X_j \text{ and } W_n(t) = n^{-1/2}\sum_{i=1}^{[nt]} X_i\,.$$

In this paper we are concerned with the moderate deviation principle for the normalized partial sums process W_n, considered as an element of $D([0,1])$ (functions on $[0,1]$ with left-hand limits and continuous from the right), equipped with the Skorohod topology (see Section 14 in Billingsley (1968) for the description of the topology on $D([0,1])$). More exactly, we say that the family of

[1]Supported in part by a Charles Phelps Taft Memorial Fund grant and NSA grant, H98230-07-1-0016.

random variables $\{W_n, n > 0\}$ satisfies the Moderate Deviation Principle (MDP) in $D[0,1]$ with speed $a_n \to 0$ and good rate function $I(.)$, if the level sets $\{x, I(x) \le \alpha\}$ are compact for all $\alpha < \infty$, and for all Borel sets

$$\begin{aligned} -\inf_{t\in\Gamma^0} I(t) &\le \liminf_n a_n \log \mathbf{P}(\sqrt{a_n}W_n \in \Gamma) \\ &\le \limsup_n a_n \log \mathbf{P}(\sqrt{a_n}W_n \in \Gamma) \le -\inf_{t\in\bar\Gamma} I(t)\,. \end{aligned} \tag{1}$$

The Moderate Deviation Principle is an intermediate behavior between the central limit theorem ($a_n = a$) and Large Deviations ($a_n = a/n$). 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.

De Acosta and Chen (1998) used the renewal theory to derive the MDP for bounded functionals of geometrically ergodic stationary Markov chains. Puhalskii (1994) and Dembo (1996) applied the stochastic exponential to prove the MDP for martingales. Starting from the martingale case and using the so-called coboundary decomposition due to Gordin (1969) ($X_k = M_k + Z_k - Z_{k+1}$, where M_k is a stationary martingale difference), Gao (1996) and Djellout (2002) obtained the MDP for ϕ-mixing sequences with summable mixing rate. In the context of Markov chains, the coboundary decomposition is known as the Poisson equation. Starting from this equation, Delyon, Juditsky and Liptser (2006) proved the MDP for $n^{-1/2}\sum_{k=1}^n H(Y_k)$, where H is a Lipschitz function, and $Y_n = F(Y_{n-1}, \xi_n)$, where F satisfies $|F(x,z) - F(y,t)| \le \kappa|x-y| + L|z-t|$ with $\kappa < 1$, and $(\xi_n)_{n\ge1}$ is an iid sequence of random variables independent of Y_0. In their paper, the random variables are not assumed to be bounded: the authors only assume that there exists a positive δ such that $\mathbf{E}(e^{\delta|\xi_1|}) < \infty$. They strongly used the Markov structure to derive some appropriate properties for the coboundary (see their lemma 4.2).

In this paper we propose a modification of the martingale approximation approach that allows to avoid the coboundary decomposition and thus to enlarge the class of dependent sequences known to satisfy the moderate deviation principle. Recent or new exponential inequalities are applied to justify the martingale approximation. The conditions involved in our results are well adapted to a large variety of examples, including regular functionals of linear processes, expanding maps of the interval and symmetric random walks on the circle.

The paper is organized as follows. In Section 2 we state the main results. A discussion of the conditions, clarifications, and some simple examples and extensions follow. Section 3 describes the applications, while Section 4 is dedicated to the proofs. Several technical lemmas are proved in the appendix.

2 Results

From now on, we assume that the stationary sequence $(X_i)_{i\in\mathbf{Z}}$ is given by $X_i = X_0 \circ T^i$, where $T : \Omega \mapsto \Omega$ is a bijective bimeasurable transformation preserving the probability $\mathbf{P}$ on $(\Omega, \mathcal{A})$. For a subfield $\mathcal{F}_0$ satisfying $\mathcal{F}_0 \subseteq T^{-1}(\mathcal{F}_0)$, let $\mathcal{F}_i = T^{-i}(\mathcal{F}_0)$. By $\|X\|_\infty$ we denote the $\mathbf{L}_\infty$-norm, that is the smallest u such that $\mathbf{P}(|X| > u) = 0$.

Our first theorem and its corollary treat the so-called adapted case, X_0 being $\mathcal{F}_0$–measurable and so the sequence $(X_i)_{i\in\mathbf{Z}}$ is adapted to the filtration $(\mathcal{F}_i)_{i\in\mathbf{Z}}$.

Theorem 1 *Assume that $\|X_0\|_\infty < \infty$ and that X_0 is $\mathcal{F}_0$–measurable. In addition, assume that*

$$\sum_{n=1}^{\infty} n^{-3/2}\|\mathbf{E}(S_n|\mathcal{F}_0)\|_\infty < \infty\,, \tag{2}$$

and that there exists $\sigma^2 \geq 0$ *with*

$$\lim_{n\to\infty} \|n^{-1}\mathbf{E}(S_n^2|\mathcal{F}_0) - \sigma^2\|_\infty = 0\,. \tag{3}$$

Then, for all positive sequences a_n *with* $a_n \to 0$ *and* $na_n \to \infty$*, the normalized partial sums processes* $W_n(.)$ *satisfy (1) with the good rate function* $I_\sigma(\cdot)$ *defined by*

$$I_\sigma(h) = \frac{1}{2\sigma^2}\int_0^1 (h'(u))^2 du \tag{4}$$

if simultaneously $\sigma > 0$*,* $h(0) = 0$ *and* h *is absolutely continuous, and* $I_\sigma(h) = \infty$ *otherwise.*

The following corollary gives simplified conditions for the MDP principle, which will be verified in several examples later on.

Corollary 2 *Assume that* $\|X_0\|_\infty < \infty$ *and that* X_0 *is* $\mathcal{F}_0$*–measurable. In addition, assume that*

$$\sum_{n=1}^{\infty} n^{-1/2}\|\mathbf{E}(X_n|\mathcal{F}_0)\|_\infty < \infty\,, \tag{5}$$

and that for all $i, j \geq 1$*,*

$$\lim_{n\to\infty} \|\mathbf{E}(X_iX_j|\mathcal{F}_{-n}) - \mathbf{E}(X_iX_j)\|_\infty = 0\,. \tag{6}$$

Then the conclusion of Theorem 1 holds with $\sigma^2 = \sum_{k\in\mathbf{Z}} \mathbf{E}(X_0X_k)$*.*

The next theorem allows to deal with non-adapted sequences and it provides additional applications. Let $\mathcal{F}_{-\infty} = \bigcap_{n\geq 0} \mathcal{F}_{-n}$ and $\mathcal{F}_\infty = \bigvee_{k\in\mathbf{Z}} \mathcal{F}_k$.

Theorem 3 *Assume that* $\|X_0\|_\infty < \infty$*,* $\mathbf{E}(X_0|\mathcal{F}_{-\infty}) = 0$ *almost surely, and* X_0 *is* $\mathcal{F}_\infty$*-measurable. Define the projection operators by* $P_j(X) = \mathbf{E}(X|\mathcal{F}_j) - \mathbf{E}(X|\mathcal{F}_{j-1})$ *. Suppose that (6) holds and that*

$$\sum_{j\in\mathbf{Z}} \|P_0(X_j)\|_\infty < \infty\,. \tag{7}$$

Then the conclusion of Theorem 1 holds with $\sigma^2 = \sum_{k\in\mathbf{Z}} \mathbf{E}(X_0X_k)$*.*

2.1 Simple examples, comments and extensions

Comment 4 ϕ**-mixing sequences**. Recall that if Y is a random variable with values in a Polish space $\mathcal{Y}$ and if $\mathcal{M}$ is a σ-field, the ϕ-mixing coefficient between $\mathcal{M}$ and $\sigma(Y)$ is defined by

$$\phi(\mathcal{M}, \sigma(Y)) = \sup_{A\in\mathcal{B}(\mathcal{Y})} \|\mathbf{P}_{Y|\mathcal{M}}(A) - \mathbf{P}_Y(A)\|_\infty\,. \tag{8}$$

For the sequence $(X_i)_{i\in\mathbf{Z}}$ and positive integer m , let $\phi_m(n) = \sup_{i_m>\ldots>i_1\geq n} \phi(\mathcal{M}_0, \sigma(X_{i_1}, \ldots, X_{i_m}))$ and let $\phi(k) = \phi_\infty(k) = \lim_{m\to\infty} \phi_m(k)$ be the usual ϕ–mixing coefficient. It follows from Corollary 2 that if the variables are bounded, the conclusion of Theorem 1 holds as soon as

$$\sum_{k>0} k^{-1/2}\phi_1(k) < \infty \text{ and } \lim_{k\to\infty} \phi_2(k) = 0\,. \tag{9}$$

The condition (9) improves on the one imposed by Gao (1996), that is $\sum_{k>0} \phi(k) < \infty$, to get the MDP for bounded random variables (see his Theorem 1.2).

Comment 5 Application to the functional LIL. Since the variables are bounded, under the assumptions of Theorem 1 or of Theorem 3, the MDP also holds in $C[0,1]$ for the Donsker process

$$D_n(t) = W_n(t) + n^{-1/2}(nt - [nt])X_{[nt]+1} \, .$$

Hence, if $\sigma^2 > 0$, it follows from the proof of Theorem 1.4.1 in Deuschel and Stroock (1989), that the process

$$\{(2\sigma^2 \log\log n)^{-1/2} D_n(t) : t \in [0,1]\} \tag{10}$$

satisfies the functional law of the iterated logarithm. To be more precise, if $\mathcal{S}$ denotes the subset of $C[0,1]$ consisting of all absolutely continuous functions with respect to the Lebesgue measure such that $h(0) = 0$ and $\int_0^1 (h'(t))^2 dt \leq 1$, then the process defined in (10) is relatively compact with a.s. limit set $\mathcal{S}$. In the case of bounded random variables, we then get new criteria to derive the functional LIL. In particular, the functional LIL holds for ϕ-mixing bounded random variables satisfying (9).

Comment 6 Linear processes. Let $(c_i)_{i\in\mathbf{Z}}$ be a sequence of real numbers in $\ell^1(\mathbf{Z})$ (absolutely summable). Define $X_k = \sum_{i\in\mathbf{Z}} c_i \varepsilon_{k-i}$ where $(\varepsilon_k)_{k\in\mathbf{Z}}$ is a strictly stationary sequence satisfying (6) and (7). Then, so does the sequence $(X_k)_{k\in\mathbf{Z}}$, and the conclusion of Theorem 3 holds. In particular, the result applies if ε_0 is $\mathcal{F}_0$-measurable, $\mathbf{E}(\varepsilon_1|\mathcal{F}_0) = 0$ and

$$\lim_{n\to\infty} \|\mathbf{E}(\varepsilon_0^2|\mathcal{F}_{-n}) - \mathbf{E}(\varepsilon_0^2)\|_\infty = 0 \, .$$

Comment 7 Non–mixing in the ergodic sense example. The following simple example shows that Theorem 1 is applicable to non-mixing in the ergodic theoretical sense sequences. Moreover it covers a strictly larger class of examples than its Corollary 2. For all $k \in \mathbf{Z}$, let $Q_{k+1} = -Q_k$ where $\mathbf{P}(Q_0 = \pm 1) = 1/2$ and $X_k = Q_k + Y_k$ where $(Y_k)_{k\in\mathbf{Z}}$ is an iid sequence of zero mean and bounded random variables, independent of Q_0. We can easily check that all the conditions of Theorem 1 hold while the conditions of Corollary 2 are not satisfied.

Comment 8 Stationary ergodic martingales that does not satisfy MDP. Let Y_k be the stationary discrete Markov chain with the state space $\mathbf{N}$ and the transition kernel given by $\mathbf{P}(Y_1 = j-1|Y_0 = j) = 1$ for all $j \geq 1$ and $\mathbf{P}(Y_1 = j|Y_0 = 0) = \mathbf{P}(\tau = j)$ for $j \in \mathbf{N}$ with $\mathbf{E}(\tau) < \infty$ and $\mathbf{P}(\tau = 1) > 0$ which implies that (Y_k) is ergodic. Let $X_k = \xi_k I_{(Y_k \neq 0)}$ where (ξ_k) is an iid sequence independent of (Y_k) and such that $\mathbf{P}(\xi_k = \pm 1) = 1/2$. Then X_k is a stationary ergodic martingale difference which is also a bounded function of an ergodic Markov chain. Straightforward computations show that if τ does not have a finite exponential moment then there exists a positive sequence $a_n \to 0$ with $na_n \to \infty$ for which (1) does not hold. Thus the MDP principle is not true in general for the stationary sequences satisfying (5) without a certain form of condition (3). A similar example was suggested in Djellout (2002, Remark 2.6).

Comment 9 On $\mathrm{Var}(S_n)$ **and Theorem 1.** Note that if $\sum_{n=1}^\infty n^{-3/2}\|\mathbf{E}(S_n|\mathcal{F}_0)\|_2 < \infty$, then, by Peligrad and Utev (2005)

$$\lim_{n\to\infty} \frac{\mathrm{Var}(S_n)}{n} = \sigma^2 = \mathbf{E}(X_1^2) + \sum_{j=0}^\infty 2^{-j}\mathbf{E}(S_{2^j}(S_{2^{j+1}} - S_{2^j}))\, .$$

On the other hand, we shall prove later on that Condition (2) along with (6) are sufficient for the validity of (3). Therefore the conclusion of Theorem 1 holds under (2) and (6) with σ^2 identified in this remark.

Comment 10 Sequences that are not strictly stationary. The proof of Theorem 3 is based on the exponential inequality from Lemma 22, that was established without stationarity assumption. Therefore, Theorem 3 admits various extensions to non-stationary sequences. The following slight generalization is motivated by the fixed design regression problem $Z_k = \theta q_k + X_k$, where the fixed design points are of the form $q_k = 1/g(k/n)$, the error process X_k is a stationary sequence and we analyze the error of the estimator $\hat{\theta} = n^{-1}\sum_{k=1}^{n} Z_k g(k/n)$. If $\{X_i\}_{i\in\mathbf{Z}}$ satisfies the conditions of Theorem 3, and if g is a Lipschitz function, then the process $W_n = \{n^{-1/2}\sum_{i=1}^{[nt]} g(i/n)X_i, t \in [0,1]\}$ satisfies (1) with the good rate function $J(\cdot)$ defined by

$$J(h) = \frac{1}{2\sigma^2}\int_0^1 \Big(\frac{h'(u)}{g(u)}\Big)^2 du\,, \quad \text{where} \quad \sigma^2 = \sum_{k\in\mathbf{Z}} \mathbf{E}(X_0 X_k).$$

The proof of this result is omitted. It can be done by following the proof of Theorem 3. To be more precise, we start by proving the MDP for the process $\overline{W}_n(t) = n^{-1/2}\sum_{i=1}^{[nv^{-1}(t)]} g(i/n)X_i$ where $v(t) = \sigma^2\int_0^t g^2(x)dx$. For $\overline{W}_n(.)$, the rate function is $I_\sigma(.)$ as in Theorem 1. To go back to the process $W_n(.)$, use the change-of-time $W_n = \overline{W}_n \circ v$.

3 Applications

In this Section we present applications to functions of ϕ-mixing processes, contracting Markov chains, expanding maps of the interval and symmetric random walks on the circle. The proofs are given in Section 4.

3.1 Functions of ϕ-mixing sequences

In this section, we are partly motivated by Djellout et al. (2006, Theorem 2.7), who have proved the MDP for

$$X_k = f(Y_k, \ldots, Y_{k-\ell}) - \mathbf{E}(f(Y_k, \ldots, Y_{k-\ell})) \;\text{ where }\; Y_k = \sum_{i\in\mathbf{Z}} c_i \varepsilon_{k-i} \tag{11}$$

In their Theorem 2.7, Djellout et al. (2006) assume that
(i) $(\varepsilon_i)_{i\in\mathbf{Z}}$ is an iid sequence;
(ii) (condition on c_i) the spectral density of Y_k is continuous on $[-\pi, \pi[$;
(iii) (condition on ε_0) ε_0 satisfies the so-called LSI condition, which implies that $\mathbf{E}(\exp(\delta\varepsilon_0^2)) < \infty$ for some positive δ, and that the distribution ε_0 has an absolutely continuous component with respect to the Lebesgue measure with a strictly positive density on the support of μ (see their condition (2.1));
(iv) (condition on f) the functions $\partial_i f$ are Lipschitz for $i = 0, \ldots, \ell$.

By applying our main results, we derive the Propositions 11 and 12 stated below. In the case where X_k is given by (11), the Proposition 11 will allow us to obtain the MDP for a large class of functions. However, we require a stronger condition than (ii), that is we assume that the sequence $(c_i)_{i\in\mathbf{Z}}$ is in $\ell_1(\mathbf{Z})$, and instead of (iii), we suppose that ε_0 takes its values in some compact intervall $[a, b]$ (this assumption cannot be compared to the LSI condition (iii)). Our method allows to link the regularity of f to the behavior of the coefficients $(c_i)_{i\in\mathbf{Z}}$ (in that case, the condition (16) given below means that $\sum_{i\in\mathbf{Z}} w_j(2(b-a)|c_i|) < \infty$ for any $j = 0, \ldots, \ell$, where w_j is the modulus of continuity of f with respect to the j-th coordinate). In addition, our innovations maybe dependent: more precisely, $(\varepsilon_i)_{i\in\mathbf{Z}}$ is assumed to be a stationary ϕ–mixing sequence.

We now describe our general results. Let $(\varepsilon_i)_{i\in\mathbf{Z}} = (\varepsilon_0 \circ T^i)_{i\in\mathbf{Z}}$ be a stationary sequence of ϕ-mixing random variables with values in a subset A of a Polish space $\mathcal{X}$. Starting from the definition (8), we denote by $\phi_\varepsilon(n)$ the coefficient $\phi_\varepsilon(n) = \phi(\sigma(\varepsilon_i, i \le 0), \sigma(\varepsilon_i, i \ge n))$.

Our first result is for non-adapted sequences, that is satisfying the representation (12) below. Let H be a function from $A^{\mathbf{Z}}$ to $\mathbf{R}$ satisfying the condition

$$C(A): \quad \text{for any } x, y \text{ in } A^{\mathbf{Z}}, \quad |H(x) - H(y)| \le \sum_{i\in\mathbf{Z}} \Delta_i \mathbf{1}_{x_i \ne y_i}, \quad \text{where } \sum_{i\in\mathbf{Z}} \Delta_i < \infty,$$

Define the stationary sequence $X_k = X_0 \circ T^k$ by

$$X_k = H((\varepsilon_{k-i})_{i\in\mathbf{Z}}) - \mathbf{E}(H((\varepsilon_{k-i})_{i\in\mathbf{Z}}))\,. \tag{12}$$

Note that X_k is bounded in view of $C(A)$.

Proposition 11 *Let $(X_k)_{k\in\mathbf{Z}}$ be defined by (12), for a function H satisfying $C(A)$. If $\sum_{k>0}\phi_\varepsilon(k)$ is finite, then the conclusion of Theorem 1 holds with $\sigma^2 = \sum_{k\in\mathbf{Z}} \mathbf{E}(X_0X_k)$.*

For adapted sequences, that is satisfying the representation (13) below, we can assume that H satisfies another type of condition. Let H be a function from $A^{\mathbf{N}}$ to $\mathbf{R}$ satisfying the condition

$$C'(A): \quad \text{for any } i \ge 0, \quad \sup_{x\in A^{\mathbf{N}}, y\in A^{\mathbf{N}}} |H(x) - H(x^{(i)}y)| \le R_i, \quad \text{where } R_i \text{ decreases to } 0,$$

the sequence $x^{(i)}y$ being defined by $(x^{(i)}y)_j = x_j$ for $j < i$ and $(x^{(i)}y)_j = y_j$ for $j \ge i$. Define the stationary sequence $X_k = X_0 \circ T^k$ by

$$X_k = H((\varepsilon_{k-i})_{i\in\mathbf{N}}) - \mathbf{E}(H((\varepsilon_{k-i})_{i\in\mathbf{N}}))\,. \tag{13}$$

Proposition 12 *Let $(X_k)_{k\in\mathbf{Z}}$ be defined by (13), for a function H satisfying $C'(A)$. If*

$$\sum_{\ell=1}^{\infty} R_\ell \sum_{k\ge\ell} \frac{\phi_\varepsilon(k-\ell)}{\sqrt{k}} < \infty\,, \tag{14}$$

then the conclusion of Theorem 1 holds with $\sigma^2 = \sum_{k\in\mathbf{Z}} \mathbf{E}(X_0X_k)$. In particular, the condition (14) holds as soon as

1. *$\sum_{k>0}\phi_\varepsilon(k) < \infty$ and $\sum_{k>0} k^{-1/2}R_k < \infty$.*
2. *$\sum_{k>0} R_k < \infty$ and $\sum_{k>0} k^{-1/2}\phi_\varepsilon(k) < \infty$.*

Application to functions of linear processes. Assume that ε_i takes its values in a compact interval $A = [a, b]$ of $\mathbf{R}$, and let $(c_i)_{i\in\mathbf{Z}}$ be a sequence of real numbers in $\ell^1(\mathbf{Z})$. Let $m = \inf_{x\in A^{\mathbf{Z}}} \sum_{i\in\mathbf{Z}} c_i x_i$ and $M = \sup_{x\in A^{\mathbf{Z}}} \sum_{i\in\mathbf{Z}} c_i x_i$. For a function f from $[m, M]^{\mathbf{Z}}$ to $\mathbf{R}$, let w_i be the modulus of continuity of f with respect to the i-th coordinate, that is

$$w_i(h) = \sup_{x\in[m,M]^{\mathbf{Z}}, t\in[m,M], |x_i - t|\le h} |f(x) - f(x^{(i,t)})|\,,$$

the sequence $x^{(i,t)}$ being defined by $x_j^{(i,t)} = x_j$ for $j \ne i$ and $x_i^{(i,t)} = t$. Assume that

$$\text{for any } x, y \text{ in } [m, M]^{\mathbf{Z}} \quad |f(x) - f(y)| \le \sum_{i\in\mathbf{Z}} w_i(|x_i - y_i|) < \infty\,.$$

Define the random variables $Y_k = \sum_{i\in\mathbf{Z}} c_i\varepsilon_{k-i}$, and let

$$X_k = f((Y_{k-i})_{i\in\mathbf{Z}}) - \mathbf{E}(f((Y_{k-i})_{i\in\mathbf{Z}}) \tag{15}$$

(note that (15) is a generalization of (11)). Clearly, X_k may be written as in (12), for a function H from $A^{\mathbf{Z}}$ to $\mathbf{R}$. Moreover, H satisfies $C(A)$ with $\Delta_i \le \sum_{\ell\in\mathbf{Z}} w_\ell(2(b-a)|c_{i-\ell}|)$ provided that

$$\sum_{i\in\mathbf{Z}}\sum_{\ell\in\mathbf{Z}} w_\ell(2(b-a)|c_i|) < \infty\,. \tag{16}$$

From Proposition 11, if $\sum_{k>0}\phi_\varepsilon(k) < \infty$ and if (16) holds, then the conclusion of Theorem 1 holds. In particular, the condition (16) holds as soon as there exist $(b_i)_{i\in\mathbf{Z}}$ in $\ell^1(\mathbf{Z})$ and α in $]0,1]$ such that $w_\ell(h) \le b_\ell|h|^\alpha$ and $\sum_{i\in\mathbf{Z}}|c_i|^\alpha < \infty$. Two simple examples of such functions are:

1. $f(x) = \sum_{i\in\mathbf{Z}} g_i(x_i)$ for some g_i such that $|g_i(x) - g_i(y)| \le b_i|x-y|^\alpha$ for any x, y in $[m, M]$.
2. $f(x) = \Pi_{i=p}^q h_i(x_i)$ for some h_i such that $|h_i(x) - h_i(y)| \le K_i|x-y|^\alpha$ for any x, y in $[m, M]$.

Now, assume that $c_i = 0$ for $i < 0$, so that $Y_k = \sum_{i\ge 0} c_i\varepsilon_{k-i}$. If f is in fact a function of x through x_0 only, we simply denote by $w = w_0$ its modulus of continuity over $[m, M]$. In that case $X_k = f(Y_k) - \mathbf{E}(Y_k)$ may be written as in (13) for a function H satisfying $C'(A)$ with $R_i \le w(2|b-a|\sum_{k\ge i}|c_k|)$. From item 1 of Proposition 12, if $\sum_{k>0}\phi_\varepsilon(k) < \infty$ and if

$$\sum_{n\ge 1} n^{-1/2} w\Big(2|b-a|\sum_{k\ge n}|c_k|\Big) < \infty\,, \tag{17}$$

then the conclusion of Theorem 1 holds. In particular, if $|c_i| \le C\rho^i$ for some $C > 0$ and $\rho \in]0,1[$, the condition (17) holds as soon as:

$$\int_0^1 \frac{w(t)}{t\sqrt{|\log t|}}dt < \infty\,.$$

Note that this condition is satisfied as soon as $w(t) \le D|\log(t)|^{-\gamma}$ for some $D > 0$ and some $\gamma > 1/2$. In particular, it is satisfied if f is α-Hölder for some $\alpha \in]0,1]$.

3.2 Contracting Markov chains

Let $(Y_n)_{n\ge 0}$ be a stationary Markov chain of bounded random variables with invariant measure μ and transition kernel K. Denote by $\|\cdot\|_{\infty,\mu}$ the essential supremum norm with respect to μ. Let Λ_1 be the set of 1-Lipschitz functions. Assume that the chain satisfies the two following conditions:

$$\text{there exist } C > 0 \text{ and } \rho \in]0,1[\text{ such that } \quad \sup_{g\in\Lambda_1}\|K^n(g) - \mu(g)\|_{\infty,\mu} \le C\rho^n\,, \tag{18}$$

$$\text{for any } f, g \in \Lambda_1 \text{ and any } m \ge 0 \quad \lim_{n\to\infty}\|K^n(fK^m(g)) - \mu(fK^m(g))\|_{\infty,\mu} = 0\,. \tag{19}$$

We shall see in the next proposition that if (18) and (19) are satisfied, then the MDP holds in $D[0,1]$ for the sequence

$$X_n = f(Y_n) - \mu(f) \tag{20}$$

as soon as the function f belongs to the class $\mathcal{L}$ defined below.

Definition 13 *Let $\mathcal{L}$ be the class of functions f from $\mathbf{R}$ to $\mathbf{R}$ such that $|f(x)-f(y)| \leq c(|x-y|)$, for some concave and non decreasing function c satisfying*

$$\int_0^1 \frac{c(t)}{t\sqrt{|\log t|}} dt < \infty \,. \tag{21}$$

Note that (21) holds if $c(t) \leq D|\log(t)|^{-\gamma}$ for some $D > 0$ and some $\gamma > 1/2$. In particular, $\mathcal{L}$ contains the class of functions from $[0,1]$ to $\mathbf{R}$ which are α-Hölder for some $\alpha \in]0,1]$.

Proposition 14 *Assume that the stationary Markov chain $(Y_n)_{n\geq 0}$ satisfies (18) and (19), and let X_n be defined by (20). If f belongs to $\mathcal{L}$, then the conclusion of Theorem 1 holds with*

$$\sigma^2 = \sigma^2(f) = \mu((f-\mu(f))^2) + 2\sum_{n>0} \mu(K^n(f)\cdot(f-\mu(f)))\,.$$

The proof of this proposition is based on the following lemma which has interest in itself.

Lemma 15 *Let $u_n = \sup_{g\in\Lambda_1} \|K^n(g) - \mu(g)\|_{\infty,\mu}$. Let f be a function from $\mathbf{R}$ to $\mathbf{R}$ such that $|f(x)-f(y)| \leq c(|x-y|)$ for some concave and non decreasing function c. Then*

$$\|K^n(f) - \mu(f)\|_{\infty,\mu} \leq c(u_n)\,.$$

Remark 16 *If $u_n \leq C\rho^n$ for a $C > 0$ and $\rho \in]0,1[$, and if $c(t) \leq D|\log(t)|^{-\gamma}$ for $D > 0$ and $\gamma > 0$, then*

$$\|K^n(f) - \mu(f)\|_{\infty,\mu} = O(n^{-\gamma})\,.$$

We now give two conditions under which (18) and (19) hold. Let $[a,b]$ be a compact interval in which lies the support of μ. For a Lipschitz function f, let $\mathrm{Lip}(f) = \sup_{x,y\in[a,b]} |f(x)-f(y)|/|x-y|$. The chain is said to be Lipschitz contracting if there exist $\kappa > 0$ and $\rho \in]0,1[$ such that

$$\mathrm{Lip}(K^n(f)) \leq \kappa\rho^n \mathrm{Lip}(f)\,. \tag{22}$$

Let BV be the class of bounded variation functions from $[a,b]$ to $\mathbf{R}$. For any $f \in BV$, denote by $\|df\|$ the total variation norm of the measure df: $\|df\| = \sup\{\int g df, \|g\|_\infty \leq 1\}$. The chain is said to be to be BV-contracting if there exist $\kappa > 0$ and $\rho \in [0,1[$ such that

$$\|dK^n(f)\| \leq \kappa\rho^n \|df\|\,. \tag{23}$$

It is easy to see that if either (22) or (23) holds, then (18) and (19) are satisfied (to see that the condition (23) implies (19), it suffices to note that it implies the same property for two BV functions f, g (see (52)), and that any Lipshitz function from $[a,b]$ to $\mathbf{R}$ can be uniformly approximated by BV functions).

Application to iterated random functions. The stationary bounded Markov chain $(Y_n)_{n\geq 0}$ with transition kernel K is one-step Lipschitz contracting if there exists $\rho \in]0,1[$ such that

$$\mathrm{Lip}(K(f)) \leq \rho \mathrm{Lip}(f)\,.$$

Note that if K is one-step Lipschitz contracting then (22) obviously holds with $\kappa = 1$. The one-step contraction is a very restrictive assumption. However, it is satisfied if $Y_n = F(Y_{n-1}, \varepsilon_n)$ for some iid sequence $(\varepsilon_i)_{i>0}$ independent of Y_0, and some function F such that

$$\|F(x,\varepsilon_1) - F(y,\varepsilon_1)\|_1 \leq \rho|x-y| \quad \text{for any } x, y \text{ in } \mathbf{R}. \tag{24}$$

Remark 17 *Under a more restrictive condition on F than (24), namely*

$$|F(x,z) - F(y,t)| \le \rho|x-y| + L|z-t|\,, \tag{25}$$

Delyon et al (2006) have proved the MDP for $X_n = f(Y_n) - \mu(f)$ when f is a lipschitz function. In their paper, the chain is not assumed to be bounded. It is only assumed that $\mathbf{E}(e^{\delta\varepsilon_1}) < \infty$ for some $\delta > 0$, which implies the same property for X_1 (for a smaller δ) by using the inequality (25).

Application to expanding maps. Let T be a map from $[0,1]$ to $[0,1]$ preserving a probability μ on $[0,1]$, and let

$$X_k = f \circ T^{n-k+1} - \mu(f)\,,\ W_n(t) = W_n(f,t) = n^{-1/2} \sum_{i=1}^{[nt]} (f \circ T^{n-i+1} - \mu(f))$$

Define the Perron-Frobenius operator K from $L^2([0,1],\mu)$ to $L^2([0,1],\mu)$ *via* the equality

$$\int_0^1 (Kh)(x) f(x) \mu(dx) = \int_0^1 h(x)(f \circ T)(x) \mu(dx)\,. \tag{26}$$

The map T is said to be BV-contracting if its Perron-Frobenius operator is BV-contracting, that is satisfies (23). As a consequence of Proposition 14, the following corollary holds.

Corollary 18 *If T is BV-contracting, and if f belongs to $BV \cup \mathcal{L}$, then the conclusion of Theorem 1 holds with*

$$\sigma^2 = \sigma^2(f) = \mu((f-\mu(f))^2) + 2\sum_{n>0} \mu(f \circ T^n \cdot (f - \mu(f)))\,.$$

Let us present a large class of BV-contracting maps. We shall say that T is uniformly expanding if it belongs to the class $\mathcal{C}$ defined in Broise (1996), Section 2.1 page 11. Recall that if T is uniformly expanding, then there exists a probability measure μ on $[0,1]$, whose density f_μ with respect to the Lebesgue measure is a bounded variation function, and such that μ is invariant by T. Consider now the more restrictive conditions:

(a) T is uniformly expanding.

(b) The invariant measure μ is unique and (T,μ) is mixing in the ergodic-theoretic sense.

(c) $\dfrac{1}{f_\mu}\mathbf{1}_{f_\mu>0}$ is a bounded variation function.

Starting from Proposition 4.11 in Broise (1996), one can prove that if T satisfies the assumptions (a), (b) and (c) above, then it is BV contracting (see for instance Dedecker and Prieur (2007), Section 6.3). Some well known examples of maps satisfying the conditions (a), (b) and (c) are:

1. $T(x) = \beta x - [\beta x]$ for $\beta > 1$. These maps are called β-transformations.
2. I is the finite union of disjoint intervals $(I_k)_{1\le k\le n}$, and $T(x) = a_k x + b_k$ on I_k, with $|a_k| > 1$.
3. $T(x) = a(x^{-1} - 1) - [a(x^{-1} - 1)]$ for some $a > 0$. For $a = 1$, this transformation is known as the Gauss map.

Remark 19 *The case where $f(x) = x$ (that is $X_n = T^n - \mu(T)$) has already been considered by Dembo and Zeitouni (1997). However, in this paper, the assumptions on T are more restrictive than the assumptions (a), (b) and (c) above. In particular, they assume that there is a finite partition $(I_j)_{1\le j\le m}$ of $[0,1]$ on which T restricted to I_k is C^1 and $\inf_{x\in I_k}|T'(x)| > 1$, so that their result does not cover the case of the Gauss map (Example 3 above).*

3.3 Symmetric random walk on the circle

Let K be the Markov kernel defined by

$$Kf(x) = \frac{1}{2}(f(x+a) + f(x-a))$$

on the torus $\mathbf{R}/\mathbf{Z}$, with a irrational in $[0,1]$. The Lebesgue-Haar measure m is the unique probability which is invariant by K. Let $(\xi_i)_{i\in\mathbf{Z}}$ be the stationary Markov chain with transition kernel K and invariant distribution m. Let

$$X_k = f(\xi_k) - m(f)\ ,\ W_n(t) = W_n(f,t) = n^{-1/2}\sum_{i=1}^{[nt]}(f(\xi_i) - m(f))\,. \tag{27}$$

From Derriennic and Lin (2001), Section 2, we know that the central limit theorem holds for $n^{-1/2}W_n(f,1)$ as soon as the series of covariances

$$\sigma^2(f) = m((f-m(f))^2) + 2\sum_{n>0} m(fK^n(f-m(f))) \tag{28}$$

is convergent, and that the limiting distribution is $\mathcal{N}(0,\sigma^2(f))$. In fact the convergence of the series in (28) is equivalent to

$$\sum_{k\in\mathbf{Z}^*} \frac{|\hat{f}(k)|^2}{d(ka,\mathbf{Z})^2} < \infty\,, \tag{29}$$

where $\hat{f}(k)$ are the Fourier coefficients of f. Hence, for any irrational number a, the criterion (29) gives a class of function f satisfying the central limit theorem, which depends on the sequence $((d(ka,\mathbf{Z}))_{k\in\mathbf{Z}^*}$. Note that a function f such that

$$\liminf_{k\to\infty} k|\hat{f}(k)| > 0\,, \tag{30}$$

does not satisfies (29) for any irrational number a. Indeed, it is well known from the theory of continued fraction that if p_n/q_n is the n-th convergent of a, then $|p_n - q_n a| < q_n^{-1}$, so that $d(ka,\mathbf{Z}) < k^{-1}$ for an infinite number of positive integers k. Hence, if (30) holds, then $|\hat{f}(k)|/d(ka,\mathbf{Z})$ does not even tend to zero as k tends to infinity.

Our aim in this section is to give conditions on f and on the properties of the irrational number a ensuring that the MDP holds in $D[0,1]$.

a is said to be badly approximable by rationals if for any positive ε, (31)
the inequality $d(ka,\mathbf{Z}) < |k|^{-1-\varepsilon}$ has only finitely many solutions for $k\in\mathbf{Z}$.

From Roth's theorem the algebraic numbers are badly approximable (cf. Schmidt (1980)). Note also that the set of badly approximable numbers in $[0,1]$ has Lebesgue measure 1.

In Section 5.3 of Dedecker and Rio (2006), it is proved that the condition (29) (and hence the central limit theorem for $n^{-1/2}W_n(f,1)$) holds for any badly approximable number a as soon as

$$\sup_{k\neq 0} |k|^{1+\varepsilon}|\hat{f}(k)| < \infty \quad \text{for some positive } \varepsilon. \tag{32}$$

Note that, in view of (30), one cannot take $\varepsilon = 0$ in the condition (32).

In fact, for badly approximable numbers, the condition (32) implies also the MDP in $D[0,1]$:

Proposition 20 *Suppose that a is badly approximable by rationals, i.e satisfies (31). If the function f satisfies (32), then the conclusion of Theorem 1 holds with $\sigma^2 = \sigma^2(f)$.*

Note that, under the same conditions, the process $\{W_n(f,t), t \in [0,1]\}$ satisfies the weak invariance principle in $D[0,1]$. Indeed, to prove Proposition 20, we show that the conditions of Corollary 2 are satisfied, but these conditions imply the weak invariance principle (see for instance Peligrad and Utev (2005)). From Comment 5, we also infer that the Donsker process defined in (10) satisfies the functional law of the iterated logarithm.

4 Proofs

Since the proofs of our results are mainly based on some exponential bounds for the deviation probability of the maximum of the partial sums for dependent variables, we present these inequalities, which have interest in themselves.

4.1 Exponential bounds for dependent variables

We state first the exponential bound from Proposition 2 in Peligrad, Utev and Wu (2007) that we are going to use in the proof of the main theorem.

Lemma 21 *Let $(X_i)_{i\in\mathbf{Z}}$ be a stationary sequence of random variables adapted to the filtration $(\mathcal{F}_i)_{i\in\mathbf{Z}}$. Then*

$$\mathbf{P}(\max_{1\le i\le n} |S_i| \ge t) \le 4\sqrt{e}\exp\big(-t^2/2n\big[\|X_1\|_\infty + 80\sum_{j=1}^{n} j^{-3/2}\|\mathbf{E}(S_j|\mathcal{F}_0)\|_\infty\big]^2\big)$$

In the next lemma, we bound the maximal exponential moment of the stationary sequence by using the projective criteria.

Lemma 22 *Let $\{Y_k\}_{k\in\mathbf{Z}}$ be a sequence of random variables such that for all j, $\mathbf{E}(Y_j|\mathcal{F}_{-\infty}) = 0$ almost surely and Y_j is $\mathcal{F}_\infty$-measurable. Define the projection operators by $P_j(X) = \mathbf{E}(X|\mathcal{F}_j) - \mathbf{E}(X|\mathcal{F}_{j-1})$. Assume that*

$$\|P_{k-j}(Y_k)\|_\infty \le p_j \quad \textit{and} \quad D := \sum_{j=-\infty}^{\infty} p_j < \infty$$

Let $\{g_k\ ,\ k\in\mathbf{N}\}$ be a sequence of numbers and define,

$$S_k = \sum_{i=1}^{k} g_i Y_i\ ,\ M_k = \max_{1\le j\le k} S_j\ ,\ G_n^2 = \sum_{i=1}^{n} g_i^2$$

Then,

$$\mathbf{E}\exp(tM_n) \le 4\exp(\tfrac{1}{2}G_n^2 D^2 t^2)\,.$$

In particular,

$$\mathbf{P}\Big(\max_{1\le k\le n} |S_k| \ge x) \le 8\exp\Big(-\frac{x^2}{2G_n^2 D^2}\Big)\,.$$

Proof. Start with the decomposition

$$Y_k = \sum_{j=-\infty}^{\infty} P_{k-j}(Y_k) = \sum_{j=-\infty}^{\infty} b_j P_{k-j}(Y_k)/b_j$$

where $b_j = p_j/D \geq \|P_{k-j}(Y_k)\|_\infty/D$, for any $j \in \mathbf{Z}$. Then

$$S_m = \sum_{j=-\infty}^{\infty} b_j \sum_{k=1}^{m} g_k P_{k-j}(Y_k)/b_j \,.$$

Thus,

$$M_n \leq \sum_{j=-\infty}^{\infty} b_j \max_{1\leq m\leq n} \sum_{k=1}^{m} P_{k-j}(g_k Y_k)/b_j =: \sum_{j=-\infty}^{\infty} b_j M_n^{(j)}$$

where $M_n^{(j)}$ denotes $\max_{1\leq m\leq n} \sum_{k=1}^{m} g_k P_{k-j}(Y_k)/b_j$.

Since $\exp(x)$ is convex and non–decreasing and $b_j \geq 0$ with $\sum_{j\in\mathbf{Z}} b_j = 1$,

$$\mathbf{E}\exp(tM_n) \leq \mathbf{E}\exp\Big(\sum_{j=-\infty}^{\infty} b_j t M_n^{(j)}\Big) \leq \sum_{j=-\infty}^{\infty} b_j \mathbf{E}\exp(tM_n^{(j)})\,.$$

Consider the martingale difference $U_k = g_k P_{k-j}(Y_k)/b_j$. Since $Z_k = \exp(t(U_1 + \cdots + U_k)/2)$ is a submartingale, Doob's inequality yields

$$\mathbf{E}\exp(tM_n^{(j)}) = \mathbf{E}\Big(\max_{1\leq k\leq n} Z_k^2\Big) \leq 4\mathbf{E}Z_n^2 = 4\mathbf{E}\exp(t(U_1 + \cdots + U_n))\,.$$

Applying Azuma's inequality to the right-hand side, and noting that

$$\|U_k\|_\infty = |g_k|\|P_{k-j}(Y_k)\|_\infty/b_j \leq |g_k|D\,,$$

we infer that

$$\mathbf{E}\exp(tM_n^{(j)}) \leq 4\exp(\tfrac{1}{2}G_n^2 D^2 t^2)\,.$$

Since $\sum_{j\in\mathbf{Z}} b_j = 1$, we obtain that

$$\mathbf{E}\exp(tM_n) \leq \sum_{j\in\mathbf{Z}} b_j 4\exp(\tfrac{1}{2}G_n^2 D^2 t^2) = 4\exp(\tfrac{1}{2}G_n^2 D^2 t^2)\,.$$

Next, to derive the one–sided probability inequality we use the exponential bound with $t = x/(G_n^2 D^2)$, so

$$\mathbf{P}(M_n \geq x) \leq \mathbf{E}\exp(tM_n)\exp(-tx) = 4\exp\Big(-\frac{x^2}{2G_n^2 D^2}\Big)\,.$$

Finally, to derive the two–sided inequality we observe that the stationary sequence $\{-Y_j\}$ also satisfies the conditions of the lemma. The proof is complete. ⋄

The next technical lemma provides an exponential bound for any random vector plus a correction in terms of conditional expectations (see also Wu, 1999).

Lemma 23 *Let $\{X_i\}_{1\leq i\leq n}$ be a vector of real random variables adapted to the filtration $\{\mathcal{F}_n\}_{n\geq 1}$. Denote B=$\sup_{1\leq i\leq n}\|X_i\|_\infty$. Then, for all $\delta > 0$ and c a natural number with $cB/n \leq \delta/2$, we have*

$$\mathbf{P}(\max_{1\leq i\leq n}|\frac{1}{n}\sum_{u=1}^{i}X_u| \geq \delta) \leq 2\exp(-\frac{\delta^2 n}{64B^2c}) + \mathbf{P}(\sup_{1\leq i\leq [n/c]}|\frac{1}{c}\sum_{j=(i-1)c+1}^{ic}\mathbf{E}(X_j|\mathcal{F}_{(i-1)c})| \geq \frac{\delta}{4}) \quad (33)$$

Proof of Lemma 23 Let c be a fixed integer and $k = [n/c]$ (where, as before, $[x]$ denotes the integer part of x). The initial step of the proof is to divide the variables in consecutive blocks of size c and to average the variables in each block

$$Y_{i,c} = \frac{1}{c}\sum_{j=(i-1)c+1}^{ic} X_j\,,\ i\geq 1\,.$$

Then, for all $1\leq i\leq k$ we construct the martingale,

$$M_{i,c} = \sum_{j=1}^{i}(Y_{j,c} - \mathbf{E}(Y_{j,c}|\mathcal{F}_{(j-1)c})) = \sum_{j=1}^{i} D_{j,c}$$

and we use the decomposition

$$\begin{aligned}\mathbf{P}(\max_{1\leq j\leq n}|\frac{1}{n}\sum_{u=1}^{j}X_u| &\geq \delta) \leq \mathbf{P}(\max_{1\leq i\leq k}|\frac{1}{k}\sum_{j=1}^{i}Y_{j,c}| \geq \delta - \frac{cB}{n}) \leq \mathbf{P}(\max_{1\leq i\leq k}|\frac{1}{k}\sum_{j=1}^{i}Y_{j,c}| \geq \delta/2)\\ &\leq \mathbf{P}(\max_{1\leq i\leq k}\frac{1}{k}|M_{i,c}| \geq \delta/4) + \mathbf{P}(\max_{1\leq i\leq k}\frac{1}{k}|\sum_{j=1}^{i}\mathbf{E}(Y_{j,c}|\mathcal{F}_{(j-1)c})| \geq \delta/4)\\ &\leq \mathbf{P}(\max_{1\leq i\leq k}\frac{1}{k}|M_{i,c}| \geq \delta/4) + \mathbf{P}(\max_{1\leq j\leq k}|\mathbf{E}(Y_{j,c}|\mathcal{F}_{(j-1)c})| \geq \delta/4)\,.\end{aligned}$$

Next, we apply Azuma's inequality to the martingale part and obtain,

$$\mathbf{P}\left(\max_{1\leq i\leq k}|M_{i,c}| \geq \delta k/4\right) \leq 2\exp(-\frac{\delta^2k^2}{32kB^2}) \leq 2\exp(-\frac{\delta^2 n}{64cB^2})$$

which implies that

$$\mathbf{P}(\max_{1\leq i\leq n}|\frac{1}{n}\sum_{u=1}^{i}X_u| \geq \delta) \leq 2\exp(-\frac{\delta^2 n}{64B^2c}) + \mathbf{P}(\max_{1\leq i\leq k}|\mathbf{E}(Y_{i,c}|\mathcal{F}_{(i-1)c})| \geq \delta/4)$$

proving the lemma. ◇

4.2 Some facts about the moderate deviation principle

This paragraph deals with some preparatory material. The following theorem is a result concerning the MDP for a triangular array of martingale differences sequences. It follows from Theorem 3.1 and Lemma 3.1 of Puhalskii (1994), (see also Djellout (2002), Proposition 1 and Lemma 2).

Lemma 24 *Let k_n be an increasing sequence of integers going to infinity. Let $\{D_{j,n}\}_{1\leq j\leq k_n}$ be a triangular array of martingale differences adapted to a filtration $\mathcal{F}_{j,n}$. Define the normalized partial sums process $Z_n(t) = n^{-1/2}\sum_{i=1}^{[k_nt]} D_{i,n}$. Let a_n be a sequence of real numbers such that $a_n \to 0$ and $na_n \to \infty$. Assume that $\|D_{j,n}\|_\infty = o(\sqrt{na_n})$ and that for all $\delta > 0$, there exists $\sigma^2 \geq 0$ such that*

$$\limsup_{n\to\infty} a_n \log \mathbf{P}\left(\Big|\frac{1}{n}\sum_{j=1}^{k_n}\mathbf{E}(D_{j,n}^2|\mathcal{F}_{(j-1),n}) - \sigma^2\Big| \geq \delta\right) = -\infty\,. \tag{34}$$

Then, for the given sequence a_n the partial sums processes $Z_n(.)$ satisfy (1) with the good rate function $I_\sigma(\cdot)$ defined in (4).

To be able to obtain the moderate deviation principle by approximation with martingales we state next a simple approximation lemma from Dembo and Zeitouni (1998, Theorem 4.2.13. p 130), called exponentially equivalence lemma.

Lemma 25 *Let $\xi_n(.) := \{\xi_n(t)\,, t \in [0,1]\}$ and $\zeta_n(.) := \{\zeta_n(t)\,, t \in [0,1]\}$ be two processes in $D([0,1])$. Assume that for any $\delta > 0$,*

$$\limsup_{n\to\infty} a_n \log \mathbf{P}(\sqrt{a_n}\sup_{t\in[0,1]}|\xi_n(t) - \zeta_n(t)| \geq \delta) = -\infty$$

Then, if the sequence of processes $\xi_n(.)$ satisfies (1) then so does the sequence of processes $\zeta_n(.)$.

In dealing with dependent random variables, to brake the dependence, a standard procedure is to divide first the variables in blocks. This technique introduces a new parameter, and so, in order to use a blocking procedure followed by a martingale approximation, we have to establish a more specific exponentially equivalent approximation, as stated in the following lemma:

Lemma 26 *For any positive integer m, let $k_{n,m}$ be an increasing sequence of integers going to infinity. Let $\{d_{j,n}^{(m)}\}_{1\leq j\leq k_{n,m}}$ be a sequence of triangular array of martingale differences adapted to a filtration $\mathcal{F}_{j,n}^{(m)}$. Define the normalized partial sums process $Z_n^{(m)}(t) = n^{-1/2}\sum_{i=1}^{[k_{n,m}t]} d_{i,n}^{(m)}$. Let a_n be a sequence of positive numbers such that $a_n \to 0$ and $na_n \to \infty$. Assume that for all $m \geq 1$*

$$\sup_{1\leq j\leq k_{n,m}} \|d_{j,n}^{(m)}\|_\infty = o(\sqrt{na_n}) \text{ as } n\to\infty \tag{35}$$

and that for all $\delta > 0$, there exists $\sigma^2 \geq 0$ such that

$$\lim_{m\to\infty}\limsup_{n\to\infty} a_n \log \mathbf{P}\left(\Big|\frac{1}{n}\sum_{j=1}^{k_{n,m}}\mathbf{E}((d_{j,n}^{(m)})^2|\mathcal{F}_{(j-1),n}^{(m)}) - \sigma^2\Big| \geq \delta\right) = -\infty\,. \tag{36}$$

Let $\{\zeta_n(t)\,, t \in [0,1]\}$ be a sequence of $D[0,1]$–valued random variables such that for all $\delta > 0$,

$$\lim_{m\to\infty}\limsup_{n\to\infty} a_n \log \mathbf{P}(\sqrt{a_n}\sup_{t\in[0,1]}|\zeta_n(t) - Z_n^{(m)}(t)| \geq \delta) = -\infty \tag{37}$$

Then, the processes $\zeta_n(.)$ satisfy (1) with the good rate function $I_\sigma(\cdot)$ defined in (4).

Proof. Define the functions

$$A_1(\delta,n,m)=a_n\log\mathbf{P}(\sup_{t\in[0,1]}|\zeta_n(t)-Z_n^{(m)}(t)|\geq\delta)\ ;$$

$$A_2(\delta,n,m)=a_n\log\mathbf{P}\left(\sup_{t\in[0,1]}\Big|\frac{1}{n}\sum_{j=1}^{[k_{n,m}t]}\mathbf{E}((d_{j,n}^{(m)})^2|\mathcal{F}_{(j-1),n}^{(m)})-t\sigma^2\Big|\geq\delta\right)$$

$$A_3(\delta,n,m)=\log(\sup_{1\leq j\leq k_{n,m}}\|d_{j,n}^{(m)}\|_\infty)-\log(\sqrt{a_n n})\ .$$

Observe that the functions $A_i, i=1,2,3$ satisfy the conditions of Lemma 30 from Appendix and so, we can find a sequence $m_n\to\infty$ such that the martingale difference sequence $(d_{j,n}^{(m_n)})$ satisfies the conditions of Lemma 24. We then derive that the sequence of processes $Z_n^{(m_n)}(.)$ satisfies (1) and, by applying Lemma 25, so does the sequence $\zeta_n(.)$. ◇

4.3 Proof of Theorem 1

Let m be an integer and $k=k_{n,m}=[n/m]$ (where, as before, $[x]$ denotes the integer part of x).

The initial step of the proof is to divide the variables in blocks of size m and to make the sums in each block

$$X_{i,m}=\sum_{j=(i-1)m+1}^{im}X_j\ ,\ i\geq 1.$$

Then we construct the martingales,

$$M_k^{(m)}=\sum_{i=1}^{[n/m]}(X_{i,m}-\mathbf{E}(X_{i,m}|\mathcal{F}_{(i-1)m}):=\sum_{i=1}^{[n/m]}D_{i,m}$$

and we define the process $\{M_k^{(m)}(t):\ t\in[0,1]\}$ by

$$M_k^{(m)}(t):=M_{[kt]}^{(m)}\ .$$

Now, we shall use Lemma 26 applied with $d_{j,n}^{(m)}=D_{j,m}$, and verify the conditions (36) and (37).

We start by proving (36). Notice first that $\{D_{i,m}\}_{i\geq 1}$ is a rowwise stationary sequence of bounded martingale differences. We have to verify

$$\lim_{m\to\infty}\limsup_{n\to\infty}a_n\log\mathbf{P}\left(\Big|\frac{1}{n}\sum_{j=1}^{[n/m]}\mathbf{E}(D_{j,m}^2|\mathcal{F}_{(j-1)m})-\sigma^2\Big|\geq\delta\right)=-\infty\,. \qquad (38)$$

Notice that

$$\mathbf{E}(D_{j,m}^2|\mathcal{F}_{(j-1)m})=\mathbf{E}(X_{j,m}^2|\mathcal{F}_{(j-1)m})-(\mathbf{E}(X_{j,m}|\mathcal{F}_{(j-1)m}))^2$$

and that, by stationarity

$$\frac{1}{n}\Big\|\sum_{j=1}^{[n/m]}(\mathbf{E}(X_{j,m}|\mathcal{F}_{(j-1)m}))^2\Big\|_\infty\leq\frac{\|\mathbf{E}(S_m|\mathcal{F}_0)\|_\infty^2}{m}\,.$$

Also

$$\Big\|\frac{1}{n}\sum_{j=1}^{[n/m]}\mathbf{E}(X_{j,m}^2|\mathcal{F}_{(j-1)m})-\sigma^2\Big\|_\infty\le\|m^{-1}\mathbf{E}(S_m^2|\mathcal{F}_0)-\sigma^2\|_\infty+(1-km/n)\sigma^2.$$

Consequently

$$\limsup_{n\to\infty}\Big\|\frac{1}{n}\sum_{j=1}^{[n/m]}(\mathbf{E}(D_{j,m}^2|\mathcal{F}_{(j-1)m})-\sigma^2\Big\|_\infty\le\frac{\|\mathbf{E}(S_m|\mathcal{F}_0)\|_\infty^2}{m}+\|m^{-1}\mathbf{E}(S_m^2|\mathcal{F}_0)-\sigma^2\|_\infty$$

which is smaller than $\delta/2$ provided m is large enough, by the first part of Lemma 29 from Appendix and condition (3). This proves (38).

It remains to prove (37), that means in our notation that for any $\delta>0$

$$\lim_{m\to\infty}\limsup_{n\to\infty}a_n\log\mathbf{P}\left(\sqrt{\frac{a_n}{n}}\sup_{t\in[0,1]}|S_{[nt]}-M_k^{(m)}(t)|\ge\delta\right)=-\infty\,. \tag{39}$$

Notice first that

$$\begin{aligned}\sup_{t\in[0,1]}|S_{[nt]}-M_k^{(m)}(t)|&\le\sup_{t\in[0,1]}|\sum_{i=[k\,t]m+1}^{[nt]}X_i|+\sup_{t\in[0,1]}|\sum_{i=1}^{[k\,t]}\mathbf{E}(X_{i,m}|\mathcal{F}_{(i-1)m})|\\&\le o(\sqrt{na_n})+\max_{1\le j\le[n/m]}|\sum_{i=1}^{j}\mathbf{E}(X_{i,m}|\mathcal{F}_{(i-1)m})|.\end{aligned}$$

Then, by using Lemma 21 we derive that

$$\begin{aligned}&a_n\log\mathbf{P}\left(\sqrt{\frac{a_n}{n}}\max_{1\le j\le[n/m]}|\sum_{i=1}^{j}\mathbf{E}(X_{i,m}|\mathcal{F}_{(i-1)m})|\ge\delta\right)\\&\quad\le a_n\log(4\sqrt{e})-\frac{\delta^2m}{2(\|\mathbf{E}(S_m|\mathcal{F}_0)\|_\infty+80\sum_{j=1}^\infty j^{-3/2}\|\mathbf{E}(S_{jm}|\mathcal{F}_0)\|_\infty)^2}\,.\end{aligned}$$

which is convergent to $-\infty$ when $n\to\infty$ followed by $m\to\infty$, by Lemma 29.

4.4 Proof of Corollary 2 and Remark 5

Notice that obviously, by triangle inequality and changing the order of summation, (5) implies (2). So, in order to establish both Corollary 2 and Remark 5, we just have to show that condition (2) together with (6) imply condition (3). This will be achieved by using the following two lemmas.

First let us introduce some notations. Let $S_{a,b}=S_b-S_a$ and set

$$\widetilde{\Delta}_{r,\infty}=\sum_{j=r}^{\infty}2^{-j/2}\|\mathbf{E}(S_{2^j}|\mathcal{F}_0)\|_\infty\,,\ \Delta_\infty=\|E(X_1^2|\mathcal{F}_0)\|_\infty^{1/2}+\sum_{j=0}^{\infty}2^{-j/2}\|\mathbf{E}(S_{2^j}|\mathcal{F}_0)\|_\infty\,.$$

By Peligrad and Utev (2005), $\widetilde{\Delta}_{0,\infty}<\infty$ is equivalent to (2).

Lemma 27 *Assume that X_0 is $\mathcal{F}_0$–measurable and that $\|\mathbf{E}(X_1^2|\mathcal{F}_0)\|_\infty < \infty$. Let n, r be integers such that $2^{r-1} < n \le 2^r$. Then*

$$\|\mathbf{E}(S_n^2|\mathcal{F}_0)\|_\infty \le n\Big(\|E(X_1^2|\mathcal{F}_0)\|_\infty^{1/2} + \tfrac{1}{2}\sum_{j=0}^{r-1} 2^{-j/2}\|\mathbf{E}(S_{2^j}|\mathcal{F}_0)\|_\infty\Big)^2 \le n\Delta_\infty^2\,.$$

Moreover, under (2),

$$\|n^{-1}\mathbf{E}(S_n^2|\mathcal{I}) - \eta\|_\infty \to 0 \quad as \quad n \to \infty\,,$$

where $\mathcal{I}$ is the σ-field of all T-invariant sets and

$$\eta = \mathbf{E}(X_1^2|\mathcal{I}) + \sum_{j=0}^{\infty} 2^{-j}\mathbf{E}(S_{2^j}(S_{2^{j+1}} - S_{2^j})|\mathcal{I})\,.$$

In particular, if $\mathbf{E}(X_iX_j|\mathcal{F}_{-\infty}) = \mathbf{E}(X_iX_j)$, for any i, j in $\mathbf{Z}$, then

$$\eta = \sigma^2 = \mathbf{E}(X_1^2) + \sum_{j=0}^{\infty} 2^{-j}\mathbf{E}(S_{2^j}(S_{2^{j+1}} - S_{2^j}))\,.$$

Proof. The proofs of the first three statements are almost identical to the proof of the corresponding facts in Proposition 2.1 of Peligrad and Utev (2005). The only changes are to replace everywhere the L_2–norm $\|x\|$ by the L_∞–norm $\|x\|_\infty$ and the usual expectation $\mathbf{E}(X)$ by the conditional expectation $\mathbf{E}(X) = \mathbf{E}(X|\mathcal{F}_0)$. The last statement follows from Proposition 2.12 in Bradley (2002), since for all i, j,

$$\mathbf{E}(X_iX_j|\mathcal{I}) = \mathbf{E}(\mathbf{E}(X_iX_j|\mathcal{F}_{-\infty})|\mathcal{I}) = \mathbf{E}(X_iX_j)\,. \quad \diamond$$

Lemma 28 *Assume that X_0 is $\mathcal{F}_0$–measurable and that $\|\mathbf{E}(X_1^2|\mathcal{F}_0)\|_\infty < \infty$. Suppose that the conditions (2) and (6) are satisfied. Then,*

$$\|n^{-1}\mathbf{E}(S_n^2|\mathcal{F}_0) - \sigma^2\|_\infty \to 0 \quad as \quad n \to \infty$$

Proof. By Lemma 27, it is enough to show that

$$\frac{1}{n}\|\mathbf{E}(S_n^2|\mathcal{F}_0) - \mathbf{E}(S_n^2)\|_\infty \to 0 \text{ as } n \to \infty$$

We prove this lemma by diadic recurrence. For t integer, denote

$$A_{t,k} = \|\mathbf{E}(S_t^2|\mathcal{F}_{-k}) - \mathbf{E}(S_t^2)\|_\infty\,.$$

Then, by the properties of conditional expectation and stationarity, for all $t \ge 1$

$$\begin{aligned} A_{2t,k} &= \|\mathbf{E}(S_{2t}^2|\mathcal{F}_{-k}) - \mathbf{E}(S_{2t}^2)\|_\infty \le \|\mathbf{E}(S_t^2|\mathcal{F}_{-k}) - \mathbf{E}(S_t^2)\|_\infty \\ &\quad + \|\mathbf{E}(S_{t,2t}^2|\mathcal{F}_{-k}) - \mathbf{E}(S_t^2)\|_\infty + 2\|\mathbf{E}(S_tS_{t,2t}|\mathcal{F}_{-k}) - \mathbf{E}(S_tS_{t,2t})\|_\infty \\ &\le 2\|\mathbf{E}(S_t^2|\mathcal{F}_{-k}) - \mathbf{E}(S_t^2)\|_\infty + 2\|\mathbf{E}(S_tS_{t,2t}|\mathcal{F}_{-k})\|_\infty + 2|\mathbf{E}(S_tS_{t,2t})|\,. \end{aligned}$$

Using for the last two terms the bound from Lemma 27, the Cauchy-Schwartz inequality and stationarity, we have

$$A_{2t,k} \le 2A_{t,k} + 4t^{1/2}\Delta_\infty\|\mathbf{E}(S_t|\mathcal{F}_0)\|_\infty$$

Whence, with the notation

$$B_{r,k} = 2^{-r}\|\mathbf{E}(S_{2^r}^2|\mathcal{F}_{-k}) - \mathbf{E}(S_{2^r}^2)\|_\infty = 2^{-r}A_{2^r,k}$$

by recurrence, for all $r \geq m$ and all $k > 0$, we derive

$$B_{r,k} \leq B_{r-1,k} + 2^{\frac{-r+3}{2}}\Delta_\infty\|\mathbf{E}(S_{2^{r-1}}|\mathcal{F}_0)\|_\infty \leq B_{m,k} + 2\Delta_\infty\sum_{j=m}^{r} 2^{-j/2}\|\mathbf{E}(S_{2^j}|\mathcal{F}_0)\|_\infty\,.$$

Therefore

$$2^{-r}\|\mathbf{E}(S_{2^r}^2|\mathcal{F}_{-k}) - \mathbf{E}(S_{2^r}^2)\|_\infty \leq B_{m,k} + 2\Delta_\infty\widetilde{\Delta}_{m,\infty}\,. \tag{40}$$

Now notice that, by stationarity and triangle inequality

$$\|\mathbf{E}(S_{2^r}^2|\mathcal{F}_0) - \mathbf{E}(S_{2^r}^2)\|_\infty \leq \|\mathbf{E}(S_{2^r}^2|\mathcal{F}_{-k}) - \mathbf{E}(S_{2^r}^2)\|_\infty + \|\mathbf{E}(S_{2^r}^2 - S_{k,k+2^r}^2|\mathcal{F}_0)\|_\infty\,, \tag{41}$$

and that by Lemma 27

$$\begin{aligned}\|\mathbf{E}(S_{2^r}^2 - S_{k,k+2^r}^2|\mathcal{F}_0)\|_\infty &\leq \|\mathbf{E}((S_{2^r} - S_{k,k+2^r})^2|\mathcal{F}_0)\|_\infty^{1/2}\|\mathbf{E}((S_{2^r} + S_{k,k+2^r})^2|\mathcal{F}_0)\|_\infty^{1/2} \\ &\leq 4k\|\mathbf{E}(X_1^2|\mathcal{F}_0)\|_\infty^{1/2}\|\mathbf{E}(S_{2^r}^2|\mathcal{F}_0)\|_\infty^{1/2} \\ &\leq 2^{2+r/2}k\|\mathbf{E}(X_1^2|\mathcal{F}_0)\|_\infty^{1/2}\Delta_\infty\,. \end{aligned} \tag{42}$$

Then, starting from (41) and using (40) and (42), we derive that for $r \geq m+1$,

$$2^{-r}\|\mathbf{E}(S_{2^r}^2|\mathcal{F}_0) - \mathbf{E}(S_{2^r}^2)\|_\infty \leq B_{m,k} + 2\Delta_\infty\widetilde{\Delta}_{m,\infty} + 2^{-r/2+2}k\|\mathbf{E}(X_1^2|\mathcal{F}_0)\|_\infty^{1/2}\Delta_\infty\,.$$

As a consequence

$$\limsup_{r\to\infty} 2^{-r}\|\mathbf{E}(S_{2^r}^2|\mathcal{F}_0) - \mathbf{E}(S_{2^r}^2)\|_\infty \leq B_{m,k} + 2\Delta_\infty\widetilde{\Delta}_{m,\infty}\,.$$

Then, we first let $k \to \infty$ and by Condition 6 it follows that $\lim_{k\to\infty} B_{m,k} = 0$. Then, we let m tend to infinity and by Condition (2), we derive

$$\lim_{r\to\infty} 2^{-r}\|\mathbf{E}(S_{2^r}^2|\mathcal{F}_0) - \mathbf{E}(S_{2^r}^2)\|_\infty = 0\,.$$

To complete the proof of the lemma we use the diadic expansion $n = \Sigma_{k=0}^{r-1}2^k a_k$ where $a_{r-1} = 1$ and $a_k \in \{0,1\}$ and continue the proof as in Proposition 2.1 in Peligrad and Utev (2005). $\diamond$

4.5 Proof of Theorem 3

Fix a positive integer m and define the stationary sequence

$$\xi_{j,m} := \mathbf{E}(X_j|\mathcal{F}_{j+m-1}) - \mathbf{E}(X_j|\mathcal{F}_{j-m})$$

Using a standard martingale decomposition (see also Hall and Heyde, 1980), we define

$$\theta_{j,m} = \sum_{t=0}^{\infty}\mathbf{E}(\xi_{j+t,m}|\mathcal{F}_{j+m-1}) = \sum_{k=0}^{2m-2}\mathbf{E}(\xi_{j+k,m}|\mathcal{F}_{j+m-1})\,.$$

and observe that

$$\|\theta_{0,m}\|_\infty = \| \sum_{k=0}^{2m-2} \sum_{i=k-m+1}^{m-1} P_i(X_k)\|_\infty \le 2m \sum_{i\in\mathbf{Z}} \|P_0(X_i)\|_\infty < \infty \,. \tag{43}$$

Then, $\mathbf{E}(\theta_{j+1,m}|\mathcal{F}_{j+m-1}) = \theta_{j,m} - \xi_{j,m}$ and thus,

$$\sum_{j=1}^{k} \xi_{j,m} = \theta_{1,m} - \theta_{k+1,m} + \sum_{j=1}^{k} d_{j,m} \,. \tag{44}$$

where $d_{j,m} := \theta_{j+1,m} - \mathbf{E}(\theta_{j+1,m}|\mathcal{F}_{j+m-1})$ is a stationary bounded martingale difference.

Moreover,

$$\sum_{j=1}^{k} X_j = \sum_{j=1}^{k} d_{j,m} + R_{k,m} \,, \tag{45}$$

where

$$R_{k,m} := \theta_{1,m} - \theta_{k+1,m} + \sum_{j=1}^{k}[X_j - E(X_j|\mathcal{F}_{j+m-1}) + \mathbf{E}(X_j|\mathcal{F}_{j-m})] \,.$$

First, we show that $R_{k,m}$ is negligible for the moderate deviation principle. We notice that by (43) it is enough to establish that

$$R'_{k,m} := \sum_{j=1}^{k}[X_j - E(X_j|\mathcal{F}_{j+m-1}) + \mathbf{E}(X_j|\mathcal{F}_{j-m})]$$

is negligible. Observe that

$$X_j - E(X_j|\mathcal{F}_{j+m-1}) + \mathbf{E}(X_j|\mathcal{F}_{j-m}) = \sum_{|t|\ge m} P_{j-t}(X_j) \quad \text{and}$$

$$\sum_{j\in\mathbf{Z}} \|P_0\big(X_j - E(X_j|\mathcal{F}_{j+m-1}) + \mathbf{E}(X_j|\mathcal{F}_{j-m})\big)\|_\infty \le \sum_{|k|\ge m}^{\infty} \|P_0(X_k)\|_\infty =: D_m \tag{46}$$

Now, the exponential inequality given in Lemma 22 entails that

$$\mathbf{P}\Big(\max_{1\le k\le n}\Big|\sum_{j=1}^{k}\mathbf{E}[X_j - E(X_j|\mathcal{F}_{j+m-1}) + \mathbf{E}(X_j|\mathcal{F}_{j-m})]\Big| \ge \delta\sqrt{n/a_n}\Big) \le 8\exp\Big(-\frac{\delta^2 n}{a_n 2n D_m^2}\Big)$$

The last inequality together with (7) and Lemma 25 reduces the theorem to the MDP principle for bounded stationary martingale difference $\{d_{j,m}\,;\, j\in\mathbf{Z}\}$.

Then, by Lemma 26, it remains to verify that

$$\lim_{m\to\infty}\limsup_{n\to\infty} a_n \ln \mathbf{P}\Big(\Big|\frac{1}{n}\sum_{j=1}^{n}(\mathbf{E}(d_{j,m}^2|\mathcal{F}_{j+m-1}) - \sigma^2)\Big| \ge \delta\Big) = -\infty \,.$$

In order to prove this convergence, by Lemma 23, applied with $B = 2\big(\sum_{\ell\in\mathbf{Z}}\|P_0(X_\ell)\|_\infty\big)^2$, it is enough to establish that

$$\lim_{m\to\infty}\limsup_{n\to\infty}\Big\|\frac{1}{n}\sum_{j=1}^{n}(\mathbf{E}(d_{j,m}^2|\mathcal{F}_{m-1}) - \sigma^2)\Big\|_\infty = 0 \,.$$

Since $\{d_{j,m}\}$ is a martingale difference, it follows from the decomposition (44) and (43), that it remains to prove that

$$\lim_{m\to\infty}\limsup_{n\to\infty}\Big\|\frac{1}{n}\mathbf{E}\Big(\Big(\sum_{j=2m-1}^{n}\xi_{j,m}\Big)^2|\mathcal{F}_{m-1}\Big)-\sigma^2\Big\|_\infty=0\,. \tag{47}$$

Write

$$\Big(\sum_{j=2m-1}^{n}\xi_{j,m}\Big)^2=\sum_{i=2m-1}^{n}\xi_{i,m}^2+2\sum_{i=2m-1}^{n}\sum_{j=i+1}^{(N+i)\wedge n}\xi_{i,m}\xi_{j,m}+2\sum_{i=2m-1}^{n}\sum_{j=N+i+1}^{n}\xi_{i,m}\xi_{j,m}\,.$$

Notice that, since $\xi_{j,m}=\Sigma_{k=j-m+1}^{j+m-1}P_k(X_j)$, we get

$$\frac{1}{n}\|\sum_{i=2m-1}^{n}\sum_{j=N+i+1}^{n}\mathbf{E}\big(\xi_{i,m}\xi_{j,m}|\mathcal{F}_{m-1}\big)\|_\infty\le\frac{1}{n}\sum_{i=2m-1}^{n}\sum_{j=N+i+1}^{\infty}\|\mathbf{E}\big(\xi_{i,m}\xi_{j,m}|\mathcal{F}_{m-1}\big)\|_\infty$$
$$\le\frac{1}{n}\sum_{i=2m-1}^{n}\sum_{k=i-m+1}^{i+m-1}\|P_k(X_i)\|_\infty\sum_{\ell\ge N}\|P_k(X_{i+\ell})\|_\infty\ \le\ \sum_{i\in\mathbf{Z}}\|P_0(X_i)\|_\infty\sum_{|\ell|\ge N/2}\|P_0(X_\ell)\|_\infty\to 0$$

as $N\to\infty$, uniformly in n, and so, (47) is implied by

$$\lim_{m\to\infty}\limsup_{n\to\infty}\Big\|\frac{1}{n}\mathbf{E}\Big(\sum_{i=2m-1}^{n}\xi_{i,m}^2+2\sum_{i=2m-1}^{n}\sum_{j=i+1}^{(N+i)\wedge n}\xi_{i,m}\xi_{j,m}|\mathcal{F}_{m-1}\Big)-\sigma_N^2\Big\|_\infty=0\,, \tag{48}$$

where $\sigma_N^2=\mathbf{E}(X_0^2)+2\mathbf{E}(X_0X_1)+\cdots+2\mathbf{E}(X_0X_{N-1})$. Write $\xi_{i,m}=X_i+(\xi_{i,m}-X_i)$. By condition (6), we easily get that

$$\lim_{m\to\infty}\limsup_{n\to\infty}\Big\|\frac{1}{n}\mathbf{E}\Big(\sum_{i=2m-1}^{n}X_i^2+2\sum_{i=2m-1}^{n}\sum_{j=i+1}^{(N+i)\wedge n}X_iX_j|\mathcal{F}_{m-1}\Big)-\sigma_N^2\Big\|_\infty=0\,,$$

hence (48) holds since

$$\|X_i-\xi_{i,m}\|_\infty\le\sum_{|k|\ge m}\|P_0(X_k)\|_\infty\to 0\quad\text{as}\quad m\to\infty\,.\quad\diamond$$

4.6 Proof of Proposition 11

Let $\mathcal{F}_k=\sigma(\varepsilon_i,i\le k)$. From Theorem 4.4.7 in Berbee (1979), there exists $(\varepsilon_i')_{i>0}$ distributed as $(\varepsilon_i)_{i>0}$ and independent of $\mathcal{F}_0$ such that

$$\|\mathbf{E}(\mathbf{1}_{\{\varepsilon_k\neq\varepsilon_k',\text{ for some }k\ge n\}}|\mathcal{F}_0)\|_\infty=\phi_\varepsilon(n)\,.$$

Let $(\varepsilon_i^{(0)})_{i\in\mathbf{Z}}$ be the sequence defined by $\varepsilon_i^{(0)}=\varepsilon_i$ if $i\le 0$ and $\varepsilon_i^{(0)}=\varepsilon_i'$ if $i>0$. Let $(\varepsilon_i^{(-1)})_{i\in\mathbf{Z}}$ be the sequence defined by $\varepsilon_i^{(-1)}=\varepsilon_i$ if $i<0$, $\varepsilon_i^{(-1)}=\varepsilon_i'$ if $i>0$ and $\varepsilon_0^{(-1)}=x$ where $x\in A$. Define now $Z_k=H((\varepsilon_{k-i})_{i\in\mathbf{Z}})$, $Z_k^{(0)}=H((\varepsilon_{k-i}^{(0)})_{i\in\mathbf{Z}})$ and $Z_k^{(-1)}=H((\varepsilon_{k-i}^{(-1)})_{i\in\mathbf{Z}})$. We shall apply Theorem 3. Note first that (7) is equivalent to $\sum_{i\in\mathbf{Z}}\|P_0(Z_i)\|_\infty<\infty$. Now

$$P_0(Z_i)=\mathbf{E}(Z_i^{(0)}|\mathcal{F}_0)-\mathbf{E}(Z_i^{(-1)}|\mathcal{F}_{-1})+\mathbf{E}(Z_i-Z_i^{(0)}|\mathcal{F}_0)-\mathbf{E}(Z_i-Z_i^{(-1)}|\mathcal{F}_{-1})\,.$$

Denoting by $\mathbf{E}_\varepsilon(\cdot)$ the conditional expectation with respect to ε, we infer from $C(A)$ that

$$|\mathbf{E}(Z_i^{(0)}|\mathcal{F}_0) - \mathbf{E}(Z_i^{(-1)}|\mathcal{F}_{-1})| = |\mathbf{E}_\varepsilon(H((\varepsilon_{i-j}^{(0)})_{j\in\mathbf{Z}}) - H((\varepsilon_{i-j}^{(-1)})_{j\in\mathbf{Z}}))| \leq \Delta_i\,.$$

Now, from $C(A)$ again,

$$|\mathbf{E}(Z_i - Z_i^{(0)}|\mathcal{F}_0)| \leq \sum_{k=1}^{\infty}\Delta_{i-k}\mathbf{E}(\mathbf{1}_{\varepsilon_k\neq\varepsilon_k'}|\mathcal{F}_0) \text{ and } \mathbf{E}(Z_i - Z_i^{(-1)}|\mathcal{F}_{-1}) \leq \Delta_i + \sum_{k=1}^{\infty}\Delta_{i-k}\mathbf{E}(\mathbf{1}_{\varepsilon_k\neq\varepsilon_k'}|\mathcal{F}_{-1}).$$

Consequently, by the ϕ-mixing property, we obtain the upper bound

$$\sum_{i\in\mathbf{Z}}\|P_0(Z_i)\|_\infty \leq 2\sum_{i\in\mathbf{Z}}\Delta_i + 2\sum_{i\in\mathbf{Z}}\sum_{k=1}^{\infty}\Delta_{i-k}\phi_\varepsilon(k)\,,$$

which is finite provided that $\sum_{i\in\mathbf{Z}}\Delta_i < \infty$ and $\sum_{k>0}\phi_\varepsilon(k) < \infty$. It remains to prove (6). Let $X_k^{(0)} = Z_k^{(0)} - \mathbf{E}(Z_k^{(0)})$. We have

$$\begin{aligned}\|\mathbf{E}(X_kX_l|\mathcal{F}_0) - \mathbf{E}(X_kX_l)\|_\infty &\leq \|\mathbf{E}(X_k^{(0)}X_l^{(0)}|\mathcal{F}_0) - \mathbf{E}(X_k^{(0)}X_l^{(0)})\|_\infty \\ &+ \|\mathbf{E}(X_k(X_l - X_l^{(0)})|\mathcal{F}_0) - \mathbf{E}(X_k(X_l - X_l^{(0)}))\|_\infty \\ &+ \|\mathbf{E}(X_l^{(0)}(X_k - X_k^{(0)})|\mathcal{F}_0) - \mathbf{E}(X_l^{(0)}(X_k - X_k^{(0)}))\|_\infty\,. \quad (49)\end{aligned}$$

Clearly, by $C(A)$ and the ϕ-mixing property,

$$\|\mathbf{E}(X_k(X_l - X_l^{(0)})|\mathcal{F}_0) - \mathbf{E}(X_k(X_l - X_l^{(0)}))\|_\infty \leq 4\|X_k\|_\infty\sum_{k=1}^{\infty}\Delta_{l-k}\phi_\varepsilon(k)\,,$$

which tends to zero as l tends to infinity. In the same way

$$\lim_{k\to 0}\|\mathbf{E}(X_l^{(0)}(X_k - X_k^{(0)})|\mathcal{F}_0) - \mathbf{E}(X_l^{(0)}(X_k - X_k^{(0)}))\|_\infty = 0\,.$$

Let $H_k = H - \mathbf{E}(Z_k^{(0)})$. Let $(\eta_i)_{i\in\mathbf{Z}}$ be distributed as $(\varepsilon_i)_{i\in\mathbf{Z}}$ and independent of $((\varepsilon_i)_{i\in\mathbf{Z}}, (\varepsilon_i')_{i>0})$, and let $(\eta_i^{(0)})_{i\in\mathbf{Z}}$ be the sequence defined by $\eta_i^{(0)} = \eta_i$ if $i \leq 0$ and $\eta_i^{(0)} = \varepsilon_i'$ if $i > 0$. With this notations, we have

$$\begin{aligned}\mathbf{E}(X_k^{(0)}X_l^{(0)}|\mathcal{F}_0) - \mathbf{E}(X_k^{(0)}X_l^{(0)}) &= \mathbf{E}_\varepsilon(H_k((\varepsilon_{k-i}^{(0)})_{i\in Z})(H_l((\varepsilon_{l-i}^{(0)})_{i\in Z}) - H_l((\eta_{l-i}^{(0)})_{i\in Z}))) \\ &+ \mathbf{E}_\varepsilon(H_l((\eta_{l-i}^{(0)})_{i\in Z})(H_k((\varepsilon_{k-i}^{(0)})_{i\in Z}) - H_k((\eta_{k-i}^{(0)})_{i\in Z}))) \quad (50)\end{aligned}$$

Consequently, applying $C(A)$ once more, we have that

$$\|\mathbf{E}(X_k^{(0)}X_l^{(0)}|\mathcal{F}_0) - \mathbf{E}(X_k^{(0)}X_l^{(0)})\|_\infty \leq \|X_k^{(0)}\|_\infty\sum_{i\geq l}\Delta_i + \|X_l^{(0)}\|_\infty\sum_{i\geq k}\Delta_i\,,$$

which tends to zero as k and l tends to infinity. This completes the proof. $\diamond$

4.7 Proof of Proposition 12

We shall apply Corollary 2. We use the same notations as for the proof of Proposition 11. With these notations, we have

$$\mathbf{E}(X_k|\mathcal{F}_0) = \mathbf{E}(X_k^{(0)}|\mathcal{F}_0) + \mathbf{E}(X_k - X_k^{(0)}|\mathcal{F}_0)\,.$$

Now, applying $C'(A)$,

$$|\mathbf{E}(X_k - X_k^{(0)}|\mathcal{F}_0)| \leq \sum_{i=1}^{k} R_{k-i}\mathbf{E}(\mathbf{1}_{\varepsilon_i \neq \varepsilon_i'}|\mathcal{F}_0) + \sum_{i=1}^{k} R_{k-i}\mathbf{P}(\varepsilon_i \neq \varepsilon_i')\,,$$

and by the ϕ-mixing property,

$$\|\mathbf{E}(X_k - X_k^{(0)}|\mathcal{F}_0)\|_\infty \leq 2\sum_{i=1}^{k} R_{k-i}\phi_\varepsilon(i)\,. \tag{51}$$

Now, by $C'(A)$ again,

$$\|\mathbf{E}(X_k^{(0)}|\mathcal{F}_0)\|_\infty = \|\mathbf{E}_\varepsilon(H((\varepsilon_{k-i}^{(0)})_{i\in\mathbf{Z}}) - H((\eta_{k-i}^{(0)})_{i\in\mathbf{Z}}))\|_\infty \leq R_k\,.$$

Consequently, since $\phi_\varepsilon(0) > 0$, the condition (5) is implied by (14). It remains to prove (6). We start from the decomposition (49). By (51),

$$\begin{aligned} \|\mathbf{E}(X_k(X_l - X_l^{(0)})|\mathcal{F}_0) - \mathbf{E}(X_k(X_l - X_l^{(0)}))\|_\infty &\leq 4\|X_k\|_\infty \sum_{i=1}^{l} R_{l-i}\phi_\varepsilon(i)\,, \\ \text{and}\quad \|\mathbf{E}(X_l^{(0)}(X_k - X_k^{(0)})|\mathcal{F}_0) - \mathbf{E}(X_l^{(0)}(X_k - X_k^{(0)}))\|_\infty &\leq 4\|X_l^{(0)}\|_\infty \sum_{i=1}^{k} R_{k-i}\phi_\varepsilon(i)\,. \end{aligned}$$

Hence, in view of (14), these two terms converges to zero as k and l tend to infinity. From (50) and condition $C'(A)$, we have that

$$\|\mathbf{E}(X_k^{(0)}X_l^{(0)}|\mathcal{F}_0) - \mathbf{E}(X_k^{(0)}X_l^{(0)})\|_\infty \leq \|X_k^{(0)}\|_\infty R_l + \|X_l^{(0)}\|_\infty R_k\,,$$

which again converges to zero as k and l tend to infinity. This completes the proof. $\diamond$

4.8 Proof of Proposition 14

It suffices to prove that for any f in $\mathcal{L}$, the sequence $X_i = f(Y_i) - \mu(f)$ satisfies the conditions (5) and (6) of Corollary 2.

Note first that (6) holds because of (19) and because any continuous function from $[0,1]$ to $\mathbf{R}$ can be uniformly approximated by Lipschitz functions.

From Lemma 15, we have that

$$\|K^n(f) - \mu(f)\|_{\infty,\mu} \leq c(C\rho^n)\,,$$

for some concave non decreasing function c. Consequently (5) holds as soon as $\sum_{k>0} k^{-1/2}c(C\rho^k)$ is finite, which in turn is equivalent to (21).

4.9 Proof of Lemma 15

Let $(Y_i)_{i\geq 1}$ be the Markov chain with transition Kernel K and and invariant measure μ. From Lemma 1 in Dedecker and Merlevède (2006), we know that there exists Y_k^* distributed as Y_k and independent of Y_0 such that

$$\sup_{g\in\Lambda_1} \|K^k(g) - \mu(g)\|_{\infty,\mu} = \|\mathbf{E}(|Y_k - Y_k^*||Y_0)\|_\infty .$$

For any f such that $|f(x) - f(y)| \leq c(|x - y|)$, we have

$$\begin{aligned} \|K^k(f) - \mu(f)\|_{\infty,\mu} &= \|\mathbf{E}(f(Y_k)|Y_0) - \mathbf{E}(f(Y_k^*)|Y_0)\|_\infty \\ &\leq \|\mathbf{E}(c(|Y_k - Y_k^*|)|Y_0)\|_\infty . \end{aligned}$$

Since c is concave and non decreasing, we get that

$$\|K^k(f) - \mu(f)\|_{\infty,\mu} \leq \|c(\mathbf{E}(|Y_k - Y_k^*||Y_0))\|_\infty \leq c(\|\mathbf{E}(|Y_k - Y_k^*||Y_0)\|_\infty) ,$$

and the proof is complete. $\diamond$

4.10 Proof of Corollary 18

Let $(Y_i)_{i\geq 1}$ be the Markov chain with transition Kernel K and invariant measure μ. Using the equation (26) is easy to see that $(Y_0, \ldots, Y_n)$ is distributed as $(T^{n+1}, \ldots, T)$. Consequently, for f in $\mathcal{L}$, Corollary 18 follows from Proposition 14 and Condition (23).

Assume now that f is BV. We shall prove that the sequence $X_i = f(Y_i) - \mu(f)$ satisfies the conditions (5) and (6) of Corollary 2. Since K is BV-contracting we have that

$$\|\mathbf{E}(X_k|Y_0)\|_\infty = \|K^k(f) - \mu(f)\|_{\infty,\mu} \leq \|dK^k(f)\| \leq C\rho^k\|df\| ,$$

so that (5) is satisfied. On the other hand, applying Lemma 1 in Dedecker and Prieur (2007), we have that, for any $l > k \geq 0$,

$$\|\mathbf{E}(X_kX_l|Y_0) - \mathbf{E}(X_kX_l)\|_\infty \leq C(1 + C)\rho^k\|df\|^2 , \tag{52}$$

so that (6) holds. This completes the proof of Corollary 18 when f is BV. $\diamond$

4.11 Proof of Proposition 20

To prove Proposition 20, it suffices to prove that the sequence $X_i = f(\xi_i) - m(f)$ satisfies the conditions (5) and (6) of Corollary 2. Let $\|\cdot\|_{\infty,m}$ be the essential supremum norm with respect to m.

Note that $\|\mathbf{E}(X_n|\xi_0)\|_\infty = \|K^n(f) - m(f)\|_{\infty,m}$, and that

$$K^n(f)(x) - m(f) = \sum_{k\in\mathbf{Z}^*} \cos^n(2\pi ka)\hat{f}(k)\exp(2i\pi kx) .$$

By assumption, there exists $C > 0$ such that $\sup_{k\neq 0} |k|^{1+\varepsilon}|\hat{f}(k)| \leq C$. Hence

$$\sum_{n>0} \frac{\|K^n(f) - m(f)\|_{\infty,m}}{\sqrt{n}} \leq C\sum_{k\in\mathbf{Z}^*} |k|^{-1-\varepsilon}\sum_{n>0} \frac{|\cos(2\pi ka)|^n}{\sqrt{n}} . \tag{53}$$

Here, note that there exists a positive constant K such that, for any $0 < a < 1$, we have $\sum_{n>0} n^{-1/2}a^n \leq Ka(1-a)^{-1/2}$ (to see this, it suffices to compare the sum with the integral of the function $h(x) = x^{-1/2}a^x$). Consequently, we infer from (53) that

$$\begin{aligned}\sum_{n>0} \frac{\|K^n(f) - m(f)\|_{\infty,m}}{\sqrt{n}} &\leq CK \sum_{k\in\mathbf{Z}^*} \frac{1}{|k|^{1+\varepsilon}\sqrt{1-|\cos(2\pi ka)|}} \\ &\leq CK \sum_{k\in\mathbf{Z}^*} \frac{1}{|k|^{1+\varepsilon} d(2ka, \mathbf{Z})}\,, \end{aligned} \tag{54}$$

the last inequality being true because $(1-|\cos(\pi u)|) \geq \pi(d(u,\mathbf{Z}))^2$. Since a is badly approximable by rationals, then so is $2a$. Hence, arguing as in the proof of Lemma 5.1 in Dedecker and Rio (2006), we infer that for any positive η there exists a constant D such that

$$\sum_{k=2^N}^{2^{N+1}-1} \frac{1}{d(2ka,\mathbf{Z})} \leq D2^{(N+2)(1+\eta)}N\,.$$

Applying this result with $\eta = \varepsilon/2$, we infer from (54) that

$$\sum_{n>0} \frac{\|K^n(f) - m(f)\|_{\infty,m}}{\sqrt{n}} \leq 2CKD \sum_{N\geq 0} 2^{(N+2)(1+\varepsilon/2)}N \max_{2^N\leq k\leq 2^{N+1}} k^{-1-\varepsilon} < \infty\,,$$

so that the condition (5) of Corollary 2 is satisfied. The condition (6) of Corollary 2 follows from the inequality (5.18) in Dedecker and Rio (2006). $\diamond$

4.12 Appendix

This section collects some technical lemmas.

The proof of the following lemma is left to the reader since it uses the same arguments as in the proof of Proposition 2.5 in Peligrad and Utev (2005) by replacing the $\mathbf{L}_2$ norm by the $\mathbf{L}_\infty$ norm.

Lemma 29 *Under condition (2),*

$$\frac{\|\mathbf{E}(S_m|\mathcal{F}_0)\|_\infty}{\sqrt{m}} \to 0 \quad \text{and} \quad \frac{1}{\sqrt{m}}\sum_{j=1}^{\infty} \frac{\|\mathbf{E}(S_{mj}|\mathcal{F}_0)\|_\infty}{j^{3/2}} \to 0 \quad \text{as } m\to\infty.$$

The following lemma gives a simple fact about convergence.

Lemma 30 *Let $A_j(x,n,m)$, $j=1,\ldots,J$, $x>0$, be real valued functions such that for each j,n,m the function $A_j(x,n,m)$ is non-increasing in $x>0$ and assume that, for any $x>0$,*

$$\limsup_{m\to\infty}\limsup_{n\to\infty} A_j(x,n,m) = -\infty\,.$$

Then for any $u_n\to\infty$, there exists $m_n\to\infty$ such that $m_n\leq u_n$ and, for any $x>0$ and $j=1,\ldots,J$,

$$\limsup_{n\to\infty} A_j(x,n,m_n) = -\infty$$

Proof. First, we observe that by considering the function

$$A(x,n,m) = \max_{1\le j\le J} A_j(x,n,m) \ ,$$

the lemma reduces to the case $J = 1$.

Construct a strictly increasing positive integer sequences ψ_k and n_k such that for all $n \ge n_k$,

$$A(1/k, n, \psi_k) \le -k \ .$$

Let $g(n) = k$ for $n_k < n \le n_{k+1}$ starting with $k = 1$ and $g(n) = 1$ for $n \le n_1$. Then, $g(n)$ is non-decreasing, $g(n) \to \infty$ and for all $n > n_1$ such that $n_k < n \le n_{k+1}$ (and so $g(n) = k$).

$$n_{g(n)} = n_k < n$$

Now, let $G(n)$ be a positive integer sequence such that $G(n) \le g(n)$ and $G(n) \to \infty$. Then,

$$n_{G(n)} \le n_{g(n)} = n_k < n$$

Hence, there exists $G(n)$ such that

$$\psi_{G(n)} \le u_n \ , \ n_{G(n)} \le n \quad \text{and} \quad G(n) \to \infty \ .$$

Finally, let $m_n = \psi_{G(n)}$. Then, obviously

$$m_n \le u_n \quad \text{and} \quad m_n \to \infty \ .$$

On the other hand, for any $x > 0$ and n such that $x \ge 1/G(n)$, since $A(x,n,m)$ is non-increasing in x, we have

$$\begin{aligned} A(x,n,m_n) &\le A(1/G(n), n, m_n) = A(1/G(n), n, \psi_{G(n)}) \\ &\le -G(n) \to -\infty \end{aligned}$$

which proves the lemma. $\diamond$

References

[1] Arcones, M.A. (2003). The large deviation principle for stochastic processes I. *Theory of Probability and its Applications* **47**, 567-583.

[2] Billingsley, P. (1968). Convergence of Probability Measures. Wiley, New-York.

[3] Bradley, R.C. (2002). *Introduction to Strong Mixing Conditions, Volume 1.* Technical Report, Department of Mathematics, Indiana University, Bloomington. Custom Publishing of I.U., Bloomington, March 2002.

[4] Broise, A. (1996). Transformations dilatantes de l'intervalle et théorèmes limites. Études spectrales d'opérateurs de transfert et applications. Astérisque. 238 1-109.

[5] de Acosta, A. and Chen X. (1998). Moderate deviations for empirical measure of Markov chains: upper bound. *J. Theor. Probab.* **4** 75–110.

[6] Dedecker, J. and Merlevède F. (2006). Inequalities for partial sums of Hilbert-valued dependent sequences and applications. *Math. Methods Statist.* **15**, 176–206.

[7] Dedecker, J. and Prieur, C. (2007). An empirical central limit theorem for dependent sequences. *Stoch. Processes. Appl.* **117**, 121-142.

[8] Dedecker, J. and Rio, E. (2006). On mean central limit theorems for stationary sequences. To appear in *Ann. Inst. H. Poincaré Probab. Statist.* http://www.lsta.upmc.fr/Dedecker/publi.html

[9] Delyon, B., Juditsky, A. and Liptser, R. (2006). Moderate deviation principle for ergodic Markov chain. Lipschitz summands. In: *From stochastic calculus to mathematical finance,* 189–209, Springer, Berlin.

[10] Dembo, A. (1996). Moderate deviations for martingales with bounded jumps. *Elect. Comm. Probab.***1**, 11-17.

[11] Dembo, A. and Zeitouni, O. (1997). Moderate deviations of iterates of expanding maps. *Statistics and control of stochastic processes.* 1-11, World Sci. Publi., River Edge, NJ.

[12] Dembo, A. and Zeitouni, O. (1998). Large Deviations Techniques and Applications, 2nd edition, Springer New York.

[13] Derriennic, Y. and Lin, M. (2001). The central limit theorem for Markov chains with normal transition operators, started at a point. *Probab. Theory Relat. Fields.* **119**, 508-528.

[14] Deuschel, J.D. and Stroock, D.W. (1989). Large deviations, vol. 137 of *Pure and Applied Mathematics.* Academic Press Inc., Boston, M.A..

[15] Djellout, H. (2002). Moderate Deviations for Martingale Differences and applications to $\phi-$mixing sequences. *Stoch. Stoch. Rep.* **73**, No.1-2, 37-63.

[16] Djellout, H., Guillin, A. and Wu, L. (2006). Moderate Deviations of empirical periodogram and non-linear functionals of moving average processes. *Ann. Inst. H. Poincaré Probab. Statist.* **42**, No. 4, 393-416.

[17] Gao, F-Q. (1996). Moderate deviations for martingales and mixing random processes. *Stochastic Process. Appl.* **61**, 263–275.

[18] Gordin, M. I. (1969). The central limit theorem for stationary processes. *Dokl. Akad. Nauk SSSR* **188**, 739-741.

[19] Peligrad, M. and Utev, S. (2005). A new maximal inequality and invariance principle for stationary sequences. *The Annals of Probability* **33**, 798–815.

[20] Peligrad, M. and Utev, S. (2006). Central limit theorem for stationary linear processes. *The Annals of Probability* **34**, 1608–1622.

[21] Peligrad, M., Utev S and Wu W. B. (2007). A maximal L_p–inequality for stationary sequences and its applications. *Proc. Amer. Math. Soc.* **135**, 541–550.

[22] Puhalskii, A. (1994) Large deviations of semimartingales via convergence of the predictable characteristics, *Stoch. Stoch. Rep.* **49**, 2785.

[23] Schmidt, W.M. (1980). Diophantine approximation. Lectures Notes in Mathematics. 785.

[24] Wu, L. (1999). Exponential convergence in probability for empirical means of Brownian motion and of random walks. *J. Theor. Probab.* **12**, 661-673.