

# MODERATE DEVIATIONS FOR THE DURBIN-WATSON STATISTIC RELATED TO THE FIRST-ORDER AUTOREGRESSIVE PROCESS

S.VALÈRE BITSEKI PENDA, HACÈNE DJELLOUT, AND FRÉDÉRIC PROÏA

ABSTRACT. The purpose of this paper is to investigate moderate deviations for the Durbin-Watson statistic associated with the stable first-order autoregressive process where the driven noise is also given by a first-order autoregressive process. We first establish a moderate deviation principle for both the least squares estimator of the unknown parameter of the autoregressive process as well as for the serial correlation estimator associated with the driven noise. It enables us to provide a moderate deviation principle for the Durbin-Watson statistic in the case where the driven noise is normally distributed and in the more general case where the driven noise satisfies a less restrictive Chen-Ledoux type condition.

*AMS 2000 subject classifications: 60F10, 60G42, 62M10, 62G05.*

## 1. INTRODUCTION

This paper is focused on the stable first-order autoregressive process where the driven noise is also given by a first-order autoregressive process. The purpose is to investigate moderate deviations for both the least squares estimator of the unknown parameter of the autoregressive process as well as for the serial correlation estimator associated with the driven noise. Our goal is to establish moderate deviations for the Durbin-Watson statistic [11], [12], [13], in a lagged dependent random variables framework. First of all, we shall assume that the driven noise is normally distributed. Then, we will extend our investigation to the more general framework where the driven noise satisfies a less restrictive Chen-Ledoux type condition [5], [17]. We are inspired by the recent paper of Bercu and Proïa [3], where the almost sure convergence and the central limit theorem are established for both the least squares estimators and the Durbin-Watson statistic. Our results are proved *via* an extensive use of the results of Dembo [6], Dembo and Zeitouni [7] and Worms [24], [25] on the one hand, and of the paper of Puhalskii [21] and Djellout [8] on the other hand, about moderate deviations for martingales. In order to introduce the Durbin-Watson statistic, the first-order autoregressive process of interest is as follows, for all  $n \geq 1$ ,

$$\begin{cases} X_n = \theta X_{n-1} + \varepsilon_n \\ \varepsilon_n = \rho \varepsilon_{n-1} + V_n \end{cases} \quad (1.1)$$

where we shall assume that the unknown parameters  $|\theta| < 1$  and  $|\rho| < 1$  to ensure the stability of the model. In all the sequel, we also assume that  $(V_n)$  is a sequence of independent and identically distributed random variables with zero mean and positive variance  $\sigma^2$ . The square-integrable initial values  $X_0$  and  $\varepsilon_0$  may be arbitrarily chosen. We have decided to

---

*Key words and phrases.* Durbin-Watson statistic, Moderate deviation principle, First-order autoregressive process, Serial correlation.

estimate  $\theta$  by the least squares estimator

$$\hat{\theta}_n = \frac{\sum_{k=1}^n X_k X_{k-1}}{\sum_{k=1}^n X_{k-1}^2}. \quad (1.2)$$

Then, we also define a set of least squares residuals given, for all  $1 \leq k \leq n$ , by

$$\hat{\varepsilon}_k = X_k - \hat{\theta}_n X_{k-1}, \quad (1.3)$$

which leads to the estimator of  $\rho$ ,

$$\hat{\rho}_n = \frac{\sum_{k=1}^n \hat{\varepsilon}_k \hat{\varepsilon}_{k-1}}{\sum_{k=1}^n \hat{\varepsilon}_{k-1}^2}. \quad (1.4)$$

Finally, the Durbin-Watson statistic is defined, for  $n \geq 1$ , as

$$\hat{D}_n = \frac{\sum_{k=1}^n (\hat{\varepsilon}_k - \hat{\varepsilon}_{k-1})^2}{\sum_{k=0}^n \hat{\varepsilon}_k^2}. \quad (1.5)$$

This well-known statistic was introduced by the pioneer work of Durbin and Watson [11], [12], [13], in the middle of last century, to test the presence of a significative first order serial correlation in the residuals of a regression analysis. A wide range of literature is available on the asymptotic behavior of the Durbin-Watson statistic, frequently used in Econometry. While it appeared to work pretty well in the classical independent framework, Malinvaud [18] and Nerlove and Wallis [19] observed that, for linear regression models containing lagged dependent random variables, the Durbin-Watson statistic may be asymptotically biased, potentially leading to inadequate conclusions. Durbin [10] proposed alternative tests to prevent this misuse, such as the *h-test* and the *t-test*, then substantial contributions were brought by Inder [15], King and Wu [16] and more recently Stocker [22]. Lately, a set of results have been established by Bercu and Proïa in [3] for the first-order autoregressive process, and by Proïa [20] for the autoregressive process of any order, in particular a test procedure as powerful as the *h-test* and more accurate than the usual *portmanteau* tests, and they will be summarized thereafter as a basis for this paper in the one-dimensional case. This work can be seen as an extension of [3] in the sense that more powerful convergences are reached and that a better precision than the central limit theorem is provided for the same random sequences. Hence, the establishment of moderate deviations is the natural continuation following the proof of central limit theorems and laws of iterated logarithm. We are now interested in the asymptotic estimation of

$$\mathbb{P} \left( \frac{\sqrt{n}}{b_n} (\Theta_n - \Theta) \in A \right)$$

where  $\Theta_n$  denotes the estimator of the unknown parameter of interest  $\Theta$ ,  $A$  is a given domain of deviations and  $(b_n)$  denotes the scale of deviations. When  $b_n = 1$ , this is exactly the estimation of the *central limit theorem* (CLT). When  $b_n = \sqrt{n}$ , it becomes a *large deviation principle* (LDP). And when  $1 \ll b_n \ll \sqrt{n}$ , this is the so-called *moderate deviation principle* (MDP). Usually, an MDP has a simpler rate function inherited from the approximated gaussian process which does not necessarily depend on the parameters under investigation and holds for a larger class of dependent random variables than the LDP. Furthermore, an MDP can be seen as a refinement of the CLT in the sense that the MDP tells us that the gaussian estimation still holds up to the scale of large deviations. For the sake of clarity, all useful definitions will be given later.

The paper is organized as follows. First of all, we recall the results recently established by Bercu and Proïa [3]. In Section 2, we propose moderate deviation principles for the estimators of  $\theta$  and  $\rho$  and for the Durbin-Watson statistic, given by (1.2), (1.4) and (1.5), under the normality assumption on the driven noise. Section 3 deals with the generalization of the latter results under a less restrictive Chen-Ledoux type condition on  $(V_n)$ . Finally, all technical proofs are postponed to Section 4.

**Lemma 1.1.** *Assume that  $(V_n)$  is independent and identically distributed with positive finite variance. Then, we have the almost sure convergence of the autoregressive estimator,*

$$\lim_{n \rightarrow \infty} \widehat{\theta}_n = \theta^* \quad \text{a.s.}$$

where the limiting value

$$\theta^* = \frac{\theta + \rho}{1 + \theta\rho}. \quad (1.6)$$

In addition, as soon as  $\mathbb{E}[V_1^4] < \infty$ , we also have the asymptotic normality,

$$\sqrt{n} \left( \widehat{\theta}_n - \theta^* \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\theta^2)$$

where the asymptotic variance

$$\sigma_\theta^2 = \frac{(1 - \theta^2)(1 - \theta\rho)(1 - \rho^2)}{(1 + \theta\rho)^3}. \quad (1.7)$$

**Lemma 1.2.** *Assume that  $(V_n)$  is independent and identically distributed with positive finite variance. Then, we have the almost sure convergence of the serial correlation estimator,*

$$\lim_{n \rightarrow \infty} \widehat{\rho}_n = \rho^* \quad \text{a.s.}$$

where the limiting value

$$\rho^* = \theta\rho\theta^*. \quad (1.8)$$

Moreover, as soon as  $\mathbb{E}[V_1^4] < \infty$ , we have the asymptotic normality,

$$\sqrt{n} \left( \widehat{\rho}_n - \rho^* \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\rho^2)$$

with the asymptotic variance

$$\sigma_\rho^2 = \frac{(1 - \theta\rho)}{(1 + \theta\rho)^3} \left( (\theta + \rho)^2(1 + \theta\rho)^2 + (\theta\rho)^2(1 - \theta^2)(1 - \rho^2) \right). \quad (1.9)$$

In addition, we have the joint asymptotic normality,

$$\sqrt{n} \begin{pmatrix} \widehat{\theta}_n - \theta^* \\ \widehat{\rho}_n - \rho^* \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma)$$

where the covariance matrix

$$\Gamma = \begin{pmatrix} \sigma_\theta^2 & \theta\rho\sigma_\theta^2 \\ \theta\rho\sigma_\theta^2 & \sigma_\rho^2 \end{pmatrix}. \quad (1.10)$$

**Lemma 1.3.** *Assume that  $(V_n)$  is independent and identically distributed with positive finite variance. Then, we have the almost sure convergence of the Durbin-Watson statistic,*

$$\lim_{n \rightarrow \infty} \widehat{D}_n = D^* \quad \text{a.s.}$$

where the limiting value

$$D^* = 2(1 - \rho^*). \quad (1.11)$$

In addition, as soon as  $\mathbb{E}[V_1^4] < \infty$ , we have the asymptotic normality,

$$\sqrt{n} \left( \widehat{D}_n - D^* \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_D^2)$$

where the asymptotic variance

$$\sigma_D^2 = 4\sigma_\rho^2. \quad (1.12)$$

*Proof.* The proofs of Lemma 1.1, Lemma 1.2 and Lemma 1.3 may be found in [3].  $\square$

Our objective is now to establish a set of moderate deviation principles on these estimates in order to get a better asymptotic accuracy than the central limit theorem.

In the whole paper, for any matrix  $M$ ,  $M'$  and  $\|M\|$  stand for the transpose and the euclidean norm of  $M$ , respectively. In addition, for a sequence of random variables  $(Z_n)_n$  on  $\mathbb{R}^{d \times p}$ , we say that  $(Z_n)_n$  converges  $(a_n)$ -superexponentially fast in probability to some random variable  $Z$  with  $a_n \rightarrow \infty$  if, for all  $\delta > 0$ ,

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} \log \mathbb{P} \left( \|Z_n - Z\| > \delta \right) = -\infty.$$

This *exponential convergence* with speed  $a_n$  will be shortened as

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

The *exponential equivalence* with speed  $a_n$  between two sequences of random variables  $(Y_n)_n$  and  $(Z_n)_n$ , whose precise definition is given in Definition 4.2.10 of [7], will be shortened as

$$Y_n \overset{\text{superexp}}{\underset{a_n}{\sim}} Z_n.$$

We start by recalling some useful definitions.

**Definition 1.1** (Large Deviation Principle). *We say that a sequence of random variables  $(M_n)_n$  with topological state space  $(S, \mathcal{S})$  satisfies an LDP with speed  $a_n$  and rate function  $I : S \rightarrow \mathbb{R}^+$  if  $a_n \rightarrow \infty$  and, for each  $A \in \mathcal{S}$ ,*

$$-\inf_{x \in A^\circ} I(x) \leq \liminf_{n \rightarrow \infty} \frac{1}{a_n} \log \mathbb{P}(M_n \in A) \leq \limsup_{n \rightarrow \infty} \frac{1}{a_n} \log \mathbb{P}(M_n \in A) \leq -\inf_{x \in \bar{A}} I(x)$$

where  $A^\circ$  and  $\bar{A}$  denote the interior and the closure of  $A$ , respectively. The rate function  $I$  is lower semicontinuous, i.e. all the sub-level sets  $\{x \in S \mid I(x) \leq c\}$  are closed, for  $c \geq 0$ .

Let  $(b_n)$  be a sequence of increasing positive numbers satisfying  $1 = o(b_n^2)$  and  $b_n^2 = o(n)$ ,

$$b_n \longrightarrow \infty, \quad \frac{b_n}{\sqrt{n}} \longrightarrow 0. \quad (1.13)$$

**Definition 1.2** (Moderate Deviation Principle). *We say that a sequence of random variables  $(M_n)_n$  with topological state space  $(S, \mathcal{S})$  satisfies an MDP with speed  $b_n^2$  such that (1.13) holds, and rate function  $I : S \rightarrow \mathbb{R}^+$  if the sequence  $(\sqrt{n}M_n/b_n)_n$  satisfies an LDP with speed  $b_n^2$  and rate function  $I$ .*

Formally, our main results about the MDP for a sequence of random variables  $(M_n)_n$  will be stated as the LDP for the sequence  $(\sqrt{n}M_n/b_n)_n$ .

## 2. ON MODERATE DEVIATIONS UNDER THE GAUSSIAN CONDITION

In this first part, we focus our attention on moderate deviations for the Durbin-Watson statistic in the easy case where the driven noise  $(V_n)$  is normally distributed. This restrictive assumption allows us to reduce the set of hypotheses to the existence of  $t > 0$  such that

**G1**

$$\mathbb{E} \left[ \exp(t\varepsilon_0^2) \right] < \infty,$$

**G2**

$$\mathbb{E} \left[ \exp(tX_0^2) \right] < \infty.$$

**Theorem 2.1.** *Assume that there exists  $t > 0$  such that **G1** and **G2** are satisfied and that  $(V_n)$  follows the  $\mathcal{N}(0, \sigma^2)$  distribution. Then, the sequence*

$$\left( \frac{\sqrt{n}}{b_n} \left( \hat{\theta}_n - \theta^* \right) \right)_{n \geq 1}$$

*satisfies an LDP on  $\mathbb{R}$  with speed  $b_n^2$  and rate function*

$$I_\theta(x) = \frac{x^2}{2\sigma_\theta^2} \tag{2.1}$$

*where  $\sigma_\theta^2$  is given by (1.7).*

**Theorem 2.2.** *Assume that there exists  $t > 0$  such that **G1** and **G2** are satisfied and that  $(V_n)$  follows the  $\mathcal{N}(0, \sigma^2)$  distribution. Then, as soon as  $\theta \neq -\rho$ , the sequence*

$$\left( \frac{\sqrt{n}}{b_n} \begin{pmatrix} \hat{\theta}_n - \theta^* \\ \hat{\rho}_n - \rho^* \end{pmatrix} \right)_{n \geq 1}$$

*satisfies an LDP on  $\mathbb{R}^2$  with speed  $b_n^2$  and rate function*

$$K(x) = \frac{1}{2}x'\Gamma^{-1}x \tag{2.2}$$

*where  $\Gamma$  is given by (1.10). In particular, the sequence*

$$\left( \frac{\sqrt{n}}{b_n} \left( \hat{\rho}_n - \rho^* \right) \right)_{n \geq 1}$$

*satisfies an LDP on  $\mathbb{R}$  with speed  $b_n^2$  and rate function*

$$I_\rho(x) = \frac{x^2}{2\sigma_\rho^2} \tag{2.3}$$

*where  $\sigma_\rho^2$  is given by (1.9).*

**Remark 2.1.** *The covariance matrix  $\Gamma$  is invertible if and only if  $\theta \neq -\rho$  since one can see by a straightforward calculation that its determinant is given by*

$$\det(\Gamma) = \frac{\sigma_\theta^2(\theta + \rho)^2(1 - \theta\rho)}{(1 + \rho^2)}.$$

Moreover, in the particular case where  $\theta = -\rho$ , the sequences

$$\left(\frac{\sqrt{n}}{b_n}(\hat{\theta}_n - \theta^*)\right)_{n \geq 1} \quad \text{and} \quad \left(\frac{\sqrt{n}}{b_n}(\hat{\rho}_n - \rho^*)\right)_{n \geq 1}$$

satisfy LDP on  $\mathbb{R}$  with speed  $b_n^2$  and rate functions respectively given by

$$I_\theta(x) = \frac{x^2(1 - \theta^2)}{2(1 + \theta^2)} \quad \text{and} \quad I_\rho(x) = \frac{x^2(1 - \theta^2)}{2\theta^4(1 + \theta^2)}.$$

**Theorem 2.3.** *Assume that there exists  $t > 0$  such that **G1** and **G2** are satisfied and that  $(V_n)$  follows the  $\mathcal{N}(0, \sigma^2)$  distribution. Then, the sequence*

$$\left(\frac{\sqrt{n}}{b_n}(\hat{D}_n - D^*)\right)_{n \geq 1}$$

satisfies an LDP on  $\mathbb{R}$  with speed  $b_n^2$  and rate function

$$I_D(x) = \frac{x^2}{2\sigma_D^2} \tag{2.4}$$

where  $\sigma_D^2$  is given by (1.12).

*Proof.* Theorem 2.1, Theorem 2.2 and Theorem 2.3 are proved in Section 4. □

### 3. ON MODERATE DEVIATIONS UNDER THE CHEN-LEDoux TYPE CONDITION

Via an extensive use of Puhalskii's result, we will now focus our attention on the more general framework where the driven noise  $(V_n)$  is assumed to satisfy the Chen-Ledoux type condition. Accordingly, one shall introduce the following hypothesis, for any  $a > 0$ .

**CL1(a)** Chen-Ledoux.

$$\limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log n\mathbb{P}\left(|V_1|^a > b_n\sqrt{n}\right) = -\infty.$$

**CL2(a)**

$$\frac{|\varepsilon_0|^a}{b_n\sqrt{n}} \xrightarrow[b_n^2]{\text{superexp}} 0.$$

**CL3(a)**

$$\frac{|X_0|^a}{b_n\sqrt{n}} \xrightarrow[b_n^2]{\text{superexp}} 0.$$

**Remark 3.1.** *If the random variable  $V_1$  satisfies **CL1(2)**, then*

$$\limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log n\mathbb{P}\left(|V_1^2 - \mathbb{E}[V_1^2]| > b_n\sqrt{n}\right) = -\infty, \tag{3.1}$$

which implies in particular that  $\text{Var}(V_1^2) < \infty$ . Moreover, if the random variable  $V_1^2$  has exponential moments, i.e. if there exists  $t > 0$  such that

$$\mathbb{E}\left[\exp(tV_1^2)\right] < \infty,$$

then **CL1(2)** is satisfied for every increasing sequence  $(b_n)$ . From [1], [2], [14], condition (3.1) is equivalent to say that the sequence

$$\left(\frac{1}{b_n\sqrt{n}} \sum_{k=1}^n (V_k^2 - \mathbb{E}[V_k^2])\right)_{n \geq 1}$$

satisfies an LDP on  $\mathbb{R}$  with speed  $b_n^2$  and rate function

$$I(x) = \frac{x^2}{2\text{Var}(V_1^2)}.$$

**Remark 3.2.** If we choose  $b_n = n^\alpha$  with  $0 < \alpha < 1/2$ , **CL1(2)** is immediately satisfied if there exists  $t > 0$  and  $0 < \beta < 1$  such that

$$\mathbb{E}\left[\exp(tV_1^{2\beta})\right] < \infty,$$

which is clearly a weaker assumption than the existence of  $t > 0$  such that

$$\mathbb{E}\left[\exp(tV_1^2)\right] < \infty,$$

imposed in the previous section.

**Remark 3.3.** If **CL1(a)** is satisfied, then **CL1(b)** is also satisfied for all  $0 < b < a$ .

**Remark 3.4.** In the technical proofs that will follow, rather than **CL1(4)**, the weakest assumption really needed is summarized by the existence of a large constant  $C$  such that

$$\limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{1}{n} \sum_{k=1}^n V_k^4 > C\right) = -\infty.$$

**Theorem 3.1.** Assume that **CL1(4)**, **CL2(4)** and **CL3(4)** are satisfied. Then, the sequence

$$\left(\frac{\sqrt{n}}{b_n} (\hat{\theta}_n - \theta^*)\right)_{n \geq 1}$$

satisfies the LDP on  $\mathbb{R}$  stated in Theorem 2.1.

**Theorem 3.2.** Assume that **CL1(4)**, **CL2(4)** and **CL3(4)** are satisfied. Then, as soon as  $\theta \neq -\rho$ , the sequence

$$\left(\frac{\sqrt{n}}{b_n} \begin{pmatrix} \hat{\theta}_n - \theta^* \\ \hat{\rho}_n - \rho^* \end{pmatrix}\right)_{n \geq 1}$$

satisfies the LDP on  $\mathbb{R}^2$  stated in Theorem 2.2. In particular, the sequence

$$\left(\frac{\sqrt{n}}{b_n} (\hat{\rho}_n - \rho^*)\right)_{n \geq 1}$$

satisfies the LDP on  $\mathbb{R}$  also stated in Theorem 2.2.

**Remark 3.5.** *We have already seen in Remark 2.1 that the covariance matrix  $\Gamma$  is invertible if and only if  $\theta \neq -\rho$ . In the particular case where  $\theta = -\rho$ , the sequences*

$$\left(\frac{\sqrt{n}}{b_n}(\widehat{\theta}_n - \theta^*)\right)_{n \geq 1} \quad \text{and} \quad \left(\frac{\sqrt{n}}{b_n}(\widehat{\rho}_n - \rho^*)\right)_{n \geq 1}$$

*satisfy the LDP on  $\mathbb{R}$  stated in Remark 2.1.*

**Theorem 3.3.** *Assume that **CL1**(4), **CL2**(4) and **CL3**(4) are satisfied. Then, the sequence*

$$\left(\frac{\sqrt{n}}{b_n}(\widehat{D}_n - D^*)\right)_{n \geq 1}$$

*satisfies the LDP on  $\mathbb{R}$  stated in Theorem 2.3.*

*Proof.* Theorem 3.1, Theorem 3.2 and Theorem 3.3 are proved in Section 4. □

#### 4. PROOF OF THE MAIN RESULTS

For a matter of readability, some notations commonly used in the following proofs have to be introduced. First, for all  $n \geq 1$ , let

$$L_n = \sum_{k=1}^n V_k^2. \tag{4.1}$$

Then, let us define  $M_n$ , for all  $n \geq 1$ , as

$$M_n = \sum_{k=1}^n X_{k-1} V_k \tag{4.2}$$

where  $M_0 = 0$ . For all  $n \geq 1$ , denote by  $\mathcal{F}_n$  the  $\sigma$ -algebra of the events occurring up to time  $n$ ,  $\mathcal{F}_n = \sigma(X_0, \varepsilon_0, V_1, \dots, V_n)$ . We infer from (4.2) that  $(M_n)_{n \geq 0}$  is a locally square-integrable real martingale with respect to the filtration  $\mathbb{F} = (\mathcal{F}_n)_{n \geq 0}$  with predictable quadratic variation given by  $\langle M \rangle_0 = 0$  and for all  $n \geq 1$ ,  $\langle M \rangle_n = \sigma^2 S_{n-1}$ , where

$$S_n = \sum_{k=0}^n X_k^2. \tag{4.3}$$

Moreover,  $(N_n)_{n \geq 0}$  is defined, for all  $n \geq 2$ , as

$$N_n = \sum_{k=2}^n X_{k-2} V_k \tag{4.4}$$

and  $N_0 = N_1 = 0$ . It is not hard to see that  $(N_n)_{n \geq 0}$  is also a locally square-integrable real martingale sharing the same properties than  $(M_n)_{n \geq 0}$ . More precisely, its predictable quadratic variation is given by  $\langle N \rangle_n = \sigma^2 S_{n-2}$ . To conclude, let  $P_0 = 0$  and, for all  $n \geq 1$ ,

$$P_n = \sum_{k=1}^n X_{k-1} X_k. \tag{4.5}$$

To smooth the reading of the following proofs, we introduce some relations.



**Lemma 4.1.** For any  $\eta > 0$ ,

$$\sum_{k=0}^n |X_k|^\eta \leq (1 + \alpha(\eta))|X_0|^\eta + \alpha(\eta)\beta(\eta)|\varepsilon_0|^\eta + \alpha(\eta)\beta(\eta) \sum_{k=1}^n |V_k|^\eta$$

where

$$\alpha(\eta) = (1 - |\theta|)^{-\eta} \quad \text{and} \quad \beta(\eta) = (1 - |\rho|)^{-\eta}.$$

In addition,

$$\max_{1 \leq k \leq n} X_k^2 \leq \alpha(1) X_0^2 + \alpha(2)\beta(1) \varepsilon_0^2 + \alpha(2)\beta(2) \max_{1 \leq k \leq n} V_k^2.$$

*Proof.* The proof follows from (1.1). Details are given in the proof of Lemma A.2 in [3].  $\square$

**Lemma 4.2.** For all  $n \geq 2$ ,

$$\frac{S_n}{n} - \ell = \frac{\ell}{\sigma^2} \left[ \left( \frac{L_n}{n} - \sigma^2 \right) + 2\theta^* \frac{M_n}{n} - 2\theta\rho \frac{N_n}{n} + \frac{R_n}{n} \right] \quad (4.6)$$

where  $L_n$ ,  $M_n$ ,  $S_n$  and  $N_n$  are respectively given by (4.1), (4.2), (4.3) and (4.4),

$$R_n = [2(\theta + \rho)\rho^* - (\theta + \rho)^2 - (\theta\rho)^2]X_n^2 - (\theta\rho)^2 X_{n-1}^2 + 2\rho^* X_n X_{n-1} + \xi_1,$$

and where the remainder term

$$\xi_1 = (1 - 2\theta\rho - \rho^2)X_0^2 + \rho^2\varepsilon_0^2 + 2\theta\rho X_0\varepsilon_0 - 2\rho\rho^*(\varepsilon_0 - X_0)X_0 + 2\rho(\varepsilon_0 - X_0)V_1.$$

In addition, for all  $n \geq 1$ ,

$$\frac{P_n}{n} - \theta^* \frac{S_n}{n} = \frac{1}{1 + \theta\rho} \frac{M_n}{n} + \frac{1}{1 + \theta\rho} \frac{R_n(\theta)}{n} - \theta^* \frac{X_n^2}{n} \quad (4.7)$$

with

$$R_n(\theta) = \theta\rho X_n X_{n-1} + \rho X_0(\varepsilon_0 - X_0).$$

*Proof.* The results follow from direct calculation.  $\square$

#### 4.1. Proof of Theorem 2.1.

Before starting the proof of Theorem 2.1, we need to introduce some technical tools. Denote by  $\ell$  the almost sure limit of  $S_n/n$  [3], given by

$$\ell = \frac{\sigma^2(1 + \theta\rho)}{(1 - \theta^2)(1 - \theta\rho)(1 - \rho^2)}. \quad (4.8)$$

**Lemma 4.3.** Under the assumptions of Theorem 2.1, we have the exponential convergence

$$\frac{S_n}{n} \xrightarrow[b_n^2]{\text{superexp}} \ell \quad (4.9)$$

where  $\ell$  is given by (4.8).

*Proof.* First of all,  $(V_n)$  is a sequence of independent and identically distributed gaussian random variables with zero mean and variance  $\sigma^2 > 0$ . It immediately follows from Cramér-Chernoff's Theorem, expounded *e.g.* in [7], that for all  $\delta' > 0$ ,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P} \left( \left| \frac{L_n}{n} - \sigma^2 \right| > \delta' \right) < 0. \quad (4.10)$$

Since  $b_n^2 = o(n)$ , the latter convergence leads to

$$\frac{L_n}{n} \xrightarrow[b_n^2]{\text{superexp}} \sigma^2, \quad (4.11)$$

ensuring the exponential convergence of  $L_n/n$  to  $\sigma^2$  with speed  $b_n^2$ . Moreover, for all  $\delta > 0$  and a suitable  $t > 0$ , we clearly obtain from Markov's inequality that

$$\mathbb{P}\left(\frac{X_0^2}{n} > \delta\right) \leq \exp(-tn\delta) \mathbb{E}\left[\exp(tX_0^2)\right],$$

which immediately implies *via* **G2**,

$$\frac{X_0^2}{n} \xrightarrow[b_n^2]{\text{superexp}} 0, \quad (4.12)$$

and we get the exponential convergence of  $X_0^2/n$  to 0 with speed  $b_n^2$ . The same is true for  $V_1^2/n$ ,  $\varepsilon_0^2/n$  and more generally for any isolated term in  $\xi_1$  given after (4.6). Let us now focus our attention on  $X_n^2/n$ . The model (1.1) can be rewritten in the vectorial form,

$$\Phi_n = A\Phi_{n-1} + W_n \quad (4.13)$$

where  $\Phi_n = (X_n \ X_{n-1})'$  stands for the lag vector of order 2,  $W_n = (V_n \ 0)'$  and

$$A = \begin{pmatrix} \theta + \rho & -\theta\rho \\ 1 & 0 \end{pmatrix}. \quad (4.14)$$

It is easy to show that the spectral radius of  $A$  is given by  $\rho(A) = \max(|\theta|, |\rho|) < 1$  under the stability conditions. Then,

$$\frac{\|\Phi_n\|^2}{n} \xrightarrow[b_n^2]{\text{superexp}} 0,$$

according to [23], which is clearly sufficient to deduce that

$$\frac{X_n^2}{n} \xrightarrow[b_n^2]{\text{superexp}} 0. \quad (4.15)$$

The exponential convergence of  $R_n/n$  to 0 with speed  $b_n^2$  is achieved following exactly the same lines. To conclude the proof of Lemma 4.3, it remains to study the exponential asymptotic behavior of  $M_n/n$ . For all  $\delta > 0$  and a suitable  $y > 0$ ,

$$\begin{aligned} \mathbb{P}\left(\frac{M_n}{n} > \delta\right) &= \mathbb{P}\left(\frac{M_n}{n} > \delta, \langle M \rangle_n \leq y\right) + \mathbb{P}\left(\frac{M_n}{n} > \delta, \langle M \rangle_n > y\right), \\ &\leq \exp\left(-\frac{n^2\delta^2}{2y}\right) + \mathbb{P}\left(\langle M \rangle_n > y\right), \end{aligned} \quad (4.16)$$

by application of Theorem 4.1 of [4] in the case of a gaussian martingale, and Remark 4.2 that follows. From Lemma 4.1, one can find  $\alpha$  and  $\beta$  such that, for a suitable  $t > 0$ ,

$$\begin{aligned} \mathbb{P}\left(\langle M \rangle_n > y\right) &\leq \mathbb{P}\left(X_0^2 > \frac{y}{3\alpha\sigma^2}\right) + \mathbb{P}\left(\varepsilon_0^2 > \frac{y}{3\beta\sigma^2}\right) + \mathbb{P}\left(L_{n-1} > \frac{y}{3\beta\sigma^2}\right), \\ &\leq 3 \max\left(\exp\left(\frac{-yt}{3\alpha\sigma^2}\right) \mathbb{E}\left[\exp(tX_0^2)\right], \exp\left(\frac{-yt}{3\beta\sigma^2}\right) \mathbb{E}\left[\exp(t\varepsilon_0^2)\right], \right. \\ &\quad \left. \mathbb{P}\left(L_{n-1} > \frac{y}{3\beta\sigma^2}\right)\right). \end{aligned}$$

Let us choose  $y = nx$ , assuming  $x > 3\beta\sigma^4$ . It follows that

$$\frac{1}{b_n^2} \log \mathbb{P}\left(\langle M \rangle_n > nx\right) \leq \frac{\log 3}{b_n^2} + \frac{1}{b_n^2} \max\left(\frac{-nxt}{3\alpha\sigma^2} + \log \mathbb{E}\left[\exp(tX_0^2)\right], \frac{-nxt}{3\beta\sigma^2} + \log \mathbb{E}\left[\exp(t\varepsilon_0^2)\right], \log \mathbb{P}\left(L_{n-1} > \frac{nx}{3\beta\sigma^2}\right)\right).$$

Since  $b_n^2 = o(n)$  and by virtue of (4.10) with  $\delta' = x/(3\beta\sigma^2) - \sigma^2 > 0$ , we obtain that

$$\limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\langle M \rangle_n > nx\right) = -\infty. \quad (4.17)$$

It enables us by (4.16) to deduce that for all  $\delta > 0$ ,

$$\limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{M_n}{n} > \delta\right) = -\infty. \quad (4.18)$$

The same result is also true replacing  $M_n$  by  $-M_n$  in (4.18) since  $M_n$  and  $-M_n$  share the same distribution. Therefore, we find that

$$\frac{M_n}{n} \xrightarrow[b_n^2]{\text{superexp}} 0. \quad (4.19)$$

A similar reasoning leads to the exponential convergence of  $N_n/n$  to 0, with speed  $b_n^2$ . Finally, we obtain (4.9) from Lemma 4.2 together with (4.11), (4.12), (4.15) and (4.19) which achieves the proof of Lemma 4.3.  $\square$

**Corollary 4.4.** *By virtue of Lemma 4.3 and under the same assumptions, we have the exponential convergence*

$$\frac{P_n}{n} \xrightarrow[b_n^2]{\text{superexp}} \ell_1 \quad (4.20)$$

where  $\ell_1 = \theta^*\ell$ .

*Proof.* The proof is immediately derived from previous statements and Lemma 4.2.  $\square$

We are now in the position to prove Theorem 2.1. We shall make use of the following MDP for martingales established by Worms [23].

**Theorem 4.5** (Worms). *Let  $(Y_n)$  be an adapted sequence with values in  $\mathbb{R}^p$ , and  $(V_n)$  a gaussian noise with variance  $\sigma^2 > 0$ . We suppose that  $(Y_n)$  satisfies, for some invertible square matrix  $C$  of order  $p$  and a speed sequence  $(b_n^2)$  such that  $b_n^2 = o(n)$ , the exponential convergence for any  $\delta > 0$ ,*

$$\lim_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\left\|\frac{1}{n} \sum_{k=0}^{n-1} Y_k Y_k' - C\right\| > \delta\right) = -\infty. \quad (4.21)$$

Then, the sequence

$$\left(\frac{1}{b_n \sqrt{n}} \sum_{k=1}^n Y_{k-1} V_k\right)_{n \geq 1}$$

satisfies an LDP on  $\mathbb{R}^p$  of speed  $b_n^2$  and rate function

$$I(x) = \frac{1}{2\sigma^2} x' C^{-1} x. \quad (4.22)$$

*Proof.* The proof of Theorem 4.5 is contained in the one of Theorem 5 of [23] with  $d = 1$ .  $\square$

**Proof of Theorem 2.1.** Let us consider the decomposition

$$\frac{\sqrt{n}}{b_n} \left( \widehat{\theta}_n - \theta^* \right) = \frac{n}{\langle M \rangle_n} A_n + B_n, \quad (4.23)$$

with

$$A_n = \left( \frac{\sigma^2}{1 + \theta\rho} \right) \frac{M_n}{b_n \sqrt{n}} \quad \text{and} \quad B_n = \frac{\sqrt{n}}{b_n} \left( \frac{1}{1 + \theta\rho} \right) \frac{R_n(\theta)}{S_{n-1}},$$

that can be obtained by a straightforward calculation, where the remainder term  $R_n(\theta)$  is defined after (4.7). First, by using the same methodology as in convergence (4.12), we obtain that for all  $\delta > 0$  and for a suitable  $t > 0$ ,

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P} \left( \frac{X_0^2}{b_n \sqrt{n}} > \delta \right) &\leq \lim_{n \rightarrow \infty} \left( -t\delta \frac{\sqrt{n}}{b_n} \right) + \lim_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{E} \left[ \exp(tX_0^2) \right], \\ &= -\infty, \end{aligned} \quad (4.24)$$

since  $b_n = o(\sqrt{n})$ , and the same is true for any isolated term in (4.23) of order 2 whose numerator does not depend on  $n$ . Moreover, under the gaussian assumption on the driven noise  $(V_n)$ , it is not hard to see that

$$\frac{1}{b_n \sqrt{n}} \max_{1 \leq k \leq n} V_k^2 \xrightarrow[b_n^2]{\text{superexp}} 0. \quad (4.25)$$

As a matter of fact, for all  $\delta > 0$  and for all  $t > 0$ ,

$$\begin{aligned} \mathbb{P} \left( \max_{1 \leq k \leq n} V_k^2 \geq \delta b_n \sqrt{n} \right) &= \mathbb{P} \left( \bigcup_{k=1}^n \{V_k^2 \geq \delta b_n \sqrt{n}\} \right) \leq \sum_{k=1}^n \mathbb{P} \left( V_k^2 \geq \delta b_n \sqrt{n} \right), \\ &\leq n \exp(-t\delta b_n \sqrt{n}) \mathbb{E} \left[ \exp(tV_1^2) \right]. \end{aligned}$$

In addition, as soon as  $0 < t < 1/(2\sigma^2)$ ,  $\mathbb{E}[\exp(tV_1^2)] < \infty$ . Consequently,

$$\frac{1}{b_n^2} \log \mathbb{P} \left( \max_{1 \leq k \leq n} V_k^2 \geq \delta b_n \sqrt{n} \right) \leq \frac{\sqrt{n}}{b_n} \left( \frac{\log n}{b_n \sqrt{n}} - t\delta + \frac{\log \mathbb{E}[\exp(tV_1^2)]}{b_n \sqrt{n}} \right)$$

which clearly leads to (4.25). Then, we deduce from (4.24), (4.25) and Lemma 4.1 that

$$\frac{1}{b_n \sqrt{n}} \max_{1 \leq k \leq n} X_k^2 \xrightarrow[b_n^2]{\text{superexp}} 0, \quad (4.26)$$

which of course imply the exponential convergence of  $X_n^2/(b_n \sqrt{n})$  to 0, with speed  $b_n^2$ . Therefore, we obtain that

$$\frac{R_n(\theta)}{b_n \sqrt{n}} \xrightarrow[b_n^2]{\text{superexp}} 0. \quad (4.27)$$

We infer from Lemma 4.3 and Lemma 2 of [23] that the following convergence is satisfied,

$$\frac{n}{S_n} \xrightarrow[b_n^2]{\text{superexp}} \frac{1}{\ell} \quad (4.28)$$

where  $\ell > 0$  is given by (4.8). According to (4.27), the latter convergence and again Lemma 2 of [23], we deduce that

$$B_n \xrightarrow[b_n^2]{\text{superexp}} 0. \quad (4.29)$$

Hence, we obtain from (4.28) that the same is true for

$$A_n \left( \frac{n}{\langle M \rangle_n} - \frac{1}{\sigma^2 \ell} \right) \xrightarrow[b_n^2]{\text{superexp}} 0, \quad (4.30)$$

since Lemma 4.3 together with Theorem 4.5 with  $p = 1$  directly show that  $(M_n/(b_n \sqrt{n}))$  satisfies an LDP with speed  $b_n^2$  and rate function given, for all  $x \in \mathbb{R}$ , by

$$J(x) = \frac{x^2}{2\ell\sigma^2}. \quad (4.31)$$

As a consequence,

$$\frac{\sqrt{n}}{b_n} (\hat{\theta}_n - \theta^*) \xrightarrow[b_n^2]{\text{superexp}} \frac{1}{\ell(1 + \theta\rho)} \frac{M_n}{b_n \sqrt{n}}, \quad (4.32)$$

and this implies that both of them share the same LDP, see Theorem 4.2.13 in [7]. One shall now take advantage of the contraction principle (Theorem 4.2.1 in [7]), to establish that  $(\sqrt{n}(\hat{\theta}_n - \theta^*)/b_n)$  satisfies an LDP with speed  $b_n^2$  and rate function  $I_\theta(x) = J(\ell(1 + \theta\rho)x)$  given by (2.1), that is

$$I_\theta(x) = \frac{x^2}{2\sigma_\theta^2},$$

which achieves the proof of Theorem 2.1.  $\square$

#### 4.2. Proof of Theorem 2.2.

We need to introduce some more notations. For all  $n \geq 2$ , let

$$Q_n = \sum_{k=2}^n X_{k-2} X_k. \quad (4.33)$$

In addition, for all  $n \geq 1$ , denote

$$T_n = 1 + \theta^* \rho^* - \left(1 + \rho^* (\hat{\theta}_n + \theta^*)\right) \frac{S_n}{S_{n-1}} + \left(2\rho^* + \hat{\theta}_n + \theta^*\right) \frac{P_n}{S_{n-1}} - \frac{Q_n}{S_{n-1}}, \quad (4.34)$$

where  $S_n$  and  $P_n$  are respectively given by (4.3) and (4.5). Finally, for all  $n \geq 0$ , let

$$J_n = \sum_{k=0}^n \hat{\varepsilon}_k^2 \quad (4.35)$$

where the residual sequence  $(\hat{\varepsilon}_n)$  is given in (1.3). A set of additional technical tools has to be expounded to make the proof of Theorem 2.2 more tractable.

**Corollary 4.6.** *By virtue of Lemma 4.3 and under the same assumptions, we have the exponential convergence*

$$\frac{Q_n}{n} \xrightarrow[b_n^2]{\text{superexp}} \ell_2$$

where  $\ell_2 = ((\theta + \rho)\theta^* - \theta\rho)\ell$ .

*Proof.* The proof of Corollary 4.6 immediately follows from the relation

$$\frac{Q_n}{n} - ((\theta + \rho)\theta^* - \theta\rho) \frac{S_n}{n} = \theta^* \frac{M_n}{n} + \frac{N_n}{n} + \frac{\xi_n^Q}{n} \quad (4.36)$$

where  $\xi_n^Q$  is a residual term made of isolated terms such that

$$\frac{\xi_n^Q}{n} \xrightarrow[b_n^2]{\text{superexp}} 0,$$

see *e.g.* the proof of Theorem 3.2 in [3] where more details are given on  $\xi_n^Q$ .  $\square$

**Lemma 4.7.** *Under the assumptions of Theorem 2.2, we have the exponential convergence*

$$A_n \xrightarrow[b_n^2]{\text{superexp}} A$$

where

$$A_n = \frac{n}{1 + \theta\rho} \begin{pmatrix} 1 & 0 \\ S_{n-1} & \\ T_n & -(\theta + \rho) \\ J_{n-1} & J_{n-1} \end{pmatrix}, \quad (4.37)$$

and

$$A = \frac{1}{\ell(1 + \theta\rho)(1 - (\theta^*)^2)} \begin{pmatrix} 1 - (\theta^*)^2 & 0 \\ \theta\rho + (\theta^*)^2 & -(\theta + \rho) \end{pmatrix}. \quad (4.38)$$

*Proof.* Via (4.28), we directly obtain the exponential convergence,

$$\frac{1}{(1 + \theta\rho)} \frac{n}{S_{n-1}} \xrightarrow[b_n^2]{\text{superexp}} \frac{1}{\ell(1 + \theta\rho)}. \quad (4.39)$$

The combination of Lemma 4.3, Corollary 4.4, Corollary 4.6 and Lemma 2 of [23] shows, after a simple calculation, that

$$T_n \xrightarrow[b_n^2]{\text{superexp}} (\theta^*)^2 + \theta\rho. \quad (4.40)$$

Moreover,  $J_n$  given by (4.35) can be rewritten as

$$J_n = S_n - 2\hat{\theta}_n P_n + \hat{\theta}_n^2 S_{n-1},$$

which leads, *via* Lemma 2 of [23], to

$$\frac{J_n}{n} \xrightarrow[b_n^2]{\text{superexp}} \ell(1 - (\theta^*)^2). \quad (4.41)$$

Convergences (4.40) and (4.41) imply

$$\left( \frac{n}{1 + \theta\rho} \right) \frac{T_n}{J_{n-1}} \xrightarrow[b_n^2]{\text{superexp}} \frac{(\theta^*)^2 + \theta\rho}{\ell(1 + \theta\rho)(1 - (\theta^*)^2)}, \quad (4.42)$$

and finally,

$$\left( \frac{n}{1 + \theta\rho} \right) \frac{\theta + \rho}{J_{n-1}} \xrightarrow[b_n^2]{\text{superexp}} \frac{\theta + \rho}{\ell(1 + \theta\rho)(1 - (\theta^*)^2)}. \quad (4.43)$$

Finally, (4.39) together with (4.42) and (4.43) achieve the proof of Lemma 4.7.  $\square$

**Proof of Theorem 2.2.** We shall make use of the decomposition

$$\frac{\sqrt{n}}{b_n} \begin{pmatrix} \hat{\theta}_n - \theta^* \\ \hat{\rho}_n - \rho^* \end{pmatrix} = \frac{1}{b_n \sqrt{n}} A_n Z_n + B_n, \quad (4.44)$$

where  $A_n$  is given by (4.37),  $(Z_n)_{n \geq 0}$  is the 2-dimensional vector martingale given by

$$Z_n = \begin{pmatrix} M_n \\ N_n \end{pmatrix}, \quad (4.45)$$

and where the remainder term

$$B_n = \frac{1}{(1 + \theta\rho)} \frac{\sqrt{n}}{b_n} \begin{pmatrix} \frac{R_n(\theta)}{S_{n-1}} \\ \frac{R_n(\rho)}{J_{n-1}} \end{pmatrix}. \quad (4.46)$$

The first component  $R_n(\theta)$  is given in (4.7) while  $R_n(\rho)$ , whose definition may be found in the proof of Theorem 3.2 in [3], is made of isolated terms. Consequently, (4.24) and (4.27) are sufficient to ensure that

$$\frac{R_n(\theta)}{b_n\sqrt{n}} \xrightarrow[b_n^2]{\text{superexp}} 0 \quad \text{and} \quad \frac{R_n(\rho)}{b_n\sqrt{n}} \xrightarrow[b_n^2]{\text{superexp}} 0.$$

Therefore, we obtain that

$$B_n \xrightarrow[b_n^2]{\text{superexp}} 0. \quad (4.47)$$

In addition, it follows from Lemma 4.7 and Theorem 4.5 with  $p = 2$  that  $(Z_n/(b_n\sqrt{n}))$  satisfies an LDP on  $\mathbb{R}^2$  with speed  $b_n^2$  and rate function given, for all  $x \in \mathbb{R}^2$ , by

$$J(x) = \frac{1}{2\sigma^2} x' \Lambda^{-1} x, \quad (4.48)$$

where

$$\Lambda = \ell \begin{pmatrix} 1 & \theta^* \\ \theta^* & 1 \end{pmatrix}, \quad (4.49)$$

since we have the exponential convergence

$$\frac{\langle Z \rangle_n}{n} \xrightarrow[b_n^2]{\text{superexp}} \sigma^2 \Lambda \quad (4.50)$$

by application of Lemma 4.3 and Corollary 4.4. One observes that  $\Lambda$  is invertible. As a consequence,

$$\frac{1}{b_n\sqrt{n}} (A_n - A) Z_n \xrightarrow[b_n^2]{\text{superexp}} 0, \quad (4.51)$$

and we deduce from (4.44) that

$$\frac{\sqrt{n}}{b_n} \begin{pmatrix} \widehat{\theta}_n - \theta^* \\ \widehat{\rho}_n - \rho^* \end{pmatrix} \xrightarrow[b_n^2]{\text{superexp}} \frac{1}{b_n\sqrt{n}} A Z_n. \quad (4.52)$$

This of course implies that both of them share the same LDP, see Theorem 4.2.13 in [7]. The contraction principle (Theorem 4.2.1 in [7]) enables us to conclude that the rate function of the LDP on  $\mathbb{R}^2$  with speed  $b_n^2$  associated with equivalence (4.52) is given, for all  $x \in \mathbb{R}^2$ , by  $K(x) = J(A^{-1}x)$ , that is

$$K(x) = \frac{1}{2} x' \Gamma^{-1} x,$$

where  $\Gamma = \sigma^2 A \Lambda A'$  is given by (1.10), and where we shall suppose that  $\theta \neq -\rho$  to ensure that  $A$  is invertible. In particular, the latter result also implies that the rate function of the LDP on  $\mathbb{R}$  with speed  $b_n^2$  associated with  $(\sqrt{n}(\hat{\rho}_n - \rho^*)/b_n)$  is given, for all  $x \in \mathbb{R}$ , by

$$I_\rho(x) = \frac{x^2}{2\sigma_\rho^2},$$

where  $\sigma_\rho^2$  is the last element of the matrix  $\Gamma$ . This achieves the proof of Theorem 2.2.  $\square$

### 4.3. Proof of Theorem 2.3.

For all  $n \geq 1$ , denote by  $f_n$  the explosion coefficient associated with  $J_n$  given by (4.35), that is

$$f_n = \frac{J_n - J_{n-1}}{J_n} = \frac{\hat{\varepsilon}_n^2}{J_n}. \quad (4.53)$$

It follows from decomposition (C.4) in [3] that

$$\frac{\sqrt{n}}{b_n} (\hat{D}_n - D^*) = -2 \frac{\sqrt{n}}{b_n} (1 - f_n) (\hat{\rho}_n - \rho^*) + \frac{\sqrt{n}}{b_n} \zeta_n, \quad (4.54)$$

where the remainder term  $\zeta_n$  is made of negligible terms, that is

$$\zeta_n = 2(\rho^* - 1)f_n + \frac{\hat{\varepsilon}_n^2 - \hat{\varepsilon}_0^2}{J_n}.$$

From the definition of  $(\hat{\varepsilon}_n)$  in (1.3), from (4.26), (4.41) and considering that  $\hat{\varepsilon}_0 = X_0$ , we clearly have that

$$\frac{\sqrt{n}}{b_n} \zeta_n \xrightarrow[b_n^2]{\text{superexp}} 0 \quad \text{and} \quad f_n \xrightarrow[b_n^2]{\text{superexp}} 0.$$

As a consequence,

$$\frac{\sqrt{n}}{b_n} (\hat{D}_n - D^*) \overset{\sim}{\xrightarrow[b_n^2]{\text{superexp}}} -2 \frac{\sqrt{n}}{b_n} (\hat{\rho}_n - \rho^*), \quad (4.55)$$

and this implies that both of them share the same LDP. The contraction principle [7] enables us to conclude that the rate function of the LDP on  $\mathbb{R}$  with speed  $b_n^2$  associated with equivalence (4.55) is given, for all  $x \in \mathbb{R}$ , by  $I_D(x) = I_\rho(-x/2)$ , that is

$$I_D(x) = \frac{x^2}{2\sigma_D^2},$$

which achieves the proof of Theorem 2.3.  $\square$

### 4.4. Proofs of Theorem 3.1, Theorem 3.2 and Theorem 3.3.

We shall now propose a technical lemma ensuring that all results already proved under the gaussian assumption still hold under the Chen-Ledoux type condition.

**Lemma 4.8.** *Under CL1(4), CL2(4) and CL3(4), all exponential convergences of Lemma 4.3, Corollary 4.4, Corollary 4.6 and Lemma 4.7 still hold.*



*Proof.* Following the same methodology as the one used to establish (4.27), we get

$$\mathbb{P}\left(\max_{1 \leq k \leq n} V_k^2 \geq \delta b_n \sqrt{n}\right) \leq \sum_{k=1}^n \mathbb{P}\left(V_k^2 \geq \delta b_n \sqrt{n}\right) = n \mathbb{P}\left(V_1^2 \geq \delta b_n \sqrt{n}\right).$$

Via **CL1(2)**, **CL2(2)**, **CL3(2)** and the same reasoning,

$$\frac{X_n^2}{b_n \sqrt{n}} \xrightarrow[b_n^2]{\text{superexp}} 0, \quad (4.56)$$

and Cauchy-Schwarz inequality implies that this is also the case for any isolated term of order 2, such as  $X_n X_{n-1} / (b_n \sqrt{n})$ . This allows us to control each remainder term. Note that under **CL2(4)** and **CL3(4)** and using (4.56),  $\varepsilon_0^4/n$ ,  $X_0^4/n$ ,  $\varepsilon_0^2/n$ ,  $X_0^2/n$  and  $X_n^2/n$  also exponentially converge to 0, since  $b_n \sqrt{n} = o(n)$ . Moreover, it follows from Theorem 2.2 of [14] under **CL1(2)**, that

$$\frac{L_n}{n} \xrightarrow[b_n^2]{\text{superexp}} \sigma^2. \quad (4.57)$$

Furthermore, since  $(M_n)$  is a locally square integrable martingale, we infer from Theorem 2.1 of [4] that for all  $x, y > 0$ ,

$$\mathbb{P}\left(|M_n| > x, \langle M \rangle_n + [M]_n \leq y\right) \leq 2 \exp\left(-\frac{x^2}{2y}\right), \quad (4.58)$$

where the predictable quadratic variation  $\langle M \rangle_n = \sigma^2 S_{n-1}$  is described in (4.3) and the total quadratic variation is given by  $[M]_0 = 0$  and, for all  $n \geq 1$ , by

$$[M]_n = \sum_{k=1}^n X_{k-1}^2 V_k^2. \quad (4.59)$$

According to (4.58), we have for all  $\delta > 0$  and a suitable  $b > 0$ ,

$$\begin{aligned} \mathbb{P}\left(\frac{|M_n|}{n} > \delta\right) &\leq \mathbb{P}\left(|M_n| > \delta n, \langle M \rangle_n + [M]_n \leq nb\right) + \mathbb{P}\left(\langle M \rangle_n + [M]_n > nb\right), \\ &\leq 2 \exp\left(-\frac{n\delta^2}{2b}\right) + \mathbb{P}\left(\langle M \rangle_n + [M]_n > nb\right). \end{aligned}$$

Consequently,

$$\limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{|M_n|}{n} > \delta\right) \leq \limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\langle M \rangle_n + [M]_n > nb\right). \quad (4.60)$$

Moreover, for all  $n \geq 1$ , let us define

$$T_n = \sum_{k=0}^n X_k^4 \quad \text{and} \quad \Gamma_n = \sum_{k=1}^n V_k^4.$$

From Lemma 4.1 and for  $n$  large enough, one can find  $\gamma > 0$  such that

$$T_n \leq \gamma \Gamma_n$$

under **CL2(4)** and **CL3(4)**. According to Theorem 2.2 of [14] under **CL1(4)**, we also have the exponential convergence,

$$\frac{\Gamma_n}{n} \xrightarrow[b_n^2]{\text{superexp}} \tau^4, \quad (4.61)$$

where  $\tau^4 = \mathbb{E}[V_1^4]$ , leading, *via* Cauchy-Schwarz inequality, to

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P} \left( \frac{[M]_n}{n} > \delta \right) &\leq \limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P} \left( \frac{\Gamma_n}{n} > \frac{\delta}{\sqrt{\gamma}} \right), \\ &= -\infty, \end{aligned} \quad (4.62)$$

where  $\delta > \tau^4 \sqrt{\gamma}$ . Exploiting (4.57) and again Lemma 4.1, the same result can be achieved for  $\langle M \rangle_n/n$  under **CL1**(2) and  $\delta > \sigma^4 \gamma$ . As a consequence, it follows from (4.62) that

$$\limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P} \left( \frac{\langle M \rangle_n + [M]_n}{n} > b \right) = -\infty, \quad (4.63)$$

as soon as  $b > \sigma^4 \gamma + \tau^4 \sqrt{\gamma}$ . Therefore, the exponential convergence of  $M_n/n$  to 0 with speed  $b_n^2$  is obtained *via* (4.60) and (4.63), that is, for all  $\delta > 0$ ,

$$\limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P} \left( \frac{|M_n|}{n} > \delta \right) = -\infty. \quad (4.64)$$

Explicitly, (4.64) is equivalent of (4.19) which was the main element for the proof of Lemma 4.3, and the same obviously holds for  $N_n/n$ . In consequence, one can proceed similarly to establish Corollary 4.4, Corollary 4.6 and Lemma 4.7. Indeed, hypotheses **CL2**(4) and **CL3**(4) together with exponential convergences (4.56), (4.57) and (4.64) are sufficient to achieve the proof of Lemma 4.8.  $\square$

Let us introduce a simplified version of Puhalskii's result [21] applied to a sequence of martingale differences, and two technical lemmas that shall help us to prove our results.

**Theorem 4.9** (Puhalskii). *Let  $(m_j^n)_{1 \leq j \leq n}$  be a triangular array of martingale differences with values in  $\mathbb{R}^d$ , with respect to a filtration  $(\mathcal{F}_n)_{n \geq 1}$ . Let  $(b_n)$  be a sequence of real numbers satisfying (1.13). Suppose that there exists a symmetric positive-semidefinite matrix  $Q$  such that*

$$\frac{1}{n} \sum_{k=1}^n \mathbb{E} \left[ m_k^n (m_k^n)' \mid \mathcal{F}_{k-1} \right] \xrightarrow[b_n^2]{\text{superexp}} Q. \quad (4.65)$$

Suppose that there exists a constant  $c > 0$  such that, for each  $1 \leq k \leq n$ ,

$$|m_k^n| \leq c \frac{\sqrt{n}}{b_n} \quad \text{a.s.} \quad (4.66)$$

Suppose also that, for all  $a > 0$ , we have the exponential Lindeberg's condition

$$\frac{1}{n} \sum_{k=1}^n \mathbb{E} \left[ |m_k^n|^2 \mathbf{I}_{\{|m_k^n| \geq a \frac{\sqrt{n}}{b_n}\}} \mid \mathcal{F}_{k-1} \right] \xrightarrow[b_n^2]{\text{superexp}} 0. \quad (4.67)$$

Then, the sequence

$$\left( \frac{1}{b_n \sqrt{n}} \sum_{k=1}^n m_k^n \right)_{n \geq 1}$$

satisfies an LDP on  $\mathbb{R}^d$  with speed  $b_n^2$  and rate function

$$\Lambda^*(v) = \sup_{\lambda \in \mathbb{R}^d} \left( \lambda' v - \frac{1}{2} \lambda' Q \lambda \right).$$

In particular, if  $Q$  is invertible,

$$\Lambda^*(v) = \frac{1}{2} v' Q^{-1} v. \quad (4.68)$$

*Proof.* The proof of Theorem 4.9 is contained *e.g.* in the proof of Theorem 3.1 in [21].  $\square$

**Lemma 4.10.** *Under  $\mathbf{CL1}(a)$ ,  $\mathbf{CL2}(a)$  and  $\mathbf{CL3}(a)$  for any  $a > 2$ , we have for all  $\delta > 0$ ,*

$$\limsup_{R \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P} \left( \frac{1}{n} \sum_{k=1}^n X_k^2 \mathbf{I}_{\{|X_k| > R\}} > \delta \right) = -\infty.$$

**Remark 4.1.** *Lemma 4.10 implies that the exponential Lindeberg's condition given by (4.67) is satisfied.*

*Proof.* From Lemma 4.1, for any  $\eta > 0$  and  $n$  large enough, one can find  $\gamma > 0$  such that

$$\sum_{k=0}^n |X_k|^{2+\eta} \leq \gamma \sum_{k=1}^n |V_k|^{2+\eta} \quad (4.69)$$

under  $\mathbf{CL2}(2 + \eta)$  and  $\mathbf{CL3}(2 + \eta)$ . If we suppose that  $\mathbf{CL1}(2 + \eta)$  holds, then it follows that, for  $R > 0$ ,

$$R^\eta \sum_{k=1}^n X_{k-1}^2 \mathbf{I}_{\{|X_{k-1}| > R\}} \leq \sum_{k=1}^n |X_{k-1}|^{2+\eta} \leq \gamma \sum_{k=1}^n |V_k|^{2+\eta},$$

for  $n$  large enough and  $\eta > 0$ , leading to

$$\frac{1}{b_n^2} \log \mathbb{P} \left( \frac{1}{n} \sum_{k=1}^n X_{k-1}^2 \mathbf{I}_{\{|X_{k-1}| > R\}} > \delta \right) \leq \frac{1}{b_n^2} \log \mathbb{P} \left( \frac{1}{n} \sum_{k=1}^n |V_k|^{2+\eta} > \frac{\delta}{\gamma} R^\eta \right).$$

Using Theorem 2.2 of [14] and letting  $R$  go to infinity, we immediately reach the end of the proof of Lemma 4.10.  $\square$

**Remark 4.2.** *The same result can be achieved under the less restrictive  $\mathbf{CL1}(2)$  condition, via a technical proof using the empirical measure associated with the geometric ergodic Markov chain  $(X_n)_{n \geq 0}$ . A same reasoning can be found in [9].*

**Lemma 4.11.** *Under  $\mathbf{CL1}(4)$ ,  $\mathbf{CL2}(4)$  and  $\mathbf{CL3}(4)$ , the sequence*

$$\left( \frac{M_n}{b_n \sqrt{n}} \right)_{n \geq 1}$$

*satisfies an LDP on  $\mathbb{R}$  with speed  $b_n^2$  and rate function*

$$J(x) = \frac{x^2}{2\ell\sigma^2} \quad (4.70)$$

*where  $\ell$  is given by (4.8).*

*Proof.* From now on, in order to apply Puhalskii's result concerning MDP for martingales, we introduce the following modification of the martingale  $(M_n)_{n \geq 0}$ , for  $r > 0$  and  $R > 0$ ,

$$M_n^{(r,R)} = \sum_{k=1}^n X_{k-1}^{(r)} V_k^{(R)} \quad (4.71)$$

where, for all  $1 \leq k \leq n$ ,

$$X_k^{(r)} = X_k \mathbf{I}_{\{|X_k| \leq r \frac{\sqrt{n}}{b_n}\}} \quad \text{and} \quad V_k^{(R)} = V_k \mathbf{I}_{\{|V_k| \leq R\}} - \mathbb{E} \left[ V_k \mathbf{I}_{\{|V_k| \leq R\}} \right]. \quad (4.72)$$

Then, we have to prove that for all  $r > 0$  the sequence  $(M_n^{(r,R)})$  is an exponentially good approximation of  $(M_n)$  as  $R$  goes to infinity, see *e.g.* Definition 4.2.14 in [7]. This approximation, in the sense of the large deviations, is described by the following convergence, for all  $r > 0$  and all  $\delta > 0$ ,

$$\limsup_{R \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P} \left( \frac{|M_n - M_n^{(r,R)}|}{b_n \sqrt{n}} > \delta \right) = -\infty. \quad (4.73)$$

From Lemma 4.8, and since  $\langle M \rangle_n = \sigma^2 S_{n-1}$ , we have

$$\frac{\langle M \rangle_n}{n} \xrightarrow{\text{superexp}}_{b_n^2} \sigma^2 \ell. \quad (4.74)$$

From Lemma 4.10 and Remark 4.1, we also have for all  $r > 0$ ,

$$\frac{1}{n} \sum_{k=0}^n X_k^2 \mathbf{I}_{\{|X_k| > r \frac{\sqrt{n}}{b_n}\}} \xrightarrow{\text{superexp}}_{b_n^2} 0. \quad (4.75)$$

We introduce the following notations,

$$\sigma_R^2 = \mathbb{E} \left[ (V_1^{(R)})^2 \right] \quad \text{and} \quad S_n^{(r)} = \sum_{k=0}^n (X_k^{(r)})^2.$$

Then, we easily transfer properties (4.74) and (4.75) to the truncated martingale  $(M_n^{(r,R)})_{n \geq 0}$ . We have for all  $R > 0$  and all  $r > 0$ ,

$$\frac{\langle M^{(r,R)} \rangle_n}{n} = \sigma_R^2 \frac{S_{n-1}^{(r)}}{n} = -\sigma_R^2 \left( \frac{S_{n-1}}{n} - \frac{S_{n-1}^{(r)}}{n} \right) + \sigma_R^2 \frac{S_{n-1}}{n} \xrightarrow{\text{superexp}}_{b_n^2} \sigma_R^2 \ell$$

which ensures that (4.65) is satisfied for the martingale  $(M_n^{(r,R)})_{n \geq 0}$ . Note also that Lemma 4.8 and Remark 4.1 work for the martingale  $(M_n^{(r,R)})_{n \geq 0}$ . So, for all  $r > 0$ , the exponential Lindeberg's condition and thus (4.67) are satisfied for  $(M_n^{(r,R)})_{n \geq 0}$ . By Theorem 4.9, we deduce that  $(M_n^{(r,R)}/b_n \sqrt{n})$  satisfies an LDP on  $\mathbb{R}$  with speed  $b_n^2$  and rate function

$$J_R(x) = \frac{x^2}{2\sigma_R^2 \ell}. \quad (4.76)$$

We intend to transfer the MDP result for the martingale  $(M_n)_{n \geq 0}$  by proving relation (4.73). For that purpose, let us now introduce the following decomposition,

$$M_n - M_n^{(r,R)} = L_n^{(r)} + F_n^{(r,R)}$$

where

$$L_n^{(r)} = \sum_{k=1}^n \left( X_{k-1} - X_{k-1}^{(r)} \right) V_k \quad \text{and} \quad F_n^{(r,R)} = \sum_{k=1}^n \left( V_k - V_k^{(R)} \right) X_{k-1}^{(r)}.$$

One has to show that for all  $r > 0$ ,

$$\frac{L_n^{(r)}}{b_n \sqrt{n}} \xrightarrow{\text{superexp}}_{b_n^2} 0, \quad (4.77)$$

and, for all  $r > 0$  and all  $\delta > 0$ , that

$$\limsup_{R \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P} \left( \frac{|F_n^{(r,R)}|}{b_n \sqrt{n}} > \delta \right) = -\infty. \quad (4.78)$$

Via inequality (4.69), for  $n$  large enough,

$$\begin{aligned} \frac{|L_n^{(r)}|}{b_n \sqrt{n}} &= \frac{1}{b_n \sqrt{n}} \left| \sum_{k=1}^n X_{k-1} \mathbf{1}_{\{|X_{k-1}| > r \frac{\sqrt{n}}{b_n}\}} V_k \right|, \\ &\leq \frac{1}{b_n \sqrt{n}} \left( r \frac{\sqrt{n}}{b_n} \right)^{-\eta} \left( \sum_{k=1}^n |X_{k-1}|^{2+\eta} \right)^{1/2} \left( \sum_{k=1}^n V_k^2 |X_{k-1}|^\eta \right)^{1/2}, \\ &\leq \lambda(r, \eta, \gamma) \left( \frac{b_n}{\sqrt{n}} \right)^{\eta-1} \frac{1}{n} \sum_{k=1}^n |V_k|^{2+\eta} \end{aligned} \quad (4.79)$$

by virtue of Hölder's inequality, where  $\lambda(r, \eta, \gamma) > 0$  can be easily evaluated. As a consequence, for all  $\delta > 0$ ,

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P} \left( \frac{|L_n^{(r)}|}{b_n \sqrt{n}} > \delta \right) &\leq \limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P} \left( \frac{1}{n} \sum_{k=1}^n |V_k|^{2+\eta} > \frac{\delta}{\lambda(r, \eta, \gamma)} \left( \frac{\sqrt{n}}{b_n} \right)^{\eta-1} \right), \\ &= -\infty, \end{aligned} \quad (4.80)$$

as soon as  $\eta > 1$ , by application of Theorem 2.2 of [14] under **CL1**( $2 + \eta$ ), since

$$\lim_{n \rightarrow \infty} \left( \frac{\sqrt{n}}{b_n} \right)^{\eta-1} = \infty.$$

We deduce that

$$\frac{L_n^{(r)}}{b_n \sqrt{n}} \xrightarrow[b_n^2]{\text{superexp}} 0, \quad (4.81)$$

which achieves the proof of (4.77), under **CL1**( $2 + \eta$ ), **CL2**( $2 + \eta$ ) and **CL3**( $2 + \eta$ ) for  $\eta > 1$ . On the other hand,  $(F_n^{(r,R)})_{n \geq 0}$  is a locally square-integrable real martingale whose predictable quadratic variation is given by  $\langle F^{(r,R)} \rangle_0 = 0$  and, for all  $n \geq 1$ , by

$$\langle F^{(r,R)} \rangle_n = \mathbb{E} \left[ \left( V_1 - V_1^{(R)} \right)^2 \right] S_{n-1}^{(r)}.$$

To prove (4.78), we will use Theorem 1 of [8]. For  $R$  large enough and all  $k \geq 1$ , we have

$$\begin{aligned} \mathbb{P} \left( \left| X_{k-1}^{(r)} \left( V_k - V_k^{(R)} \right) \right| > b_n \sqrt{n} \mid \mathcal{F}_{k-1} \right) &\leq \mathbb{P} \left( \left| V_k - V_k^{(R)} \right| > \frac{b_n^2}{r} \right), \\ &= \mathbb{P} \left( \left| V_1 - V_1^{(R)} \right| > \frac{b_n^2}{r} \right) = 0. \end{aligned}$$

This implies that

$$\limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \left( n \operatorname{ess\,sup}_{k \geq 1} \mathbb{P} \left( \left| X_{k-1}^{(r)} \left( V_k - V_k^{(R)} \right) \right| > b_n \sqrt{n} \mid \mathcal{F}_{k-1} \right) \right) = -\infty. \quad (4.82)$$

For all  $\nu > 0$  and all  $\delta > 0$ , we obtain from Lemma 4.10 and Remark 4.1, that

$$\limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P} \left( \frac{1}{n} \sum_{k=1}^n \left( X_{k-1}^{(r)} \right)^2 \mathbb{I}_{\{|X_{k-1}^{(r)}| > \nu \frac{\sqrt{n}}{b_n}\}} > \delta \right) \leq \\ \limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P} \left( \frac{1}{n} \sum_{k=1}^n X_{k-1}^2 \mathbb{I}_{\{|X_{k-1}| > \nu \frac{\sqrt{n}}{b_n}\}} > \delta \right) = -\infty.$$

Finally, from Lemma 4.8, Lemma 4.10 and Remark 4.1, it follows that

$$\frac{\langle F^{(r,R)} \rangle_n}{n} = Q_R \frac{S_{n-1}^{(r)}}{n} = -Q_R \left( \frac{S_{n-1}}{n} - \frac{S_{n-1}^{(r)}}{n} \right) + Q_R \frac{S_{n-1}}{n} \xrightarrow[b_n^2]{\text{superexp}} Q_R \ell$$

where

$$Q_R = \mathbb{E} \left[ \left( V_1 - V_1^{(R)} \right)^2 \right],$$

and  $\ell$  is given by (4.8). Moreover, it is clear that  $Q_R$  converges to 0 as  $R$  goes to infinity. Consequently, we infer from Theorem 1 of [8] that  $(F_n^{(r,R)}/(b_n \sqrt{n}))$  satisfies an LDP on  $\mathbb{R}$  of speed  $b_n^2$  and rate function

$$I_R(x) = \frac{x^2}{2Q_R \ell}.$$

In particular, this implies that for all  $\delta > 0$ ,

$$\limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P} \left( \frac{|F_n^{(r,R)}|}{b_n \sqrt{n}} > \delta \right) = -\frac{\delta^2}{2Q_R \ell}, \quad (4.83)$$

and letting  $R$  go to infinity clearly leads to the end of the proof of (4.78). We are able to conclude now that  $(M_n^{(r,R)}/(b_n \sqrt{n}))$  is an exponentially good approximation of  $(M_n/(b_n \sqrt{n}))$ . By application of Theorem 4.2.16 in [7], we find that  $(M_n/(b_n \sqrt{n}))$  satisfies an LDP on  $\mathbb{R}$  with speed  $b_n^2$  and rate function

$$\tilde{J}(x) = \sup_{\delta > 0} \liminf_{R \rightarrow \infty} \inf_{z \in B_{x,\delta}} J_R(z),$$

where  $J_R$  is given in (4.76) and  $B_{x,\delta}$  denotes the ball  $\{z : |z - x| < \delta\}$ . The identification of the rate function  $\tilde{J} = J$ , where  $J$  is given in (4.70) is done easily, which concludes the proof of Lemma 4.11.  $\square$

**Lemma 4.12.** *Under CL1(4), CL2(4) and CL3(4), the sequence*

$$\left( \frac{1}{b_n \sqrt{n}} \begin{pmatrix} M_n \\ N_n \end{pmatrix} \right)_{n \geq 1}$$

*satisfies an LDP on  $\mathbb{R}^2$  with speed  $b_n^2$  and rate function*

$$J(x) = \frac{1}{2\sigma^2} x' \Lambda^{-1} x \quad (4.84)$$

*where  $\Lambda$  is given by (4.49).*

*Proof.* We follow the same approach as in the proof of Lemma 4.11. We shall consider the 2-dimensional vector martingale  $(Z_n)_{n \geq 0}$  defined in (4.45). In order to apply Theorem 4.9, we introduce the following truncation of the martingale  $(Z_n)_{n \geq 0}$ , for  $r > 0$  and  $R > 0$ ,

$$Z_n^{(r,R)} = \begin{pmatrix} M_n^{(r,R)} \\ N_n^{(r,R)} \end{pmatrix}$$

where  $M_n^{(r,R)}$  is given in (4.71) and where  $N_n^{(r,R)}$  is defined in the same manner, that is, for all  $n \geq 2$ ,

$$N_n^{(r,R)} = \sum_{k=2}^n X_{k-2}^{(r)} V_k^{(R)} \tag{4.85}$$

with  $X_n^{(r)}$  and  $V_n^{(R)}$  given by (4.72). The exponential convergence (4.50) still holds, by virtue of Lemma 4.8, which immediately implies hypothesis (4.65). In addition, Lemma 4.10 ensures that, for all  $r > 0$ ,

$$\frac{1}{n} \sum_{k=0}^n X_k^2 \mathbf{I}_{\{|X_k| > r \frac{\sqrt{n}}{b_n}\}} \xrightarrow{\text{superexp}} 0, \tag{4.86}$$

justifying hypothesis (4.67). Via Theorem 4.9,  $(Z_n^{(r,R)} / (b_n \sqrt{n}))$  satisfies an LDP on  $\mathbb{R}^2$  with speed  $b_n^2$  and rate function  $J_R$  given by

$$J_R(x) = \frac{1}{2\sigma_R^2} x' \Lambda^{-1} x. \tag{4.87}$$

Finally, it is straightforward to prove that  $(Z_n^{(r,R)} / (b_n \sqrt{n}))$  is an exponentially good approximation of  $(Z_n / (b_n \sqrt{n}))$ . By application of Theorem 4.2.16 in [7], we deduce that  $(Z_n / (b_n \sqrt{n}))$  satisfies an LDP on  $\mathbb{R}^2$  with speed  $b_n^2$  and rate function given by

$$\tilde{J}(x) = \sup_{\delta > 0} \liminf_{R \rightarrow \infty} \inf_{z \in B_{x,\delta}} J_R(z),$$

where  $J_R$  is given in (4.87) and  $B_{x,\delta}$  denotes the ball  $\{z : |z - x| < \delta\}$ . The identification of the rate function  $\tilde{J} = J$  is done easily, which concludes the proof of Lemma 4.12.  $\square$

**Proofs of Theorem 3.1, Theorem 3.2 and Theorem 3.3.** The residuals appearing in the decompositions (4.23), (4.44) and (4.54) still converge exponentially to zero under **CL1(4)**, **CL2(4)** and **CL3(4)**, with speed  $b_n^2$ , as it was already proved. Therefore, for a better readability, we may skip the most accessible parts of these proofs whose development merely consists in following the same lines as those in the proofs of Theorem 2.1, Theorem 2.2 and Theorem 2.3, taking advantage of Lemma 4.11 and Lemma 4.12, and applying the contraction principle given *e.g.* in [7].  $\square$

**Acknowledgements.** *The authors thank Bernard Bercu and Arnaud Guillin for all their advices and suggestions during the preparation of this work. The authors are also very grateful to the Associate Editor and the Reviewer for spending time to evaluate this manuscript and for providing comments and suggestions that improve the paper substantially.*

## REFERENCES

- [1] ARCONES, M. A. The large deviation principle for stochastic processes I. *Theor. Probab. Appl.* 47 (2003), 567–583.
- [2] ARCONES, M. A. The large deviation principle for stochastic processes II. *Theor. Probab. Appl.* 48 (2003), 19–44.
- [3] BERCU, B., AND PROÏA, F. A sharp analysis on the asymptotic behavior of the Durbin-Watson statistic for the first-order autoregressive process. *ESAIM Probab. Stat.* 16 (2012).
- [4] BERCU, B., AND TOUATI, A. Exponential inequalities for self-normalized martingales with applications. *Ann. Appl. Probab.* 18, no.5 (2008), 1848–1869.
- [5] CHEN, X. Moderate deviations for  $m$ -dependent random variables with Banach space value. *Statis. and Probab. Letters.* 35 (1998), 123–134.
- [6] DEMBO, A. Moderate deviations for martingales with bounded jumps. *Electron. Comm. Probab.* 1, no. 3 (1996), 11–17.
- [7] DEMBO, A., AND ZEITOUNI, O. *Large deviations techniques and applications, second edition*, vol. 38 of *Applications of Mathematics*. Springer, 1998.
- [8] DJELLOUT, H. Moderate deviations for martingale differences and applications to  $\phi$ -mixing sequences. *Stoch. Stoch. Rep.* 73, 1-2 (2002), 37–63.
- [9] DJELLOUT, H., AND GUILLIN, A. Moderate deviations for Markov chains with atom. *Stochastic Process. Appl.* 95, no. 2 (2001), 203–217.
- [10] DURBIN, J. Testing for serial correlation in least-squares regression when some of the regressors are lagged dependent variables. *Econometrica* 38 (1970), 410–421.
- [11] DURBIN, J., AND WATSON, G. S. Testing for serial correlation in least squares regression. I. *Biometrika* 37 (1950), 409–428.
- [12] DURBIN, J., AND WATSON, G. S. Testing for serial correlation in least squares regression. II. *Biometrika* 38 (1951), 159–178.
- [13] DURBIN, J., AND WATSON, G. S. Testing for serial correlation in least squares regression. III. *Biometrika* 58 (1971), 1–19.
- [14] EICHELSBACHER, P., AND LÖWE, M. Moderate deviations for i.i.d. random variables. *ESAIM Probab. Stat.* 7 (2003), 209–218.
- [15] INDER, B. A. An approximation to the null distribution of the Durbin-Watson statistic in models containing lagged dependent variables. *Econometric Theory* 2 (1986), 413–428.
- [16] KING, M. L., AND WU, P. X. Small-disturbance asymptotics and the Durbin-Watson and related tests in the dynamic regression model. *J. Econometrics* 47 (1991), 145–152.
- [17] LEDOUX, M. Sur les déviations modérées des sommes de variables aléatoires vectorielles indépendantes de même loi. *Ann. Inst. Henri-Poincaré.* 35 (1992), 123–134.
- [18] MALINVAUD, E. Estimation et prévision dans les modèles économiques autorégressifs. *Review of the International Institute of Statistics* 29 (1961), 1–32.
- [19] NERLOVE, M., AND WALLIS, K. F. Use of the Durbin-Watson statistic in inappropriate situations. *Econometrica* 34 (1966), 235–238.
- [20] PROÏA, F. Further results on the H-Test of Durbin for stable autoregressive processes. *J. Multivariate. Anal. In Press.* (2013).
- [21] PUHALSKII, A. Large deviations of semimartingales: a maxingale problem approach. I. Limits as solutions to a maxingale problem. *Stoch. Stoch. Rep.* 61 (1997, no. 3-4), 141–243.
- [22] STOCKER, T. On the asymptotic bias of OLS in dynamic regression models with autocorrelated errors. *Statist. Papers* 48 (2007), 81–93.
- [23] WORMS, J. Moderate deviations for stable Markov chains and regression models. *Electron. J. Probab.* 4, no. 8 (1999), 1–28.
- [24] WORMS, J. Moderate deviations of some dependent variables. I. Martingales. *Math. Methods Statist.* 10, no. 1 (2001), 38–72.
- [25] WORMS, J. Moderate deviations of some dependent variables. II. Some kernel estimators. *Math. Methods Statist.* 10, no. 2 (2001), 161–193.



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

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

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

*E-mail address:* Frederic.Proia@inria.fr

UNIVERSITÉ BORDEAUX 1, INSTITUT DE MATHÉMATIQUES DE BORDEAUX, UMR 5251, AND INRIA  
BORDEAUX, TEAM ALEA, 200 AVENUE DE LA VIEILLE TOUR, 33405 TALENCE CEDEX, FRANCE.