# Moderate deviations of empirical periodogram and non-linear functionals of moving average processes 

H. Djellout ${ }^{\text {a,* }}$, A. Guillin ${ }^{\text {b }}$, L. Wu ${ }^{\text {a,c }}$<br>${ }^{\text {a }}$ Laboratoire de Mathématiques, CNRS-UMR 6620, Université Blaise Pascal, 63177 Aubière, France<br>${ }^{\text {b }}$ Ceremade, CNRS-UMR 7534, Université Paris IX Dauphine, 75775 Paris, France<br>${ }^{\text {c }}$ Department of Mathematics, Wuhan University, 430072, China

Received 28 April 2004; received in revised form 17 February 2005; accepted 8 April 2005
Available online 17 November 2005


#### Abstract

A moderate deviation principle for non-linear functionals, with at most quadratic growth, of moving average processes (or linear processes) is established. The main assumptions on the moving average process are a Logarithmic Sobolev Inequality for the driving random variables and the continuity, or some (weaker) integrability condition on the spectral density (covering some cases of long range dependence). We also obtain the moderate deviation estimate for the empirical periodogram, exhibiting an interesting new form of the rate function, i.e. with a correction term compared to the Gaussian rate functional. As statistical applications we provide the moderate deviation estimates of the least square and the Yule-Walker estimators of the parameter of a stationary autoregressive process and of the Neyman-Pearson likelihood ratio test in the Gaussian case.


© 2005 Elsevier SAS. All rights reserved.

## Résumé

Un principe de déviations modérées pour des fonctionnelles non linéaires, à croissances quadratiques, des processus de moyennes mobiles (ou processus linéaire) est établi. Les conditions imposées sur le processus de moyennes mobiles sont une inégalité de Sobolev Logarithmique sur les variables aléatoires d'innovation et la continuité, ou une condition (plus faible) d'intégrabilité sur la densité spectrale (couvrant certains cas de longue mémoire). On obtient aussi une estimation des déviations modérées pour le périodogramme empirique, faisant apparaître une nouvelle forme de la fonction de taux, avec un terme correctif comparé à la fonction de taux gaussienne. Comme applications statistiques, on donne des estimations de déviations modérées pour les estimateurs de Yule-Walker et des moindres carrés du paramètre de processus autoregressif stationnaire, ainsi que pour le test de Neyman-Pearson pour le rapport de vraisemblance dans le cadre gaussien.
© 2005 Elsevier SAS. All rights reserved.
MSC: 60F10; 60G10; 60G15
Keywords: Moderate deviations; Moving average processes; Logarithmic Sobolev inequalities; Toeplitz matrices

[^0]
## 1. Introduction

Consider the moving average process (or the linear process)

$$
\begin{equation*}
X_{n}:=\sum_{j=-\infty}^{+\infty} a_{j-n} \xi_{j}=\sum_{j=-\infty}^{+\infty} a_{j} \xi_{n+j}, \quad \forall n \in \mathbb{Z} \tag{1.1}
\end{equation*}
$$

where $\left(\xi_{n}\right)_{n \in \mathbb{Z}}$ is a sequence of $\mathbb{R}$-valued centered square integrable i.i.d.r.v., with common law $\mathcal{L}\left(\xi_{0}\right)=\mu$, and $\left(a_{n}\right)_{n \in \mathbb{Z}}$ be a sequence of real numbers such that

$$
\begin{equation*}
\sum_{n \in \mathbb{Z}}\left|a_{n}\right|^{2}<+\infty \tag{1.2}
\end{equation*}
$$

This last condition (1.2) is necessary and sufficient for the a.s. convergence or convergence in law of the series (1.1) (see [17, Chapter 2]). The sequence ( $X_{k}$ ) is strictly stationary with the spectral density given by

$$
f(\theta):=\operatorname{Var}\left(\xi_{0}\right)|g(\theta)|^{2}
$$

where

$$
\begin{equation*}
g(\theta):=\sum_{n=-\infty}^{+\infty} a_{n} \mathrm{e}^{\mathrm{i} n \theta} \tag{1.3}
\end{equation*}
$$

Moving average processes (or linear processes) are of special importance in time series analysis, filtrage of noise and they arise in a wide variety of contexts. Applications to economics, engineering and physical sciences are very broad and a vast amount of literature is devoted to the study of the limit theorems for moving average processes under various conditions (e.g. Brockwell and Davis [4] and references therein). A most important class of moving average processes is the real stationary Gaussian processes $\left(X_{n}\right)$ with a square integrable spectral density function $f$ (which can be represented as (1.1) with $\xi=\mathcal{N}(0,1)$ in law).

Let

$$
\begin{equation*}
\mathcal{I}_{n}(\theta):=\frac{1}{n}\left|\sum_{k=1}^{n} X_{k} \mathrm{e}^{\mathrm{i} k \theta}\right|^{2} \tag{1.4}
\end{equation*}
$$

be the so-called empirical periodogram of order $n$ of the process $\left(X_{k}\right)$. It is one of the main tools in the study of non-parametric statistical estimation of the unknown spectral density $f$ on the basis of the sample ( $X_{1}, \ldots, X_{n}$ ) from the process $\left(X_{n}\right)$. And for an observable $F(x)=F\left(x_{0}, \ldots, x_{l}\right)$ valued in $\mathbb{R}^{m}$, let

$$
\frac{1}{n} S_{n}(F):=\frac{1}{n} \sum_{k=1}^{n} F\left(X_{k}, X_{k+1}, \ldots, X_{k+l}\right)
$$

be the empirical mean of $F$. We begin with reviewing some known results which motivate our investigation.
(I) Linear observables $F(x)=x_{0}$.
(a) The minimal condition for the central limit theorem (CLT in short) for $\frac{1}{n} \sum_{k=1}^{n} X_{k}$ is the continuity of $g$ at $\theta=0$ (see [17, Corollary 5.2, p. 135]).
(b) Large deviations for $\frac{1}{n} \sum_{k=1}^{n} X_{k}$. See Burton and Dehling [7], Jiang, Rao and Wang [18,19], Djellout and Guillin [12] etc.
For non-linear observables $F$, the limit theorems for $\frac{1}{n} S_{n}(F)$ becomes much more difficult, even in the particular Gaussian case.
(II) Quadratic observables $F(x)=\left(x_{0}^{2}, x_{0} x_{1}, \ldots, x_{0} x_{l}\right)$ and $\mathcal{I}_{n}(\theta)$.

By Fourier series, one can often reduce the limit theorems of the empirical periodogram $\mathcal{I}_{n}(\theta)$ to those for $\frac{1}{n} S_{n}(F)$ where $F(x)=\left(x_{0}^{2}, x_{0} x_{1}, \ldots, x_{0} x_{l}\right)$. There exists an abundant literature on limit theorems of $\mathcal{I}_{n}(\theta)$ and of $\frac{1}{n} S_{n}(F)$ because of their importance in practice, especially in Gaussian case.
(a) CLT. Avram [1] and Fox and Taqqu [15] proved the CLT for $\mathcal{I}_{n}(\cdot)$ and $\frac{1}{n} S_{n}(F)$ in the Gaussian case. This CLT was generalized by Giraitis and Surgailis [16] to non-Gaussian case.
(b) Large and moderate deviations. Bryc and Dembo [6] have considered quadratic functional $F(x)=x_{0}^{2}$ of Gaussian processes both at the level of large and moderate deviations, under the boundedness of $f$ or the $L^{q}$-integrability of $f$ respectively. But for $F(x)=x_{0} x_{l}$ with $l \geqslant 1$, they have assumed that $f=1$ (i.e., $\left(X_{k}\right)$ are i.i.d.), an assumption excluding the dependent case.
Their result on large deviations (LDP in short) was generalized for general quadratic $F$ by Bercu, Gamboa and Rouault [2] under some condition on the distribution of the eigenvalues of the involved Toeplitz matrix, always in the Gaussian case. This last "technical" condition, (wrongly) omitted in the precedent works, is optimal but quite difficult to check in practice. In [2], they provided several concrete important statistical examples for which their condition is fulfilled.
In [24], the third author proved the LDP of $\mathcal{I}_{n}$ and $\frac{1}{n} S_{n}(F)$ for general quadratic $F$, without the technical condition in [2], but under the following integrability condition $\mathbb{E} \mathrm{e}^{\lambda \xi^{2}}<+\infty, \forall \lambda>0$, which excludes unfortunately the Gaussian case.
(III) General non-linear observables $F$.
(a) CLT. The literature is again abundant, we refer the reader to Rosenblatt [21] and the references therein.
(b) Large deviations. The seminal work of Donsker and Varadhan [14] established the LDP of the empirical process $R_{n}:=\frac{1}{n} \sum_{k=1}^{n} \delta_{\left(X_{k}, X_{k+1}, \ldots\right)}$ for the stationary Gaussian processes such that $f \in C_{b}(\mathbb{T})$ and $\log f \in$ $L^{1}(\mathbb{T})$. This implies the LDP of $\frac{1}{n} S_{n}(F)$ once if $F$ is continuous and bounded. Bryc and Dembo [5] showed that the continuity of the spectral density $f$ cannot be weakened but the condition $\log f \in L^{1}(\mathbb{T})$ can be removed, for the LDP result of Donsker and Varadhan. More recently, the third author [24] generalized this last result to all moving average processes such that $\mathbb{E} \mathrm{e}^{\delta \xi^{2}}<+\infty$ for some $\delta>0$.

The main purpose of this paper consists to investigate the moderate deviation principle (MDP in short) for the so-called empirical periodogram $\mathcal{I}_{n}(\theta)$ of order $n$ of the process $\left(X_{k}\right)$ defined by (1.4) in the space $L^{p}(\mathbb{T}, \mathrm{~d} \theta)$ of $p$-integrable function on the torus $\mathbb{T}$ identified with $[-\pi, \pi[$ equipped with the weak convergence topology. We establish the MDP for $\mathcal{I}_{n}(\theta)$ under some conditions such as the $L^{q}(\mathbb{T}, d \theta)$-integrability of the spectral density of $\left(X_{k}\right)$ and a Logarithmic Sobolev Inequality (in short LSI) for the law $\mu$ of the driven random variable $\xi$. Moreover our approach allows us to obtain the MDP of $\frac{1}{n} S_{n}(F)$ for non-linear $\mathbb{R}^{m}$-valued observables $F$ of at most quadratic growth.

To our knowledge, it is the first time that a MDP for a general class of non-linear observables of moving average processes is established (not only in the Gaussian case). Our investigation is a natural continuation of the known works [14,6,2,24] etc. We also consider statistical applications such as
(1) the MDP of the least square and Yule-Walker estimators of the autoregression parameter in a stationary autoregressive process, complementing known CLT results and the LDP (limited to the Gaussian case) obtained by Bercu et al. [2];
(2) the MDP in the Neyman-Pearson likelihood ratio test (largely inspired by Bercu and al. [2]) in the Gaussian case.

Besides the standard techniques in large deviations (such as approximation lemmas, projective limit etc.), our method is mainly based on the LSI technique, as developed by Ledoux [20] and al.

This paper is structured as follows. The MDP for the empirical spectral density and non-linear functionals are stated in next section. In Section 3, we provide statistical applications. We establish the key a priori estimations in Section 4. The last section is devoted to the proofs of the main results.

## 2. Main results

### 2.1. MDP for the empirical periodogram

For the sake of completeness, we recall the definition of the LDP [10] and [11]. A sequence of random variables $\left(Y_{n}\right)$ with values in a regular Hausdorff topological space $E$ is said to satisfy the LDP with speed $\lambda_{n} \rightarrow \infty$ and good rate function $I(\cdot): E \rightarrow \mathbb{R}^{+}$if: $I$ has compact level sets and for all measurable sets $A$ of $X$ :

$$
-\inf _{x \in \AA} I(x) \leqslant \liminf _{n \rightarrow+\infty} \frac{1}{\lambda_{n}} \log \mathbb{P}\left(Y_{n} \in A\right) \leqslant \limsup _{n \rightarrow+\infty} \frac{1}{\lambda_{n}} \log \mathbb{P}\left(Y_{n} \in A\right) \leqslant-\inf _{x \in \bar{A}} I(x)
$$

where $\AA, \bar{A}$ denote the interior and closure of $A$, respectively.
In the whole paper we shall study only a special type of LDP, called usually moderate deviation principle (MDP in short, cf. [10]).

Let $\left(\xi_{n}\right)_{n \in \mathbb{Z}}$ is a sequence of $\mathbb{R}$-valued centered i.i.d.r.v., with common law $\mathcal{L}\left(\xi_{0}\right)=\mu$, and let $a:=\left(a_{n}\right)_{n \in \mathbb{Z}}$ be a sequence of real numbers satisfying (1.2), and define ( $X_{n}$ ) by (1.1). Our basic assumption, supposed throughout this paper, is that $\mu$ satisfies a LSI, i.e. there exists $C>0$ such that

$$
\begin{equation*}
\operatorname{Ent}_{\mu}\left(h^{2}\right) \leqslant 2 C \mathbb{E}_{\mu}\left(|\nabla h|^{2}\right) \tag{2.1}
\end{equation*}
$$

for every smooth $h$ such that $\mathbb{E}_{\mu}\left(h^{2} \log ^{+} h^{2}\right)<\infty$, where

$$
\operatorname{Ent}_{\mu}\left(h^{2}\right)=\mathbb{E}_{\mu}\left(h^{2} \log h^{2}\right)-\mathbb{E}_{\mu}\left(h^{2}\right) \log \mathbb{E}_{\mu}\left(h^{2}\right)
$$

See Ledoux [20] for further details on LSI. Note that it implies in particular that there exists some positive $\delta$ such that

$$
\begin{equation*}
\mathbb{E}_{\mu}\left(\mathrm{e}^{\delta|x|^{2}}\right)<\infty \tag{2.2}
\end{equation*}
$$

Remark 2.1. First note that there exists some practical criteria ensuring the LSI. For example, consider a $C^{2}$ function $W$ on $\mathbb{R}^{d}$ such that $\mathrm{e}^{-W}$ is integrable with respect to Lebesgue measure and let

$$
\mathrm{d} \mu(x)=Z^{-1} \mathrm{e}^{-W(x)} \mathrm{d} x
$$

where $Z$ is the normalization constant, and suppose that for some $c \in \mathbb{R}, W^{\prime \prime}(x) \geqslant c I$ for every $x$ and that for some $\epsilon>0$,

$$
\iint \mathrm{e}^{\left(c^{-}+\epsilon\right)|x-y|^{2}} \mathrm{~d} \mu(x) \mathrm{d} \mu(y)<\infty
$$

where $c^{-}=-\min (c, 0)$. Then $\mu$ satisfies (2.1) by the criterion of Wang [20]. Obviously Gaussian variables fulfill this criterion. See Bobkov and Götze [3] for a necessary and sufficient condition in the actual one-dimensional case, relying on generalized Hardy's inequalities.

We are interested in the moderate deviation principle (MDP in short) of the empirical spectral density (or periodogram) of ( $X_{n}$ ) defined by

$$
\mathcal{I}_{n}(\theta):=\frac{1}{n}\left|\sum_{k=1}^{n} X_{k} \mathrm{e}^{\mathrm{i} k \theta}\right|^{2}
$$

which are random elements in the space $L^{p}(\mathbb{T}, \mathrm{~d} \theta)$ equipped with the weak convergence topology, where $\mathbb{T}$ is the torus identified with $[-\pi, \pi[$ in the usual way.

We first present here the MDP for the empirical autocorrelation vector which will be our main tool for the MDP of the empirical spectral density, and has its own interest in statistics. Let

$$
\kappa_{4}=\frac{\mathbb{E}\left(\xi^{4}\right)-3\left[\mathbb{E}\left(\xi^{2}\right)\right]^{2}}{\mathbb{E}\left(\xi^{2}\right)^{2}},
$$

the cumulant of order 4 of the driven random variable $\xi$.
Theorem 2.1. Assume that $\mu$ satisfies the LSI (2.1). Suppose moreover that
(H1) the spectral density function $f$ is in $L^{q}(\mathbb{T}, \mathrm{~d} \theta)$, where $2<q \leqslant+\infty$; and
(H2) the moderate deviation scale ( $b_{n}$ ) is a sequence of positive numbers satisfying $1 \ll b_{n} \ll \sqrt{n}$ (i.e. $b_{n} \rightarrow+\infty$ and $b_{n} n^{-1 / 2} \rightarrow 0$, the moderate deviation scale) and

$$
b_{n} n^{1 / q-1 / 2} \rightarrow 0
$$

then for every $\lambda \in \mathbb{R}^{m+1}$,

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \frac{1}{b_{n}^{2}} \log \mathbb{E} \exp \left(\frac{b_{n}}{\sqrt{n}} \sum_{\ell=0}^{m} \lambda_{l} \sum_{k=1}^{n}\left(X_{k} X_{k+\ell}-\mathbb{E} X_{k} X_{k+\ell}\right)\right)=\frac{1}{2} \lambda^{*} \Sigma^{2} \lambda \tag{2.3}
\end{equation*}
$$

where $\Sigma^{2}=\left(\Sigma_{k, \ell}^{2}\right)_{0 \leqslant k, \ell \leqslant m}$ is given by

$$
\begin{align*}
\Sigma_{k, \ell}^{2} & =\frac{1}{2 \pi} \int_{\mathbb{T}}\left(\mathrm{e}^{\mathrm{i}(k-\ell) \theta}+\mathrm{e}^{\mathrm{i}(k+\ell) \theta}\right) f^{2}(\theta) \mathrm{d} \theta+\kappa_{4}\left(\frac{1}{2 \pi} \int_{\mathbb{T}} f(\theta) \mathrm{e}^{\mathrm{i} k \theta} \mathrm{~d} \theta\right)\left(\frac{1}{2 \pi} \int_{\mathbb{T}} f(\theta) \mathrm{e}^{\mathrm{i} \ell \theta} \mathrm{~d} \theta\right) \\
& =\frac{1}{2 \pi} \int_{\mathbb{T}} 2 \cos (k \theta) \cos (\ell \theta) f^{2}(\theta) \mathrm{d} \theta+\kappa_{4}\left(\frac{1}{2 \pi} \int_{\mathbb{T}} f(\theta) \cos (k \theta) \mathrm{d} \theta\right)\left(\frac{1}{2 \pi} \int_{\mathbb{T}} f(\theta) \cos (l \theta) \mathrm{d} \theta\right) . \tag{2.4}
\end{align*}
$$

In particular

$$
\left(\frac{1}{b_{n} \sqrt{n}} \sum_{k=1}^{n}\left(X_{k} X_{k+\ell}-\mathbb{E} X_{k} X_{k+\ell}\right)\right)_{0 \leqslant \ell \leqslant m}
$$

satisfies the LDP on $\mathbb{R}^{m+1}$ with speed $b_{n}^{2}$ and with the rate function given by

$$
I(z)=\sup _{\lambda \in \mathbb{R}^{m+1}}\left\{\langle\lambda, z\rangle-\frac{1}{2}\left\langle\lambda, \Sigma^{2} \lambda\right\rangle\right\} .
$$

Remark 2.2. By Cauchy-Schwartz inequality we have $\left[\mathbb{E}\left(\xi^{2}\right)\right]^{2} \leqslant \mathbb{E}\left(\xi^{4}\right)$, so $\kappa_{4} \geqslant-2$ and $\kappa_{4}=-2$ iff $\xi^{2}=C$, a.s. Under the assumption (2.1), $\xi^{2}$ cannot be constant by [13, Remark 2.4], so $\kappa_{4}>-2$. Consequently the matrix $\Sigma^{2}$ is symmetric and non-negative definite. Notice that the rate function $I$ given above can be calculated explicitly as

$$
I(z)= \begin{cases}\frac{1}{2}\left\langle z, \Sigma^{-2} z\right\rangle, & \text { if } z \in \operatorname{Ran}\left(\Sigma^{2}\right) ; \\ +\infty, & \text { otherwise, }\end{cases}
$$

where $\Sigma^{-2}$ is the inverse of $\Sigma^{2}$ restricted to the range $\operatorname{Ran}\left(\Sigma^{2}\right)$ of $\Sigma^{2}$.
Remark 2.3. The assumptions (H1) and (H2) on $f$ and the scale $b_{n}$ are exactly the ones imposed in Bryc and Dembo [6, Theorem 2.3] for the MDP of $\frac{1}{n} \sum_{k=1}^{n} X_{k}^{2}$ in the Gaussian case. Their large deviations result (namely Proposition 2.5 in [6]) for the empirical autocorrelation is further restricted to the i.i.d. case.

Remark 2.4. Notice that the condition (H1) on the dependence is indeed quite weak and general. It covers not only the short-range case (i.e. $\sum\left|\operatorname{Cov}\left(X_{0}, X_{n}\right)\right|<+\infty$ ), but also some cases of long range. To illustrate this case, consider the following example: let $\left\{B_{H}(t), t \in \mathbb{R}\right\}$ be the fractional Brownian motion with Hurst parameter $0<H<1$. Consider its increments

$$
Y_{j}=B_{H}(j+1)-B_{H}(j), \quad j \in \mathbb{Z},
$$

which form a stationary Gaussian sequence with mean zero and variance $\mathbb{E}\left(B_{H}^{2}(1)\right)=\sigma_{0}^{2}$. The sequence $\left\{Y_{j}, j \in \mathbb{Z}\right\}$ has the covariance function

$$
\alpha(j)=\mathbb{E}\left(Y_{1} Y_{j+1}\right)=\frac{\sigma_{0}^{2}}{2}\left(|j+1|^{2 H}-2|j|^{2 H}+|j-1|^{2 H}\right),
$$

and the spectral density

$$
f(\lambda)=\frac{\sigma_{0}^{2}}{C^{2}}\left|\mathrm{e}^{\mathrm{i} \lambda}-1\right|^{2} \sum_{k=-\infty}^{+\infty} \frac{1}{|\lambda+2 \pi k|^{2 H+1}}, \quad-\pi \leqslant \lambda \leqslant \pi,
$$

where $C$ is a constant depending only on $H$. It is known that (see [22])

$$
\alpha(j) \sim \sigma_{0}^{2} H(2 H-1) j^{2 H-2}, \quad \text { as } j \rightarrow \infty, \text { for } H \neq 1 / 2
$$

and $f$ is continuous on $\mathbb{T} \backslash\{0\}$ and

$$
f(\lambda) \sim \sigma_{0}^{2} C^{-2}(H)|\lambda|^{1-2 H}, \quad \text { as } \lambda \rightarrow 0
$$

When $0<H \leqslant 1 / 2, f$ is continuous (and then bounded) on $\mathbb{T}$. So the MDP in Theorem 2.1 holds for every moderate deviation scale $\left(b_{n}\right)$.

When $1 / 2<H<1$, the series $\sum \alpha(j)$ diverges. In this case $\left\{Y_{j}, j \in \mathbb{Z}\right\}$ exhibits long range dependence. The condition (H1) is satisfied if $1 / 2<H<1 / 2+1 /(2 q)$, (so $1 / 2<H<3 / 4$ for $q>2$ ), and (1.2) is thus easily verified. In this case, we obtain the MDP of Theorem 2.1 for the sequence $\left\{Y_{j}, j \in \mathbb{Z}\right\}$ with $\kappa_{4}=0$ for the moderate deviation scale $\left(b_{n}\right)$ verifying (H2).

The following corollary follows from Theorem 2.1 by the contraction principle
Corollary 2.2. Under the assumptions of Theorem 2.1, we have for all $\ell \geqslant 0,\left(1 /\left(\sqrt{n} b_{n}\right) \sum_{k=1}^{n}\left(X_{k} X_{k+\ell}-\mathbb{E} X_{k} X_{k+\ell}\right)\right)$ satisfies the LDP on $\mathbb{R}$ with speed $b_{n}^{2}$ and rate function given by

$$
I^{\ell}(z)=\frac{1}{2} \frac{z^{2}}{\left.1 /(2 \pi) \int_{\mathbb{T}} 2 \cos ^{2}(\ell \theta)\right) f^{2}(\theta) \mathrm{d} \theta+\kappa_{4}\left(1 /(2 \pi) \int_{\mathbb{T}} f(\theta) \cos (\ell \theta) \mathrm{d} \theta\right)^{2}}
$$

with the convention that $a / 0=+\infty$ for $a>0$ and $0 / 0:=0$.
Let us present now the main result of this paper. From Theorem 2.1 (and its proof) together with the projective limit method, we yield the functional type's MDP below, for

$$
\mathcal{L}_{n}(\theta)=\frac{\sqrt{n}}{b_{n}}\left(\mathcal{I}_{n}(\theta)-\mathbb{E} \mathcal{I}_{n}(\theta)\right)
$$

Theorem 2.3. Suppose that $\mu$ satisfies the LSI (2.1) and (H1), (H2). Let $1 \leqslant p<2$ and $p^{\prime} \in[2,+\infty]$ the conjugated number, i.e., $1 / p+1 / p^{\prime}=1$. Assume moreover
(H3) the moderate deviation scale $b_{n}$ satisfies

$$
b_{n} n^{1 / q+1 / p^{\prime}-1 / 2} \rightarrow 0, \quad \frac{1}{p^{\prime}}+\frac{1}{q}<\frac{1}{2} .
$$

Then $\mathcal{I}_{n}(\theta)$ satisfies the MDP, i.e., $\left(\mathcal{L}_{n}\right)_{n} \geqslant 0$ satisfies the $L D P$ on $\left(L^{p}(\mathbb{T}, \mathrm{~d} \theta), \sigma\left(L^{p}(\mathbb{T}, \mathrm{~d} \theta), L^{p^{\prime}}(\mathbb{T}, \mathrm{d} \theta)\right)\right)$ with speed $b_{n}^{2}$ and with the rate function given by

$$
J(\eta)=\left\{\begin{array}{l}
\frac{1}{2 \pi} \int_{\mathbb{T}} \frac{\eta^{2}(\theta)}{4 f^{2}(\theta)} \mathrm{d} \theta-\frac{\kappa_{4}}{2+\kappa_{4}}\left(\frac{1}{2 \pi} \int_{\mathbb{T}} \frac{\eta(\theta)}{2 f(\theta)} \mathrm{d} \theta\right)^{2}, \\
\quad \text { if } \kappa_{4}>-2, \eta \text { is even, } \eta \mathrm{d} \theta \ll f \mathrm{~d} \theta \text { and } \frac{\eta}{f} \in L^{2}(\mathbb{T}, \mathrm{~d} \theta) ; \\
+\infty, \text { otherwise. }
\end{array}\right.
$$

As a consequence of Theorem 2.3 we have the following marginal MDP:
Corollary 2.4. Under the assumptions of Theorem 2.3, we have that for all $h \in L^{p^{\prime}}(\mathbb{T}, \mathrm{d} \theta)$,

$$
\limsup _{n \rightarrow \infty} \frac{1}{b_{n}^{2}} \log \mathbb{E}\left(\mathrm{e}^{b_{n}^{2} \frac{1}{2 \pi} \int_{\mathbb{T}} h(\theta) \mathcal{L}_{n}(\theta) \mathrm{d} \theta}\right)=\frac{1}{2} \sigma^{2}(h)
$$

where

$$
\sigma^{2}(h):=\frac{1}{2 \pi} \int_{\mathbb{T}} 2 \tilde{h}^{2}(\theta) f^{2}(\theta) \mathrm{d} \theta+\kappa_{4}\left(\frac{1}{2 \pi} \int_{\mathbb{T}} h(\theta) f(\theta) \mathrm{d} \theta\right)^{2}
$$

and $\tilde{h}(\theta)=(h(\theta)+h(-\theta)) / 2$. In particular $\frac{1}{2 \pi} \int_{\mathbb{T}} h(\theta) \mathcal{L}_{n}(\theta) \mathrm{d} \theta$ satisfies the LDP on $\mathbb{R}$ with speed $b_{n}^{2}$ and with the rate function given by $I_{h}(z):=(1 / 2) z^{2} /\left(\sigma^{2}(h)\right)$.

Remark 2.5. One cannot hope that the MDP in Theorem 2.3 holds w.r.t. the strong topology of $L^{p}(\mathbb{T}, \mathrm{~d} \theta)$, because the rate function $I(\eta)$ is not inf-compact w.r.t. this topology.

The assumption (H3) is stronger than (H2). When $p=1$ (and then $p^{\prime}=+\infty$ ), (H3) becomes (H2) and thus under the LSI for $\xi$ and $(\mathrm{H} 1)$ and $(\mathrm{H} 2), \mathcal{L}_{n}(\theta)$ satisfies the LDP on $L^{1}(\mathbb{T})$ w.r.t. the weak convergence topology $\sigma\left(L^{1}, L^{\infty}\right)$ in Theorem 2.3, and $\frac{1}{2 \pi} \int_{\mathbb{T}} h(\theta) \mathcal{L}_{n}(\theta) \mathrm{d} \theta$ satisfies the LDP in Corollary 2.4 for every $h \in L^{\infty}(\mathbb{T})$.

Remark 2.6. Now assume that $\left(\xi_{n}\right)$ is a sequence of real i.i.d. normal random variables, so $\left(X_{n}\right)$ is a stationary Gaussian process and inversely any real Gaussian stationary process ( $X_{n}$ ) with a square integrable spectral density function $f$ can be represented as (1.1). In this case, we have $\mathbb{E}\left(\xi^{4}\right)=3 \mathbb{E}\left(\xi^{2}\right)^{2}$ and thus $\kappa_{4}=0$, so under the assumptions of Theorem 2.3 we obtain that $\left(\mathcal{L}_{n}\right)_{n} \geqslant 0$ satisfies the LDP on $L^{p}(\mathbb{T}, d \theta)$ with speed $b_{n}^{2}$ and with the rate function given by

$$
J(\eta)= \begin{cases}\frac{1}{2 \pi} \int_{\mathbb{T}} \frac{\eta^{2}(\theta)}{4 f^{2}(\theta)} \mathrm{d} \theta, & \text { if } \eta \text { is even, } \eta \mathrm{d} \theta \ll f \mathrm{~d} \theta \text { and } \frac{\eta}{f} \in L^{2}(\mathbb{T}, \mathrm{~d} \theta), \\ +\infty, & \text { otherwise. }\end{cases}
$$

We thus give the MDP for the spectral empirical measure in the setting of Bercu and al. [2]. Note however that they only consider the marginal LDP, i.e. LDP for $\mathcal{I}_{n}(h)$ for some bounded $h$ on the torus with an extra assumption on the eigenvalues of the Toeplitz matrix, where $\mathcal{I}_{n}(h)=\frac{1}{2 \pi} \int_{\mathbb{T}} \mathcal{I}_{n}(\theta) h(\theta) \mathrm{d} \theta$.

Remark 2.7. For any real and symmetric function $h \in L^{1}(\mathbb{T}, \mathrm{~d} \theta)$, let $T_{n}(h)$ be the Toeplitz matrix of order $n$ associated with $h$ i.e. $T_{n}(h)=\left(\hat{r}_{k-l}(h)\right)_{1 \leqslant k, l \leqslant n}$ where $\hat{r}_{k}(h)$ is the $k$ th Fourier coefficient of $h$ given by

$$
\begin{equation*}
\hat{r}_{k}(h)=\frac{1}{2 \pi} \int_{\mathbb{T}} \mathrm{e}^{\mathrm{i} k \theta} h(\theta) \mathrm{d} \theta, \quad \forall k \in \mathbb{Z} \tag{2.5}
\end{equation*}
$$

The matrix $T_{n}(h)$ is obviously real and symmetric, is positive definite whenever $h \geqslant 0$.
Notice that the extra term with respect to the Gaussian case in the evaluation of the asymptotic variance has been known for a long time (see [21]). The result of [16] about CLT for $\mathcal{I}_{n}$ can be summarized as below: if

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \frac{1}{n} \operatorname{tr}\left(\left(T_{n}(f) T_{n}(h)\right)^{2}\right)=\frac{1}{2 \pi} \int_{\mathbb{T}} f^{2}(\theta) h^{2}(\theta) \mathrm{d} \theta \tag{2.6}
\end{equation*}
$$

(where $T_{n}(h)$ is the Toeplitz matrix of $h$ ) then $\sqrt{n}\left(\mathcal{I}_{n}(h)-\mathbb{E} \mathcal{I}_{n}(h)\right.$ ) converges in law (as $\left.n \rightarrow \infty\right)$ to the normal distribution $\mathcal{N}\left(0, \sigma^{2}(h)\right)$ with $\sigma^{2}(h)$ given in Corollary 2.4. In Gaussian case this result was already proved by Avram [1] and Fox and Taqqu [15].

In the next corollaries of Theorem 2.3, we replace $\mathbb{E} \mathcal{I}_{n}(\theta)$ by $f(\theta)$ in the definition of $\mathcal{L}_{n}(\theta)$, more useful in practice, but need more assumptions. More precisely we are interested in the MDP of

$$
\tilde{\mathcal{L}}_{n}(\theta)=\frac{\sqrt{n}}{b_{n}}\left(\mathcal{I}_{n}(\theta)-f(\theta)\right) .
$$

Corollary 2.5. Suppose that $\mu$ satisfies the $\operatorname{LSI}$ (2.1) and the spectral density $f$ verifies

$$
\begin{equation*}
f \in L^{\infty}(\mathbb{T}) \quad \text { and } \quad\|f(t+\cdot)-f(\cdot)\|_{L^{p}(\mathbb{T})}=\mathrm{O}(\sqrt{t}) \tag{2.7}
\end{equation*}
$$

then for every scale $1 \ll b_{n} \ll n^{1 / 2-1 / p^{\prime}}$, $\left(\tilde{\mathcal{L}}_{n}\right)_{n \geqslant 0}$ satisfies the LDP on $L^{p}(\mathbb{T}, \mathrm{~d} \theta)$ w.r.t. the weak topology $\left.\sigma\left(L^{p}(\mathbb{T}), L^{p^{\prime}}(\mathbb{T})\right)\right)$, with speed $b_{n}^{2}$ and with the rate function $J$ given in Theorem 2.3.

We have also the following consequence of Corollary 2.4 for the marginals of the empirical spectral measures
Corollary 2.6. Assume (2.1) and (H1), (H2). Suppose that

$$
\begin{equation*}
h \in L^{\infty}(\mathbb{T}) \quad \text { and } \quad\|h(t+\cdot)-h(\cdot)\|_{L^{q^{\prime}}(\mathbb{T})}=\mathrm{O}(\sqrt{t}) \tag{2.8}
\end{equation*}
$$

then the conclusion of Corollary 2.4 holds for $\int_{\pi} h(\theta) \tilde{\mathcal{L}}_{n}(\theta) \mathrm{d} \theta$ instead of $\int_{\pi} h(\theta) \mathcal{L}_{n}(\theta) \mathrm{d} \theta$.

### 2.2. MDP for non-linear functionals

We now present the MDP of $\frac{1}{n} S_{n}(F)$, i.e., the LDP of

$$
M_{n}(F)=\frac{1}{\sqrt{n} b_{n}} \sum_{k=1}^{n}\left(F\left(X_{k}, \ldots, X_{k+l}\right)-\mathbb{E}\left(F\left(X_{k}, \ldots, X_{k+l}\right)\right)\right),
$$

where the observable $F: \mathbb{R}^{l+1} \rightarrow \mathbb{R}^{m}$ is a general non-linear differentiable function.
Theorem 2.7. Suppose that $\mu$ satisfies the LSI (2.1), and $g$ given in (1.3) is continuous on $\mathbb{T}$. Assume moreover that $\partial_{x_{i}} F$ is Lipschitz for $i=0, \ldots, l$. Then

$$
\begin{equation*}
\Sigma_{F}^{2}:=\lim _{n \rightarrow+\infty} \frac{1}{n} \Gamma\left(\sum_{k=1}^{n} F\left(X_{k}, \ldots, X_{k+l}\right)\right) \tag{2.9}
\end{equation*}
$$

exists where $\Gamma(\cdot)$ is the covariance matrix of the random vector $\cdot$, and for every moderate deviation scale $1 \ll b_{n} \ll$ $\sqrt{n}, M_{n}(F)$ satisfies the LDP on $\mathbb{R}^{m}$ with speed $b_{n}^{2}$ and good rate function $I_{F}$ given by

$$
I_{F}(z)=\sup _{\lambda \in \mathbb{R}^{m}}\left\{\langle\lambda, z\rangle-\frac{1}{2}\left\langle\lambda, \Sigma_{F}^{2} \lambda\right\rangle\right\}= \begin{cases}\frac{1}{2}\left\langle z, \Sigma_{F}^{-2} z\right\rangle, & \text { if } z \in \operatorname{Ran}\left(\Sigma_{F}^{2}\right), \\ +\infty, & \text { otherwise, }\end{cases}
$$

where $\Sigma_{F}^{-2}: \operatorname{Ran}\left(\Sigma_{F}^{2}\right) \rightarrow \operatorname{Ran}\left(\Sigma_{F}^{2}\right)$ is the inverse of the limit covariance matrix $\Sigma_{F}^{2}$ restricted to $\operatorname{Ran}\left(\Sigma_{F}^{2}\right)$.
Note also the following corollary in the linear case $F\left(x_{0}, \ldots, x_{l}\right)=x_{0}$ in which the assumption on $g$ can be largely weakened.

Corollary 2.8. Suppose that $\mu$ satisfies the integrability condition (2.2), if

$$
f^{N}(0)=\sum_{|k| \leqslant N}\left(1-\frac{|k|}{N}\right) \hat{r}_{k}(f) \rightarrow \sigma^{2},
$$

then for every moderate deviation scale $1 \ll b_{n} \ll \sqrt{n}, \frac{1}{b_{n} \sqrt{n}} \sum_{k=1}^{n} X_{k}$ satisfies the LDP on $\mathbb{R}$ with speed $b_{n}^{2}$ and rate $I(z)=\frac{1}{2} \frac{z^{2}}{\sigma^{2}}$.

Remark 2.8. When $f$ admits a version which is continuous at 0 , then $f^{N}(0) \rightarrow f(0)=\sigma^{2}$. This corollary generalizes Theorem 3.1 of Djellout and Guillin [12] to the case of unbounded r.v.

## 3. Statistical applications

We now provide two statistical applications. The first deals with the least square estimator of the parameter of the autoregressive linear process and the second about the likelihood ratio test on spectral densities in the Gaussian case.

### 3.1. Autoregressive stationary process

Consider the autoregressive process (not necessarily Gaussian)

$$
X_{n+1}=\theta X_{n}+\sigma \xi_{n+1},
$$

where the noises sequence $\left(\left(\xi_{n}\right)_{n \in \mathbb{Z}}\right)$ is i.i.d. with common law $\mu$, satisfying a LSI, and $\mathbb{E}\left(\xi_{n}\right)=0, \mathbb{E}\left(\xi^{2}\right)=1, \sigma>0$ and $\theta \in(-1,1)$ is the unknown parameter. Assume that $X_{0}$ is independent of $\left(\xi_{n}\right)_{n} \geqslant 1$ and has the same law as $\sum_{k=0}^{\infty} \theta^{k} \sigma \xi_{-k} .\left(X_{n}\right)$ is thus a centered stationary process of the form (1.1), with spectral density given by

$$
f(t)=\frac{\sigma^{2}}{1+\theta^{2}-2 \theta \cos t}, \quad \forall t \in \mathbb{T}
$$

Let $\hat{\theta}_{n}$ be the least square estimator of $\theta$, given by:

$$
\hat{\theta}_{n}=\frac{\sum_{i=1}^{n} X_{i} X_{i-1}}{\sum_{i=1}^{n} X_{i-1}^{2}}
$$

It is well-known that $\hat{\theta}_{n} \rightarrow \theta$ a.s. and $\sqrt{n}\left(\hat{\theta}_{n}-\theta\right)$ satisfies the CLT. We show in the next proposition the MDP of the least square estimator.

Proposition 3.1. For every moderate deviation scale $1 \ll b_{n} \ll \sqrt{n}, \frac{\sqrt{n}}{b_{n}}\left(\hat{\theta}_{n}-\theta\right)$ satisfies a LDP on $\mathbb{R}$ with speed $b_{n}^{2}$ and with the rate function given by

$$
I(x)=\frac{x^{2}}{2\left(1-\theta^{2}\right)}
$$

Remark 3.1. Let $\tilde{\theta}_{n}$ be the Yule-Walker estimator of $\theta$ :

$$
\tilde{\theta}_{n}=\frac{\sum_{i=1}^{n} X_{i} X_{i-1}}{\sum_{i=0}^{n} X_{i}^{2}}
$$

It is well-known that the Yule-Walker estimator share the same almost sure property and the same CLT. Bercu et al. [2] showed however that the LDP of the Yule-Walker estimator is better than the one of the least-squares.

In the regime of the MDP, following the same proof as for the least square estimator we see that the Yule-Walker estimator share the same MDP.

### 3.2. Likelihood ratio test in the Gaussian case

Let $f_{0}$ and $f_{1}$ be two spectral densities which differ on a positive Lebesgue measure subset of $\mathbb{T}$. If we wish to test $\mathrm{H}_{0}: f=f_{0}$ against $\mathrm{H}_{1}: f=f_{1}$, on the basis of the stationary centered Gaussian observation $X_{1}, \ldots, X_{n}$, the Neyman-Pearson theorem tells us that the optimal strategy is the likelihood ratio test:

$$
L_{n}=\frac{1}{2 n}\left(\log \frac{\operatorname{det} T_{n}\left(f_{0}\right)}{\operatorname{det} T_{n}\left(f_{1}\right)}+\left\langle X^{(n)},\left[T_{n}\left(f_{0}\right)^{-1}-T_{n}\left(f_{1}\right)^{-1}\right] X^{(n)}\right\rangle\right)
$$

The study of the MDP properties of $\left(L_{n}\right)$ under hypothesis $\mathrm{H}_{0}$ or $\mathrm{H}_{1}$ is useful to control asymptotically the threshold or the power of the test. We now make the two following assumptions:
$\left(\mathrm{A}_{1}\right)$ the spectral density $f_{0}$ is in the Szegö class, i.e. $\log \left(f_{0}\right) \in L^{1}(\mathbb{T})$;
$\left(\mathrm{A}_{2}\right)$ the ratio $f_{0} / f_{1} \in L^{\infty}(\mathbb{T})$.
Under those assumption, Bercu and al. [2] proved that $L_{n}$ converges a.s. to

$$
\frac{1}{4 \pi}\left(\int_{\mathbb{T}} \log f_{0}(t) \mathrm{d} t-\int_{\mathbb{T}} \log f_{1}(t) \mathrm{d} t+\int_{\mathbb{T}}\left(1-\frac{f_{0}(t)}{f_{1}(t)}\right) \mathrm{d} t\right)
$$

and satisfies the LDP. Inspired by their work we have furthermore
Proposition 3.2. Assume that $\left(\mathrm{A}_{1}\right)$ and $\left(\mathrm{A}_{2}\right)$ are satisfied. Then, under the null hypothesis $\mathrm{H}_{0}$, for every moderate deviation scale $1 \ll b_{n} \ll \sqrt{n}$, the sequence $\frac{\sqrt{n}}{b_{n}}\left(L_{n}-\mathbb{E}\left(L_{n}\right)\right)$ satisfies a LDP on $\mathbb{R}$ with speed $b_{n}^{2}$ and good rate function

$$
G(x)=\frac{x^{2}}{(1 /(2 \pi)) \int_{\mathbb{T}}\left(1-f_{0} / f_{1}\right)^{2}(\theta) \mathrm{d} \theta}
$$

## 4. Several lemmas

In this section we first establish the a priori estimate, next recall several facts concerning the Toeplitz matrix and the Fejèr approximation and the MDP of $m$-dependent stationary sequences.

### 4.1. A priori estimation

We recall the following well known elementary result
Lemma 4.1. Suppose that $Y^{(n)}=\left(Y_{1}, \ldots, Y_{n}\right)^{*}$ is a standard $\mathcal{N}(0, I)$ centered Gaussian vector valued in $\mathbb{R}^{n}$ and let A be a symmetric real valued $n \times n$-matrix. Let $\lambda_{1}, \ldots, \lambda_{n}$ be the eigenvalues of the matrix $A$. Then for every $z \in \mathbb{R}$

$$
\log \mathbb{E} \exp \left(z\left\langle Y^{(n)}, A Y^{(n)}\right\rangle\right)= \begin{cases}-\frac{1}{2} \sum_{j=1}^{n} \log \left(1-2 z \lambda_{j}\right) & \text { if } \max _{1 \leqslant j \leqslant n}\left(z \lambda_{j}\right)<\frac{1}{2}  \tag{4.1}\\ +\infty, & \text { otherwise. }\end{cases}
$$

We give a crucial lemma which was first proved in Wu [24], and reproduced here for completeness.

Lemma 4.2. If the centered r.v. $\xi_{0}$ satisfies (2.2), then there is some constant $K>0$ such that

$$
\begin{equation*}
L(y):=\mathbb{E} \exp \left(y \xi_{0}\right) \leqslant \exp \left(\frac{K^{2}}{2} y^{2}\right), \quad \forall y \in \mathbb{R} \tag{4.2}
\end{equation*}
$$

Proof. Let $\delta>0$ be given in (2.2). Since

$$
2 y \xi_{0} \leqslant 2 \delta \xi_{0}^{2}+\frac{1}{2 \delta} y^{2},
$$

there is $C_{1}>0$ such that (4.2) holds for all $|y|>1$.
For $|y| \leqslant 1$, notice that $\log L(y) \in C^{\infty}(\mathbb{R})$, and

$$
\log L(0)=0,\left.\quad \frac{\mathrm{~d}}{\mathrm{~d} y} \log L(y)\right|_{y=0}=\mathbb{E} \xi_{0}=0
$$

By Taylor's formula of order 2 , we have for all $y$ with $|y| \leqslant 1$,

$$
\log L(y) \leqslant \frac{1}{2} C_{2}^{2} y^{2}
$$

where

$$
C_{2}:=\sup _{|y| \leqslant 1}\left|\frac{\mathrm{~d}^{2}}{\mathrm{~d} y^{2}} \log L(y)\right|^{1 / 2} .
$$

Thus (4.2) follows with $K:=C_{1} \vee C_{2}$.
We now extend (4.1) from Gaussian distribution to general law $\mu$ satisfying (2.2), which is a generalization of the preceding lemma.

Lemma 4.3. Let $\left(X_{k}\right)$ be the moving average process given by (1.1) and $T_{n}(f)$ the Toeplitz matrix associated with the spectral density function $f$ of $\left(X_{k}\right)$, given in Remark 2.7. Assume the integrability condition (2.2) (but not the stronger $L S I$ ).

Let $X^{(n)}=\left(X_{1}, \ldots, X_{n}\right)^{*}, B$ be a real non-negative definite symmetric $n \times n$-matrix, and $\mu_{1}^{n}, \ldots, \mu_{n}^{n}$ the eigenvalues of the matrix $\sqrt{B} T_{n}(f) \sqrt{B}$. Then for all $\lambda \geqslant 0$ satisfying $\lambda \max _{1 \leqslant j \leqslant n} \mu_{j}^{n}<1 /\left(2 K^{2}\right)$, we have

$$
\log \mathbb{E} \exp \left(\lambda\left\langle X^{(n)}, B X^{(n)}\right\rangle\right) \leqslant-\frac{1}{2} \sum_{j=1}^{n} \log \left(1-2 K^{2} \lambda \mu_{j}^{n}\right),
$$

where $K>0$ is given in Lemma 4.2.

Proof. The main difficulty resides in the non-linear property of $\langle x, B x\rangle$. The trick consists to reduce it to an estimation of linear type in the following way:

$$
\mathbb{E}\left\{\mathrm{e}^{\frac{1}{2} t^{2}\left\langle X^{(n)}, B X^{(n)}\right\rangle}\right\}=\mathbb{E}\left\{\mathrm{e}^{\frac{1}{2} t^{2}\left|\sqrt{B} X^{(n)}\right|^{2}}\right\}=\int_{\mathbb{R}^{n}} \mathbb{E}\left\{\mathrm{e}^{\left.t / \sqrt{B} X^{(n)}, y^{(n)}\right\rangle}\right\} \gamma\left(\mathrm{d} y^{(n)}\right)
$$

where $\gamma$ is the standard Gaussian law $\mathcal{N}(0, I)$ on $\mathbb{R}^{n}$.
Since

$$
\left\langle\sqrt{B} X^{(n)}, y^{(n)}\right\rangle=\left\langle X^{(n)}, \sqrt{B} y^{(n)}\right\rangle=\sum_{k=1}^{n} X_{k}\left(\sqrt{B} y^{(n)}\right)_{k}=\sum_{j \in \mathbb{Z}} \xi_{j} \sum_{k=1}^{n} a_{j-k}\left(\sqrt{B} y^{(n)}\right)_{k} .
$$

We get by Lemma 4.2 and the i.i.d. property of $\left(\xi_{j}\right)$,

$$
\mathbb{E}\left\{\exp \left[t\left\langle\sqrt{B} X^{(n)}, y^{(n)}\right\rangle\right]\right\} \leqslant \exp \left[\frac{K^{2} t^{2}}{2} \sum_{j \in \mathbb{Z}}\left|\sum_{k=1}^{n} a_{j-k}\left(\sqrt{B} y^{(n)}\right)_{k}\right|^{2}\right] .
$$

Now observe that

$$
\begin{aligned}
\sum_{j \in \mathbb{Z}}\left|\sum_{k=1}^{n} a_{j-k}\left(\sqrt{B} y^{(n)}\right)_{k}\right|^{2} & =\sum_{k, l=1}^{n} \sum_{j \in \mathbb{Z}} a_{j-k} a_{j-l}\left(\sqrt{B} y^{(n)}\right)_{k}\left(\sqrt{B} y^{(n)}\right)_{l} \\
& =\sum_{k, l=1}^{n}\left(T_{n}(f)\right)_{k, l}\left(\sqrt{B} y^{(n)}\right)_{k}\left(\sqrt{B} y^{(n)}\right)_{l} \\
& =\left\langle y^{(n)}, \sqrt{B} T_{n}(f) \sqrt{B} y^{(n)}\right\rangle .
\end{aligned}
$$

Then letting $\mu_{1}^{n}, \ldots, \mu_{n}^{n}$ be the eigenvalues of the matrix $\sqrt{B} T_{n}(f) \sqrt{B}$ (which are also the eigenvalues of $\left.T_{n}(f) B\right)$, we get for all $t$ such that $K^{2} t^{2} \max _{1 \leqslant j \leqslant n} \mu_{j}^{n}<1$,

$$
\begin{aligned}
\mathbb{E}\left\{\exp \left[\frac{1}{2} t^{2}\left\langle X^{(n)}, B X^{(n)}\right\rangle\right]\right\} & \leqslant \int_{\mathbb{R}^{n}}\left\{\exp \left[\frac{K^{2} t^{2}}{2}\left\langle y^{(n)}, \sqrt{B} T_{n}(f) \sqrt{B} y^{(n)}\right\rangle\right]\right\} \gamma\left(\mathrm{d} y^{(n)}\right) \\
& =-\frac{1}{2} \sum_{j=1}^{n} \log \left(1-K^{2} t^{2} \mu_{j}^{n}\right)
\end{aligned}
$$

where the last equality follows by Lemma 4.1. Finally the desired result follows with $\lambda=t^{2} / 2$.
Remark 4.1. If we assume $\|g\|_{\infty}=\|g(\theta)\|_{L^{\infty}(\mathbb{T}, \mathrm{d} \theta)}<+\infty$, and $B=I$ we obtain exactly the result in Wu [24]. In fact in this case, we have for any $\lambda>0$ such that $2 \lambda K^{2}\|g\|_{\infty}^{2}<1$,

$$
\begin{equation*}
\log \mathbb{E} \mathrm{e}^{\lambda\left(X^{(n)}, X^{(n)}\right\rangle} \leqslant-\frac{1}{2} \log \left(1-2 \lambda K^{2}\|g\|_{\infty}^{2}\right)^{n}, \tag{4.3}
\end{equation*}
$$

because the eigenvalues of $T_{n}(f)$ are bounded by $\|f\|_{\infty}=\|g\|_{\infty}^{2}$.
Remark 4.2. Instead of Lemma 4.2, we can use the consequence of the LSI (5.3) below to prove Lemma 4.3, but (5.3) is stronger than (2.2).

### 4.2. Preparating lemmas

For an $n \times n$ matrix $A$, we consider the usual operator norm $\|A\|=\sup _{x \in \mathbb{R}^{n}}(|A x| /|x|)$. Recall (cf. Remarks 2.7) that for any real and even function $h \in L^{1}(\mathbb{T}, \mathrm{~d} \theta), T_{n}(h)$ is the Toeplitz matrix of order $n$ associated with $h$ i.e. $T_{n}(h)=\left(\hat{r}_{k-l}(h)\right)_{1 \leqslant k, l \leqslant n}$ where $\hat{r}_{k}(h)$ is the $k$ th Fourier coefficient of $h$ given by

$$
\hat{r}_{k}(h)=\frac{1}{2 \pi} \int_{\mathbb{T}} \mathrm{e}^{\mathrm{i} k \theta} h(\theta) \mathrm{d} \theta, \quad \forall k \in \mathbb{Z} .
$$

Lemma 4.4 ((Avram [1], Lemma 1)). If $f \in L^{q}(\mathbb{T})$ where $1 \leqslant q \leqslant \infty$, then for all $n>1$ we have $\left\|T_{n}(f)\right\| \leqslant$ $n^{1 / q}\|f\|_{q}$.

Lemma 4.5 ((Avram [1], Theorem 1)). Let $f_{k} \in L^{q_{k}}(\mathbb{T}, \mathrm{~d} \theta)$ with $q_{k} \geqslant 1$ for $k=1, \ldots, p$ and $\sum_{k=1}^{p}\left(1 / q_{k}\right) \leqslant 1$. Then

$$
\lim _{n \rightarrow \infty} \frac{1}{n} \operatorname{tr}\left(\prod_{k=1}^{p} T_{n}\left(f_{k}\right)\right)=\hat{r}_{0}\left(\prod_{k=1}^{p} f_{k}\right)
$$

Introduce now the Fejèr approximation of $g$ :

$$
g^{N}(\theta)=\sum_{j \in \mathbb{Z}} a_{j}^{N} \mathrm{e}^{\mathrm{i} j \theta}, \quad \forall \theta \in \mathbb{R}, \text { where } a_{j}^{N}=a_{j}\left(1-\frac{|j|}{N}\right) 1_{|j| \leqslant N} .
$$

We recall the following (see [8])
Lemma 4.6. $g^{N}(\theta)=\int_{-\pi}^{\pi} g(\theta-t) K_{N}(t) \mathrm{d} t$ where $K_{N}$ is the Fejèr kernel of order $N$ given by

$$
K_{N}(t)=\frac{1}{2 \pi N}\left(\frac{\sin (N t / 2)}{\sin (t / 2)}\right)^{2}, \quad t \in \mathbb{T} .
$$

Furthermore for $g \in L^{p}(\mathbb{T})$ where $1 \leqslant p<\infty, g^{N} \rightarrow g$ in $L^{p}(\mathbb{T})$ and $g^{N} \rightarrow g$ uniformly on $\mathbb{T}$ if $g$ is continuous. Moreover, $K_{n}$ is even, non-negative and possesses the following properties for small $\delta$ :
(a) $\int_{\mathbb{T}} K_{n}(t) \mathrm{d} t=1$,
(b) $\int_{|t| \geqslant \delta} K_{n}(t) \mathrm{d} t \leqslant \frac{C}{n}$,
(c) $\int_{|t| \leqslant \delta} K_{n}(t) t^{\alpha} \mathrm{d} t \leqslant \begin{cases}C n^{-\alpha}, & \alpha<1, \\ C n^{-1} \ln n, & \alpha=1, \\ C n^{-1}, & \alpha>1 .\end{cases}$

Let $m$ be a given positive integer, a sequence $\left(Z_{n}\right)_{n \geqslant 1}$ of strictly stationary random variables is called $m$-dependent if for every $k \geqslant 1$ the two collections $\left\{Z_{1}, \ldots, Z_{k}\right\}$ and $\left\{Z_{k+m}, Z_{k+m+1}, \ldots\right\}$ are independent. We have the following

Lemma 4.7 ((Chen X. [9])). Let $\left(Z_{n}\right)_{n} \geqslant 1$ be a stationary sequence of m-dependent random variables taking values in $\mathbb{R}^{m}$, such that

$$
\mathbb{E}\left(\mathrm{e}^{\alpha\left|Z_{1}\right|}\right)<+\infty, \quad \text { for some } \alpha>0 \text {. }
$$

Then for all $\lambda \in \mathbb{R}^{m}$,

$$
\begin{aligned}
\lim _{n \rightarrow+\infty} \frac{1}{b_{n}^{2}} \log \mathbb{E}\left(\mathrm{e}^{b_{n}^{2}\left\langle\lambda, \frac{1}{\sqrt{n b_{n}}} \sum_{k=1}^{n}\left(Z_{k}-\mathbb{E} Z_{k}\right)\right.}\right) & =\frac{1}{2} \lim _{n \rightarrow \infty} \mathbb{E}\left\langle\lambda, \sum_{k=1}^{n}\left(Z_{k}-\mathbb{E} Z_{k}\right)\right)^{2} \\
& =\frac{1}{2}\left(\mathbb{E}\left\langle\lambda, Z_{1}\right\rangle^{2}+2 \sum_{k=2}^{m+1} \mathbb{E}\left\langle\lambda, Z_{1}\right\rangle \mathbb{E}\left\langle\lambda, Z_{k}\right\rangle\right) .
\end{aligned}
$$

## 5. Proofs of the main results in Section 2

### 5.1. Proof of Theorem 2.1

The proof is divided into three steps. In the first one, we approximate the moving average process by a bilateral moving average process of finite range $2 N$ which satisfies the MDP. Then we will show that this approximation is
a good one in the sense of the MDP. In third step, we will finally establish the convergence of the rate function and the subsequent existence of the limiting variance.

Step 1 (Approximation by bilateral moving average process of finite range $2 N$ ). Let $X_{k}^{N}=\sum_{j \in \mathbb{Z}} a_{j}^{N} \xi_{k+j}$, where $a_{j}^{N}=a_{j}\left(1-\frac{|j|}{N}\right) 1_{|j| \leqslant N}$, be the Fejèr approximation of $X_{k}$.

Put

$$
Q_{n}^{N}=\left(Q_{n}^{N, l}\right)=\left(\frac{1}{\sqrt{n} b_{n}} Z_{n}^{N, l}\right)_{l=0, \ldots, m} \quad \text { and } \quad Q_{n}=\left(Q_{n}^{l}\right)_{l=0, \ldots, m}=\left(\frac{1}{\sqrt{n} b_{n}} Z_{n}^{l}\right)
$$

where

$$
Z_{n}^{N, l}=\sum_{k=1}^{n}\left(X_{k}^{N} X_{k+l}^{N}-\mathbb{E} X_{k}^{N} X_{k+l}^{N}\right) \quad \text { and } \quad Z_{n}^{l}=\sum_{k=1}^{n}\left(X_{k} X_{k+l}-\mathbb{E} X_{k} X_{k+l}\right)
$$

The crucial remark is that the sequence $\left\{\left(X_{k}^{N} X_{k+l}^{N}\right)_{l=0, \ldots, m} \in \mathbb{R}^{m+1}, k \in \mathbb{Z}\right\}$ is a $2 N$-dependent stationary sequence. By (2.2), we get for all $N$ there is $\eta>0$ such that $\mathbb{E}\left(\mathrm{e}^{\eta\left|X_{k}^{N} X_{k+l}^{N}\right|}\right)<\infty$.

Then applying Lemma 4.7, we get that for each $N$ fixed, for all $\lambda \in \mathbb{R}^{m+1}$,

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \frac{1}{b_{n}^{2}} \log \mathbb{E}\left(\mathrm{e}^{b_{n}^{2}\left\langle\lambda, Q_{n}^{N}\right\rangle}\right)=\frac{1}{2} \lim _{n \rightarrow \infty} \frac{1}{n} \mathbb{E}\left\langle\lambda, Z_{n}^{N, \cdot}\right\rangle^{2}:=\frac{1}{2}\left\langle\lambda, \Sigma^{2, N} \lambda\right\rangle \in \mathbb{R} \tag{5.1}
\end{equation*}
$$

where $\Sigma^{2, N}$ is the limit covariance matrix given in Lemma 4.7, and that $Q_{n}^{N}$ satisfies the MDP on $\mathbb{R}^{m+1}$ with the good rate function

$$
I^{N}(x)=\sup _{\lambda \in \mathbb{R}^{m+1}}\left\{\langle\lambda, x\rangle-\frac{1}{2}\left\langle\lambda, \Sigma^{2, N} \lambda\right\rangle\right\} .
$$

Furthermore, by [21], $\Sigma_{k, l}^{2, N}$ can be expressed as (2.4) with $f$ replaced by $f^{N}$.
Step 2 (Exponential contiguity, see Section 4.2 in [10]). The purpose of this step will be to prove the asymptotic negligibility of $Q_{n}-Q_{n}^{N}$ with respect to the MDP as $N$ goes to $\infty$, i.e. we will establish that for all $\lambda \in \mathbb{R}^{m+1}$,

$$
\limsup _{N \rightarrow \infty} \limsup _{n \rightarrow \infty} \frac{1}{b_{n}^{2}} \log \mathbb{E}\left(\mathrm{e}^{b_{n}^{2}\left\langle\lambda, Q_{n}-Q_{n}^{N}\right\rangle}\right)=0
$$

As our functional $Q_{n}-Q_{n}^{N}$ are centered, by Jensen inequality we only have to establish the upper inequality in the equality above. By Jensen's inequality again,

$$
\mathbb{E}\left(\mathrm{e}^{b_{n}^{2}\left\langle\lambda, Q_{n}-Q_{n}^{N}\right\rangle}\right) \leqslant \frac{1}{m+1} \sum_{l=0}^{m} \mathbb{E}\left(\mathrm{e}^{(m+1) b_{n}^{2} \lambda_{l}\left(Q_{n}^{l}-Q_{n}^{N, l}\right)}\right),
$$

we need only to show that for each $l=0, \ldots, m$ fixed and for every $\lambda \in \mathbb{R}$,

$$
\begin{equation*}
\lim _{N \rightarrow \infty} \limsup _{n \rightarrow \infty} \frac{1}{b_{n}^{2}} \log \mathbb{E}\left(\mathrm{e}^{b_{n}^{2} \lambda\left(Q_{n}^{l}-Q_{n}^{N, l}\right)}\right) \leqslant 0 . \tag{5.2}
\end{equation*}
$$

To this end, our main tool is the following consequence of the LSI (2.1), see Ledoux [20, Theorem 2.7] (after having extended (2.1) by tensorization to the product measure of $\mu$ ): for any integrable $C^{1}$ functional $G$ of $\xi=\left(\xi_{k}\right)_{|k| \leqslant m}$,

$$
\begin{equation*}
\mathbb{E}\left(\mathrm{e}^{\lambda\left(b_{n} / \sqrt{n}\right)(G-\mathbb{E} G)}\right) \leqslant \mathbb{E}\left(\mathrm{e}^{\lambda^{2}\left(b_{n}^{2} / n\right) C\left|\nabla_{\xi} G\right|^{2}}\right), \tag{5.3}
\end{equation*}
$$

with $C$ given in (2.1), where $\left|\nabla_{\xi} G\right|^{2}:=\sum_{k}\left|\partial_{\xi_{k}} G\right|^{2}$. This inequality can be extended to all integrable functionals $G=F\left(X_{1}, \ldots, X_{n}\right)$ where $F \in C^{1}\left(\mathbb{R}^{n}\right)$ by dominated convergence (even now $X_{k}$ depends on the infinite sequence $\left(\xi_{k}\right)_{k \in \mathbb{Z}}$, the detail is left to the reader).

Let apply it to

$$
G_{n}^{N, l}\left(\left(\xi_{i}\right)_{i \in \mathbb{Z}}\right)=\sum_{k=1}^{n}\left(X_{k} X_{k+l}-X_{k}^{N} X_{k+l}^{N}\right)
$$

so that our main estimations are now transferred to the gradient of $G_{n}^{N, l}$.

## Clearly

$$
\partial_{\xi_{i}} G_{n}^{N, l}=\sum_{k=1}^{n}\left(a_{i-k} X_{k+l}+a_{i-k-l} X_{k}-a_{i-k}^{N} X_{k+l}^{N}-a_{i-k-l}^{N} X_{k}^{N}\right) ;
$$

so

$$
\begin{aligned}
\left|\nabla G_{n}^{N, l}\right|^{2} \leqslant & 4 \sum_{i \in \mathbb{Z}}\left(\left(\sum_{k=1}^{n}\left(a_{i-k}-a_{i-k}^{N}\right) X_{k+l}\right)^{2}+\left(\sum_{k=1}^{n}\left(a_{i-k-l}-a_{i-k-l}^{N}\right) X_{k}\right)^{2}\right. \\
& \left.+\left(\sum_{k=1}^{n} a_{i-k}^{N}\left(X_{k+l}-X_{k+l}^{N}\right)\right)^{2}+\left(\sum_{k=1}^{n} a_{i-k-l}^{N}\left(X_{k}-X_{k}^{N}\right)\right)^{2}\right) \\
= & (I)+(I I)+(I I I)+(I V) .
\end{aligned}
$$

By Hölder inequality,

$$
\begin{align*}
\log \mathbb{E}\left(\mathrm{e}^{\lambda\left(b_{n} / \sqrt{n}\right)\left(G_{n}^{N, l}-\mathbb{E} G_{n}^{N, l}\right)}\right) \leqslant & \log \mathbb{E}\left(\mathrm{e}^{C \lambda^{2}\left(b_{n}^{2} / n\right)\left\|\nabla_{\xi} G_{n}^{N, l}\right\|^{2}}\right) \\
\leqslant & \frac{1}{4} \log \mathbb{E}\left(\mathrm{e}^{4 C \lambda^{2}\left(b_{n}^{2} / n\right)(I)}\right)+\frac{1}{4} \log \mathbb{E}\left(\mathrm{e}^{4 C \lambda^{2}\left(b_{n}^{2} / n\right)(I I)}\right) \\
& +\frac{1}{4} \log \mathbb{E}\left(\mathrm{e}^{4 C \lambda^{2}\left(b_{n}^{2} / n\right)(I I I)}\right)+\frac{1}{4} \log \mathbb{E}\left(\mathrm{e}^{4 C \lambda^{2}\left(b_{n}^{2} / n\right)(I V)}\right) . \tag{5.4}
\end{align*}
$$

Let us deal with the first term of this inequality. Using the definition of $\hat{r}_{l}$ given in (2.5) and the fact that the spectral density of $\left(X_{k}-X_{k}^{N}\right)$ is $\left|g-g^{N}\right|^{2}$, we rewrite the expression of $(I)$ as

$$
\begin{aligned}
(I) & =4 \sum_{i \in \mathbb{Z}} \sum_{k, k^{\prime}=1}^{n}\left(a_{i-k}-a_{i-k}^{N}\right)\left(a_{i-k^{\prime}}-a_{i-k^{\prime}}^{N}\right) X_{k+l} X_{k^{\prime}+l}=4 \sum_{k, k^{\prime}=1}^{n} \hat{r}_{k^{\prime}-k}\left(\left|g-g^{N}\right|^{2}\right) X_{k+l} X_{k^{\prime}+l} \\
& =4\left\langle X_{\cdot+l}^{(n)}, T_{n}\left(\left|g-g^{N}\right|^{2}\right) X_{\cdot+l}^{(n)}\right\rangle
\end{aligned}
$$

where $X_{++l}^{(n)}=\left(X_{l+1}, \ldots, X_{l+n}\right)^{*}$. Let $\mu_{1}^{n, N}, \ldots, \mu_{n}^{n, N}$ be the eigenvalues of the matrix

$$
\sqrt{T_{n}\left(\left|g-g^{N}\right|^{2}\right)} T_{n}(f) \sqrt{T_{n}\left(\left|g-g^{N}\right|^{2}\right)}
$$

Its operator norm is bounded from above by (using Lemma 4.4)

$$
\left\|T_{n}(f)\right\| \cdot\left\|T_{n}\left(\left|g-g^{N}\right|^{2}\right)\right\| \leqslant n^{1 / q}\|f\|_{q} n^{1 / q}\left\|\left|g-g^{N}\right|^{2}\right\|_{q} .
$$

Since $\left(b_{n} / \sqrt{n}\right) n^{1 / q} \rightarrow 0$ by (H2) and $f \in L^{q}(\mathbb{T}, \mathrm{~d} \theta)$ by (H1), we have for all $n$ sufficiently large, $32 C K^{2} \lambda^{2} b_{n}^{2} / n$ $\max _{1 \leqslant j \leqslant n} \mu_{j}^{n, N}<1$. Applying the crucial Lemma 4.3, we get

$$
\begin{equation*}
\log \mathbb{E}\left(\mathrm{e}^{4 C \lambda^{2}\left(b_{n}^{2} / n\right)(I)}\right) \leqslant-\frac{1}{2} \sum_{j=1}^{n} \log \left(1-32 C K^{2} \lambda^{2} \frac{b_{n}^{2}}{n} \mu_{j}^{n, N}\right) \tag{5.5}
\end{equation*}
$$

Similarly, for all $n$ sufficiently large such that $32 C K^{2} \lambda^{2}\left(b_{n}^{2} / n\right) \max _{1 \leqslant j \leqslant n} \mu_{j}^{n, N}<1$, we have

$$
\begin{equation*}
\log \mathbb{E}\left(\mathrm{e}^{4 C \lambda^{2}\left(b_{n}^{2} / n\right)(I I)}\right)=\log \mathbb{E} \mathrm{e}^{16 C\left(b_{n}^{2} / n\right) \lambda^{2}\left\langle X^{(n)}, T_{n}\left(\left|g-g^{N}\right|^{2}\right) X^{(n)}\right\rangle} \leqslant-\frac{1}{2} \sum_{j=1}^{n} \log \left(1-32 C K^{2} \lambda^{2} \frac{b_{n}^{2}}{n} \mu_{j}^{n, N}\right) \tag{5.6}
\end{equation*}
$$

Let us deal with the third term. We rewrite the expression of (III) as

$$
(I I I)=4 \sum_{i \in \mathbb{Z}} \sum_{k, k^{\prime}=1}^{n} a_{i-k}^{N} a_{i-k^{\prime}}^{N}\left(X_{k+l}-X_{k+l}^{N}\right)\left(X_{k^{\prime}+l}-X_{k^{\prime}+l}^{N}\right)
$$

$$
\begin{aligned}
& =4 \sum_{k, k^{\prime}=1}^{n} \hat{r}_{k^{\prime}-k}\left(\left|g^{N}\right|^{2}\right)\left(X_{k+l}-X_{k+l}^{N}\right)\left(X_{k^{\prime}+l}-X_{k^{\prime}+l}^{N}\right) \\
& =4\left\langle X_{++l}^{(n)}-\left(X^{N}\right)_{+l}^{(n)}, T_{n}\left(\left|g^{N}\right|^{2}\right)\left(X_{\cdot+l}^{(n)}-\left(X^{N}\right)_{+l}^{(n)}\right)\right\rangle,
\end{aligned}
$$

where $\left(X^{N}\right)_{+l}^{(n)}=\left(X_{1+l}^{N}, \ldots, X_{n+l}^{N}\right)^{*}$. Let $v_{1}^{n, N}, \ldots, v_{n}^{n, N}$ the eigenvalues of the matrix

$$
\sqrt{T_{n}\left(\left|g^{N}\right|^{2}\right)} T_{n}\left(\left|g-g^{N}\right|^{2}\right) \sqrt{T_{n}\left(\left|g^{N}\right|^{2}\right)}
$$

Its operator norm is bounded from above by (using Lemma 4.4)

$$
\left\|T_{n}\left(\left|g^{N}\right|^{2}\right)\right\| \cdot\left\|T_{n}\left(\left|g-g^{N}\right|^{2}\right)\right\| \leqslant n^{1 / q}\left\|\left|g^{N}\right|^{2}\right\|_{q} n^{1 / q}\left\|\left|g-g^{N}\right|^{2}\right\|_{q} .
$$

By our assumptions (H1) and (H2) on $b_{n}$ and $f$, we have for all $n$ sufficiently large, $32 C K^{2} \lambda^{2}\left(b_{n}^{2} / n\right) \max _{1 \leqslant j \leqslant n} v_{j}^{n, N}$ $<1$. Applying the crucial Lemma 4.3, we get

$$
\begin{equation*}
\log \mathbb{E}\left(\mathrm{e}^{4 C \lambda^{2}\left(b_{n}^{2} / n\right)(I I I)}\right) \leqslant-\frac{1}{2} \sum_{j=1}^{n} \log \left(1-32 C K^{2} \lambda^{2} \frac{b_{n}^{2}}{n} \nu_{j}^{n, N}\right) \tag{5.7}
\end{equation*}
$$

Similarly for all $n$ sufficiently large such that $32 C K^{2} \lambda^{2}\left(b_{n}^{2} / n\right) \max _{1 \leqslant j \leqslant n} v_{j}^{n, N}<1$ and we have

$$
\begin{equation*}
\log \mathbb{E}\left(\mathrm{e}^{4 C \lambda^{2}\left(b_{n}^{2} / n\right)(I V)}\right) \leqslant-\frac{1}{2} \sum_{j=1}^{n} \log \left(1-32 C K^{2} \lambda^{2} \frac{b_{n}^{2}}{n} \nu_{j}^{n, N}\right) . \tag{5.8}
\end{equation*}
$$

By (5.4) and the previous estimations (5.5)-(5.8), we obtain

$$
\begin{equation*}
\log \mathbb{E}\left(\mathrm{e}^{\lambda b_{n}^{2}\left(Q_{n}^{l}-Q_{n}^{N, l}\right)}\right) \leqslant-\frac{1}{4} \sum_{j=1}^{n}\left(\log \left(1-32 C K^{2} \lambda^{2} \frac{b_{n}^{2}}{n} \mu_{j}^{n, N}\right)+\log \left(1-32 C K^{2} \lambda^{2} \frac{b_{n}^{2}}{n} \nu_{j}^{n, N}\right)\right) . \tag{5.9}
\end{equation*}
$$

Notice that by the Taylor's expansion of order 1, we have for $|z|<1$

$$
\log (1-z)=-z(1-t z)^{-1}
$$

where $t=t(z) \in[0,1]$. This applied here to $z_{j}^{n, N}=32 C K^{2} \lambda^{2}\left(b_{n}^{2} / n\right) \lambda_{j}^{n, N}$, where $\lambda_{j}^{n, N}=v_{j}^{n, N}$ or $\lambda_{j}^{n, N}=\mu_{j}^{n, N}$ which satisfies $\sup _{1 \leqslant j \leqslant n}\left|z_{j}^{n, N}\right| \rightarrow 0$ as $n \rightarrow \infty$, yields by (5.9),

$$
\limsup _{n \rightarrow \infty} \frac{1}{b_{n}^{2}} \log \mathbb{E}\left(\mathrm{e}^{\lambda b_{n}^{2}\left(Q_{n}^{l}-Q_{n}^{N, l}\right)}\right) \leqslant 16 C^{2} \lambda^{2} \lim _{n \rightarrow \infty}\left(\frac{1}{n} \sum_{j=1}^{n}\left(\mu_{j}^{n, N}+v_{j}^{n, N}\right)\right) .
$$

Thanks to Lemma 4.5, we have

$$
\lim _{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^{n} \mu_{j}^{n, N}=\lim _{n \rightarrow \infty} \frac{1}{n} \operatorname{tr}\left(T_{n}(f) T_{n}\left(\left|g-g^{N}\right|^{2}\right)\right)=\hat{r}_{0}\left(\left|g-g^{N}\right|^{2} f\right)
$$

Similarly

$$
\lim _{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^{n} v_{j}^{n, N}=\lim _{n \rightarrow \infty} \frac{1}{n} \operatorname{tr}\left(T_{n}\left(\left|g^{N}\right|^{2}\right) T_{n}\left(\left|g-g^{N}\right|^{2}\right)\right)=\hat{r}_{0}\left(\left|g^{N}\right|^{2}\left|g-g^{N}\right|^{2}\right)
$$

So we get

$$
\limsup _{n \rightarrow \infty} \frac{1}{b_{n}^{2}} \log \mathbb{E}\left(\mathrm{e}^{\lambda b_{n}^{2}\left(Q_{n}^{l}-Q_{n}^{N, l}\right)}\right) \leqslant 16 C^{2} \lambda^{2}\left[\hat{r}_{0}\left(\left|g-g_{N}\right|^{2} f\right)+\hat{r}_{0}\left(\left|g-g^{N}\right|^{2}\left|g_{N}\right|^{2}\right)\right]
$$

where the desired negligibility (5.2) follows.
Step 3. Now we establish (2.3), i.e., for all $\lambda \in \mathbb{R}^{m+1}$,

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \frac{1}{b_{n}^{2}} \log \mathbb{E} \mathrm{e}^{\left.b_{n}^{2} \lambda \lambda, Q_{n}\right\rangle}=\frac{1}{2}\left\langle\lambda, \Sigma^{2} \lambda\right\rangle . \tag{5.10}
\end{equation*}
$$

At first we have $\Sigma_{k, l}^{2, N} \rightarrow \Sigma_{k, l}^{2}$ for all $k, l$ as $N$ goes to infinity, by (H1) and the expression (2.4). Next for any fixed $\alpha, \beta>1$ with $\frac{1}{\alpha}+\frac{1}{\beta}=1$, by Hölder inequality we have that

$$
\log \mathbb{E} \mathrm{e}^{b_{n}^{2}\left\langle\lambda, Q_{n}\right\rangle} \leqslant \frac{1}{\alpha} \log \mathbb{E} \mathrm{e}^{\alpha b_{n}^{2}\left\langle\lambda, Q_{n}^{N}\right\rangle}+\frac{1}{\beta} \log \mathbb{E} \mathrm{e}^{\beta b_{n}^{2}\left\langle\lambda, Q_{n}-Q_{n}^{N}\right\rangle}
$$

for all $\lambda \in \mathbb{R}^{m+1}$. From (5.1) and (5.2) it follows that

$$
\limsup _{n \rightarrow \infty} \frac{1}{b_{n}^{2}} \log \mathbb{E} \mathrm{e}^{b_{n}^{2}\left\langle\lambda, Q_{n}\right\rangle} \leqslant \frac{\alpha}{2}\left\langle\lambda, \Sigma^{2, N} \lambda\right\rangle+\delta_{N}
$$

where $\delta_{N}:=\lim \sup _{n \rightarrow \infty}\left(1 / \beta b_{n}^{2}\right) \log \mathbb{E} \mathrm{e}^{\beta b_{n}^{2}\left\langle\lambda, Q_{n}-Q_{n}^{N}\right\rangle} \rightarrow 0$. Letting $N \rightarrow \infty$, we get

$$
\begin{equation*}
\limsup _{n \rightarrow \infty} \frac{1}{b_{n}^{2}} \log \mathbb{E} \mathrm{e}^{b_{n}^{2}\left\langle\lambda, Q_{n}\right\rangle} \leqslant \frac{\alpha}{2}\left\langle\lambda, \Sigma^{2} \lambda\right\rangle \tag{5.11}
\end{equation*}
$$

Similarly, by Hölder inequality, we have for every $\lambda$,

$$
\frac{1}{b_{n}^{2}} \log \mathbb{E} \mathrm{e}^{\alpha^{-1} b_{n}^{2}\left\langle\lambda, Q_{n}^{N}\right\rangle} \leqslant \frac{1}{b_{n}^{2}}\left(\frac{1}{\alpha} \log \mathbb{E} \mathrm{e}^{b_{n}^{2}\left\langle\lambda, Q_{n}\right\rangle}+\frac{1}{\beta} \log \mathbb{E} \mathrm{e}^{\left(\beta b_{n}^{2} / \alpha\right)\left\langle\lambda, Q_{n}^{N}-Q_{n}\right\rangle}\right)
$$

Taking first $\liminf _{n \rightarrow \infty}$ and next $\lim _{N \rightarrow \infty}$ we get from (5.1) and (5.2)

$$
\begin{equation*}
\frac{1}{2 \alpha^{2}}\left\langle\lambda, \Sigma^{2} \lambda\right\rangle \leqslant \liminf _{n \rightarrow \infty} \frac{1}{b_{n}^{2}} \log \mathbb{E} \mathrm{e}^{b_{n}^{2}\left\langle\lambda, Q_{n}\right\rangle} \tag{5.12}
\end{equation*}
$$

Letting $\alpha \rightarrow 1$ in (5.11) and (5.12) yields (5.10). Finally the desired MDP follows from (5.10) by Ellis-Gärtner's theorem ([10], Section 2.3).

### 5.2. Proof of Theorem 2.3

We begin with the following
Lemma 5.1. Under the hypothesis Theorem 2.3, we have that for all $h \in L^{p^{\prime}}(\mathbb{T}, \mathrm{d} \theta)$,

$$
\begin{equation*}
\Lambda(h):=\limsup _{n \rightarrow \infty} \frac{1}{b_{n}^{2}} \log \mathbb{E}\left(\mathrm{e}^{b_{n}^{2}(1 / 2 \pi) \int_{\mathbb{T}} h(\theta) \mathcal{L}_{n}(\theta) \mathrm{d} \theta}\right) \leqslant \frac{4 C K^{2}}{2 \pi} \int_{\mathbb{T}} f^{2}(t) h^{2}(t) \mathrm{d} t \tag{5.13}
\end{equation*}
$$

In particular $\mathbb{P}\left(\mathcal{L}_{n} \in \cdot\right)$ is exponentially ${ }^{*}$-tight in $\left(L^{p}(\mathbb{T}, \mathrm{~d} \theta), \sigma\left(L^{p}(\mathbb{T}, \mathrm{~d} \theta), L^{p^{\prime}}(\mathbb{T}, \mathrm{d} \theta)\right)\right)$, where $1 / p^{\prime}+1 / p=1$.
Proof. The last claim follows from (5.13) by [23, Chapter 2, Proposition 2.5] when $1<p<2$ and by [23, Chapter 2, Theorem 2.1] when $p=1$. So it is enough to prove (5.13). For every function $h \in L^{p^{\prime}}(\mathbb{T}, \mathrm{d} \theta)$, the function $\tilde{h}(\theta)=$ $\frac{1}{2}[h(\theta)+h(-\theta)]$ is even and

$$
\frac{1}{2 \pi} \int_{\mathbb{T}} h(\theta) \mathcal{I}_{n}(\theta) \mathrm{d} \theta=\frac{1}{2 \pi} \int_{\mathbb{T}} \tilde{h}(\theta) \mathcal{I}_{n}(\theta) \mathrm{d} \theta
$$

we shall hence restrict ourselves to the case where $h$ is even. Since

$$
\frac{1}{2 \pi} \int_{\mathbb{T}} h(\theta) \mathcal{L}_{n}(\theta) \mathrm{d} \theta=\frac{1}{b_{n} \sqrt{n}}\left(\left\langle X_{\cdot}^{(n)}, T_{n}(h) X_{\cdot}^{(n)}\right\rangle-\mathbb{E}\left\langle X_{\cdot}^{(n)}, T_{n}(h) X_{\cdot}^{(n)}\right\rangle\right)
$$

Applying (2.1) to $H\left(\left(\xi_{l}\right)_{l \in \mathbb{Z}}\right)=\left\langle X^{(n)}, T_{n}(h) X^{(n)}\right\rangle$, we have

$$
\mathbb{E}\left(\mathrm{e}^{b_{n}^{2}(1 /(2 \pi)) \int_{\mathbb{T}} h(\theta) \mathcal{L}_{n}(\mathrm{~d} \theta)}\right)=\mathbb{E}\left(\mathrm{e}^{\left(b_{n} / \sqrt{n}\right)(H-\mathbb{E} H)}\right) \leqslant \mathbb{E}\left(\mathrm{e}^{\left(b_{n}^{2} / n\right) C\left|\nabla_{\xi} H\right|^{2}}\right)
$$

Since $T_{n}(h)$ is symmetric, we have

$$
\partial_{\xi_{i}} H\left(\left(\xi_{l}\right)_{l \in \mathbb{Z}}\right)=\sum_{l, k=1}^{n} a_{i-k} X_{l} T_{n}(h)_{k, l}+a_{i-l} X_{k} T_{n}(h)_{k, l}=2 \sum_{l, k=1}^{n} a_{i-k} X_{l} T_{n}(h)_{k, l}
$$

Clearly

$$
\begin{aligned}
\left|\nabla_{\xi} H\right|^{2} & =\sum_{i \in \mathbb{Z}}\left(\partial_{\xi_{i}} H\right)^{2}=\sum_{i \in \mathbb{Z}}\left(2 \sum_{l, k=1}^{n} a_{i-k} X_{l} T_{n}(h)_{k, l}\right)^{2}=4 \sum_{l, l^{\prime}=1}^{n}\left(\sum_{k, k^{\prime}=1}^{n} T_{n}(h)_{k, l} T_{n}(f)_{k, k^{\prime}} T_{n}(h)_{k^{\prime}, l^{\prime}}\right) X_{l} X_{l^{\prime}} \\
& =4 \sum_{l, l^{\prime}=1}^{n}\left(T_{n}(h) T_{n}(f) T_{n}(h)\right)_{l, l^{\prime}} X_{l} X_{l^{\prime}}=4\left(X_{\cdot}^{(n)}, T_{n}(h) T_{n}(f) T_{n}(h) X_{\cdot}^{(n)}\right\rangle .
\end{aligned}
$$

Let $\alpha_{1}^{n}, \ldots, \alpha_{n}^{n}$ the eigenvalues of the matrix

$$
\sqrt{T_{n}(h) T_{n}(f) T_{n}(h)} T_{n}(f) \sqrt{T_{n}(h) T_{n}(f) T_{n}(h)}
$$

Its operator norm is bounded from above by (using Lemma 4.4)

$$
\left\|T_{n}(f)\right\| \cdot\left\|T_{n}(h) T_{n}(f) T_{n}(h)\right\| \leqslant\left(n^{1 / q}\|f\|_{q}\right)^{2}\left(n^{1 / p^{\prime}}\|h\|_{p^{\prime}}\right)^{2}
$$

Since $b_{n} n^{1 / q+1 / p^{\prime}-1 / 2} \rightarrow 0, f \in L^{q}(\mathbb{T}, \mathrm{~d} \theta)$ and $h \in L^{p^{\prime}}(\mathbb{T}, \mathrm{d} \theta)$, we have $8 C K^{2}\left(b_{n}^{2} / n\right) \max _{1 \leqslant j \leqslant n} \alpha_{j}^{n}<1$ for $n$ large enough. Applying Lemma 4.3, we get

$$
\log \mathbb{E}\left(\mathrm{e}^{\mathrm{b}_{n}^{2}(1 /(2 \pi)) \int_{\mathbb{T}} h(\theta) \mathcal{L}_{n}(\mathrm{~d} \theta)}\right) \leqslant-\frac{1}{2} \sum_{j=1}^{n} \log \left(1-8 C K^{2} \frac{b_{n}^{2}}{n} \alpha_{j}^{n}\right) .
$$

Thus

$$
\limsup _{n \rightarrow \infty} \frac{1}{b_{n}^{2}} \log \mathbb{E}\left(\mathrm{e}^{b_{n}^{2}(1 /(2 \pi)) \int_{\mathbb{T}} h(\theta) \mathcal{L}_{n}(\mathrm{~d} \theta)}\right) \leqslant 4 C K^{2} \lim _{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^{n} \alpha_{j}^{n} .
$$

Since $f \in L^{q}(\mathbb{T}, \mathrm{~d} \theta)$ and $h \in L^{p^{\prime}}(\mathbb{T}, \mathrm{d} \theta)$ with $1 / p^{\prime}+1 / q \leqslant \frac{1}{2}$, applying Lemma 4.5 , we obtain

$$
\lim _{n \rightarrow+\infty} \frac{1}{n} \sum_{j=1}^{n} \alpha_{j}^{n}=\lim _{n \rightarrow+\infty} \frac{1}{n} \operatorname{tr}\left(\left(T_{n}(f) T_{n}(h)\right)^{2}\right)=\hat{r}_{0}\left(f^{2} h^{2}\right)<+\infty .
$$

Hence (5.13) follows.
We now turn to the
Proof of Theorem 2.3. Step 1. By Lemma 4.3, $\mathbb{E} \mathrm{e}^{\lambda\left|X_{0}\right|^{2}}<+\infty$ for some $\lambda>0$. Then by Chebychev inequality

$$
\left(\frac{1}{b_{n} \sqrt{n}} \sum_{k=n-\ell+1}^{n}\left(X_{k} X_{k+\ell}-\mathbb{E} X_{k} X_{k+\ell}\right)\right)_{0 \leqslant \ell \leqslant m}
$$

is negligible with respect to the MDP. Using Theorem 2.1, we get the finite dimensional MDP on $\mathbb{R}^{m+1}$ of

$$
\left(\frac{1}{b_{n} \sqrt{n}} \sum_{k=1}^{n-\ell}\left(X_{k} X_{k+\ell}-\mathbb{E} X_{k} X_{k+\ell}\right)\right)_{0 \leqslant \ell \leqslant m}
$$

with the rate function given by

$$
I(z)=\sup _{\lambda \in \mathbb{R}^{m+1}}\left\{\langle\lambda, z\rangle-\frac{1}{2}\left\langle\lambda, \Sigma^{2} \lambda\right\rangle\right\},
$$

where

$$
\left\langle\lambda, \Sigma^{2} \lambda\right\rangle=\frac{1}{2 \pi} \int_{\mathbb{T}} 2\left(\sum_{k=0}^{m} \lambda_{k} \cos (k \theta)\right)^{2} f^{2}(\theta) \mathrm{d} \theta+\kappa_{4}\left(\frac{1}{2 \pi} \int_{\mathbb{T}}\left(\sum_{k=0}^{m} \lambda_{k} \cos (k \theta)\right) f(\theta) \mathrm{d} \theta\right)^{2} .
$$

Now notice that $\mathcal{L}_{n}(\theta)$ is even and

$$
\hat{\mathcal{L}}_{n}(\ell):=\frac{1}{2 \pi} \int_{\mathbb{T}} \cos (\ell \theta) \mathcal{L}_{n}(\theta) \mathrm{d} \theta=\frac{1}{b_{n} \sqrt{n}} \sum_{k=1}^{n-\ell}\left(X_{k} X_{k+\ell}-\mathbb{E} X_{k} X_{k+\ell}\right), \quad \ell \geqslant 0 .
$$

Thus $\left(\hat{\mathcal{L}}_{n}(\ell)\right)_{0 \leqslant \ell \leqslant m}$ satisfies the MDP on $\mathbb{R}^{m+1}$ with the same rate function. By Lemma 4.3 and the projective limit Theorem [10, Theorem 4.6.9], we deduce that $\left(\mathcal{L}_{n}\right)_{n} \geqslant 0$ satisfies the MDP on $\left(L^{p}(\mathbb{T}, \mathrm{~d} \theta), \sigma\left(L^{p}(\mathbb{T}, \mathrm{~d} \theta), L^{p^{\prime}}(\mathbb{T}, \mathrm{d} \theta)\right)\right)$ with the rate function given by

$$
I(\eta)=\left\{\begin{array}{l}
\sup _{m \geqslant 0 \lambda_{0}, \ldots, \lambda_{m} \in \mathbb{R}} \sup _{m}\left\{\frac{1}{2 \pi} \int_{\mathbb{T}}\left(\sum_{k=0}^{m} \lambda_{k} \cos (k \theta)\right) \eta(\theta) \mathrm{d} \theta-\frac{1}{2} \Lambda\left(\sum_{k=0}^{m} \lambda_{k} \cos (k \theta)\right)\right\}, \quad \text { if } \eta \text { is even, }  \tag{5.14}\\
+\infty, \quad \text { otherwise }
\end{array}\right.
$$

where

$$
\Lambda\left(\sum_{k=0}^{m} \lambda_{k} \cos (k \theta)\right)=\left\langle\lambda, \Sigma^{2} \lambda\right\rangle .
$$

Step 2. Identification of the rate function. Introduce $L_{\text {even }}^{p}(\mathbb{T}, v)=\left\{h \in L^{p}(\mathbb{T}, \nu), h\right.$ even $\}$. Remark as trigonometric polynomials are dense in $L^{2}\left(\mathbb{T}, f^{2} \mathrm{~d} \theta\right)$, one can find for every $h \in L_{\text {even }}^{2}\left(\mathbb{T}, f^{2} \mathrm{~d} \theta\right)$, an approximation by some cosine polynomials sequence $h_{n}$, such that

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \int_{\mathbb{T}}\left(h_{n}-h\right)^{2}(\theta) f^{2}(\theta) \mathrm{d} \theta=0 \tag{5.15}
\end{equation*}
$$

So we can extend continuously the definition of $\Lambda$ to all functions $h \in L_{\text {even }}^{2}\left(\mathbb{T}, f^{2} \mathrm{~d} \theta\right)$,

$$
\begin{equation*}
\Lambda(h)=\frac{1}{2 \pi} \int_{\mathbb{T}} 2 h^{2}(\theta) f^{2}(\theta) \mathrm{d} \theta+\kappa_{4}\left(\frac{1}{2 \pi} \int_{\mathbb{T}} h(\theta) f(\theta) \mathrm{d} \theta\right)^{2} \tag{5.16}
\end{equation*}
$$

(a) Suppose that $\eta$ is even, $\eta \mathrm{d} \theta$ is absolutely continuous w.r.t. $f^{2} \mathrm{~d} \theta$, and $\eta / f \in L^{2}(\mathbb{T}, \mathrm{~d} \theta)$. For any $h \in$ $L_{\text {even }}^{2}\left(\mathbb{T}, f^{2} \mathrm{~d} \theta\right)$, let $h_{n}$ the sequence defined in (5.15), by Cauchy-Schwartz inequality, we get

$$
\left(\int_{\mathbb{T}}\left|\left(h_{n}-h\right)(\theta) \eta(\theta)\right| \mathrm{d} \theta\right)^{2} \leqslant \int_{\mathbb{T}}\left|h_{n}(\theta)-h(\theta)\right|^{2} f^{2}(\theta) \mathrm{d} \theta \int_{\mathbb{T}}\left(\frac{\eta}{f}\right)^{2}(\theta) \mathrm{d} \theta \underset{n \rightarrow \infty}{\longrightarrow} 0 .
$$

So $I(\eta)$ defined in (5.14) coincides with

$$
I(\eta)=\sup _{h \in L_{\text {even }}^{2}\left(\mathbb{T}, f^{2} \mathrm{~d} \theta\right)}\left\{\frac{1}{2 \pi} \int_{\mathbb{T}} h(\theta) \eta(\theta) \mathrm{d} \theta-\frac{1}{2} \Lambda(h)\right\}:=\sup _{h \in L_{\text {even }}^{2}\left(\mathbb{T}, f^{2} \mathrm{~d} \theta\right)} D(h)
$$

Let us find explicitly the maximizer $h_{0}$ of $D(h)$. Let $k \in L_{\text {even }}^{2}\left(\mathbb{T}, f^{2} \mathrm{~d} \theta\right)$ and $\epsilon>0$,

$$
\begin{aligned}
\lim _{\epsilon \rightarrow 0} \frac{D(h+\epsilon k)-D(h)}{\epsilon}= & \frac{1}{2 \pi} \int_{\mathbb{T}} k(\theta) \eta(\theta) \mathrm{d} \theta-\frac{1}{2}\left(\frac{2}{2 \pi} \int_{\mathbb{T}} 2 f^{2}(\theta) h(\theta) k(\theta) \mathrm{d} \theta\right. \\
& \left.+2 \kappa_{4}\left(\frac{1}{2 \pi} \int_{\mathbb{T}} f(\theta) h(\theta) \mathrm{d} \theta\right)\left(\frac{1}{2 \pi} \int_{\mathbb{T}} f(\theta) k(\theta) \mathrm{d} \theta\right)\right) .
\end{aligned}
$$

So

$$
\begin{equation*}
\lim _{\epsilon \rightarrow 0} \frac{D(h+\epsilon k)-D(h)}{\epsilon}=0, \quad \forall k \in L_{\text {even }}^{2}\left(\mathbb{T}, f^{2} \mathrm{~d} \theta\right) \tag{5.17}
\end{equation*}
$$

iff

$$
\begin{equation*}
\eta(\theta)=2 f(\theta)^{2} h(\theta)+\kappa_{4}\left(\frac{1}{2 \pi} \int_{\mathbb{T}} f(\theta) h(\theta) \mathrm{d} \theta\right) f(\theta) . \tag{5.18}
\end{equation*}
$$

Dividing (5.18) by $f$ and integrating over $\mathbb{T}$, we obtain

$$
\int_{\mathbb{T}} f(\theta) h(\theta) \mathrm{d} \theta=\frac{1}{2+\kappa_{4}} \int_{\mathbb{T}} \frac{\eta(\theta)}{f(\theta)} \mathrm{d} \theta
$$

Plugging this last expression in (5.18), it is then easy to verify that the only functional $h_{0} \in L_{\text {even }}^{2}\left(\mathbb{T}, f^{2} \mathrm{~d} \theta\right)$ realizing (5.17) is given by

$$
h_{0}(\theta) f(\theta)=\frac{\eta(\theta)}{2 f(\theta)}-\frac{\kappa_{4}}{2+\kappa_{4}}\left(\frac{1}{2 \pi} \int_{\mathbb{T}} \frac{\eta(u)}{2 f(u)} \mathrm{d} u\right) .
$$

Calculating $D\left(h_{0}\right)$ gives finally the announced rate function.
(b) Now we have to treat the case where $\eta \mathrm{d} \theta$ is absolutely continuous w.r.t. $f^{2} \mathrm{~d} \theta$ but $\frac{\eta}{f} \notin L^{2}(\mathbb{T}$, $\mathrm{d} \theta)$. So there exists $g \in L_{\text {even }}^{2}(\mathbb{T}, \mathrm{~d} \theta)$ such that $\int_{\mathbb{T}} g(\theta) \frac{\eta}{f}(\theta) \mathrm{d} \theta=+\infty$, and $g \frac{\eta}{f} \geqslant 0$. Let $h:=\frac{g}{f}$, so $h \in L_{\text {even }}^{2}\left(\mathbb{T}, f^{2} \mathrm{~d} \theta\right)$, we choose $h_{n}=(h \vee(-n)) \wedge n$. We get by dominated convergence

$$
\lim _{n \rightarrow \infty} \int_{\mathbb{T}}\left(h_{n}(\theta)-h(\theta)\right)^{2} f(\theta)^{2} \mathrm{~d} \theta=0
$$

so it follows that

$$
\lim _{n \rightarrow+\infty} \Lambda\left(h_{n}\right)=\Lambda(h) .
$$

By Fatou's lemma we get

$$
\liminf _{n \rightarrow \infty} \int_{\mathbb{T}} h_{n}(\theta) \eta(\theta) \mathrm{d} \theta \geqslant \int_{\mathbb{T}} \liminf _{n \rightarrow \infty} h_{n}(\theta) \eta(\theta) \mathrm{d} \theta=+\infty
$$

Since by approximation,

$$
I(\eta) \geqslant \frac{1}{2 \pi} \int_{\mathbb{T}} h_{n}(\theta) \eta(\theta) \mathrm{d} \theta-\frac{1}{2} \Lambda\left(h_{n}\right),
$$

letting $n$ to $\infty$, we obtain $I(\eta)=\infty$.
(c) Now we have to treat the case where $\eta \mathrm{d} \theta$ is not absolutely continuous w.r.t. $f^{2} \mathrm{~d} \theta$, i.e. there exists a measurable and symmetric set $K \subset \mathbb{T}$ such that $\int_{K} f^{2}(\theta) \mathrm{d} \theta=0$ while $\int_{K} \eta(\theta) \mathrm{d} \theta>0$. For any $t>0$, we approximate the function $t 1_{K}$ by a sequence of cosine polynomials $h_{n}$ in $L^{2}\left(\mathbb{T},\left(f^{2}+|\eta|\right) \mathrm{d} \theta\right)$ and get

$$
I(\eta) \geqslant \lim _{n \rightarrow+\infty} D\left(h_{n}\right) \geqslant t \int_{K} \eta(\theta) \mathrm{d} \theta
$$

Letting $t$ to infinity, we get $I(\eta)=+\infty$.

### 5.3. Proof of the corollaries of Theorem 2.3

Proof of Corollary 2.4. It is enough to prove it for $h$ even. When $h$ is a cosine polynomial, this was established in the proof of Theorem 2.3. For general $h \in L^{p^{\prime}}(\mathbb{T}, \mathrm{d} \theta)$, let $\left(h_{N}\right)$ be a sequence of cosine polynomials such that $h_{N} \rightarrow h$ in $L^{p^{\prime}}(\mathbb{T}, \mathrm{d} \theta)$. To get the desired result, it remains to show

$$
\lim _{N \rightarrow \infty} \limsup _{n \rightarrow \infty} \frac{1}{b_{n}^{2}} \log \mathbb{E} \exp \left(\lambda b_{n}^{2} \frac{1}{2 \pi} \int_{\mathbb{T}}\left(h_{N}-h\right) \mathcal{L}_{n}(\theta) \mathrm{d} \theta\right)=0, \quad \forall \lambda \in \mathbb{R} .
$$

This follows by Lemma 5.1.

Proof of Corollary 2.5. To deduce MDP for $\left(\tilde{\mathcal{L}}_{n}\right)$ from Theorem 2.3 (with $q=+\infty$ ), we only need to prove that for all $h \in L^{p^{\prime}}(\mathbb{T})$,

$$
\begin{equation*}
\frac{\sqrt{n}}{b_{n}}\left(\int_{\mathbb{T}} h(t) \mathbb{E} \mathcal{I}_{n}(t) \mathrm{d} t-\int_{\mathbb{T}} f(t) h(t) \mathrm{d} t\right) \underset{n \rightarrow 0}{\longrightarrow} 0 \tag{5.19}
\end{equation*}
$$

It is easy to see that

$$
\int_{\mathbb{T}} h(t) \mathbb{E} \mathcal{I}_{n}(t) \mathrm{d} t=\int_{\mathbb{T}} \int_{\mathbb{T}} K_{n}(u-t) f(u) h(t) \mathrm{d} t \mathrm{~d} u=\left\langle K_{n} * f, h\right\rangle
$$

where $K_{n}$ is the Fejèr kernel function given in Lemma 4.6. Since the function $K_{n}$ is even, we have

$$
\int_{\mathbb{T}} h(t) \mathbb{E} \mathcal{I}_{n}(t) \mathrm{d} t=\frac{1}{2} \iint_{\mathbb{T}} \int_{\mathbb{T}} K_{n}(t) f(u+t) h(t) \mathrm{d} t \mathrm{~d} u+\frac{1}{2} \iint_{\mathbb{T}} \int_{\mathbb{T}} K_{n}(t) f(u) h(t+u) \mathrm{d} t \mathrm{~d} u .
$$

Taking into account the equalities $\int_{\mathbb{T}} f(u) h(u) \mathrm{d} u=\int_{\mathbb{T}} f(u+t) h(u+t) \mathrm{d} u, \int_{\mathbb{T}} K_{n}(t) \mathrm{d} t=1$ we get

$$
\begin{aligned}
\left|\int_{\mathbb{T}} h(t) \mathbb{E} \mathcal{I}_{n}(t) \mathrm{d} t-\int_{\mathbb{T}} f(t) h(t) \mathrm{d} t\right| & =\frac{1}{2}\left|\int_{\mathbb{T}} K_{n}(t) \int_{\mathbb{T}}(f(u)-f(u+t))(h(t+u)-h(u)) \mathrm{d} u \mathrm{~d} t\right| \\
& \leqslant \frac{1}{2} \int_{\mathbb{T}} K_{n}(t)\|f(\cdot)-f(\cdot+t)\|_{p}\|h(t+\cdot)-h(\cdot)\|_{p^{\prime}} \mathrm{d} t .
\end{aligned}
$$

By our assumption (2.7) on $f$, for $\delta>0$ small and $|t| \leqslant \delta$ we have

$$
\|f(\cdot)-f(\cdot+t)\|_{p} \leqslant C \sqrt{|t|} \quad \text { and } \quad\|h(t+\cdot)-h(\cdot)\|_{p^{\prime}} \leqslant 2\|h\|_{p^{\prime}}
$$

By Lemma 4.6, the last quantity above is smaller than

$$
C\|h\|_{p^{\prime}} \int_{|t| \leqslant \delta} K_{n}(t) \sqrt{|t|} \mathrm{d} u+2\|f\|_{p}\|h\|_{p^{\prime}} \int_{|t| \geqslant \delta} K_{n}(t) \mathrm{d} t=\mathrm{O}\left(\frac{1}{\sqrt{n}}\right) .
$$

Hence (5.19) follows.
Proof of Corollary 2.6. By Corollary 2.4 with $p=1$, we only need to prove (5.19) for all $h$ satisfies (2.8), and the proof of (5.19) is completely similar to that of Corollary 2.5 .

### 5.4. Proof of Theorem 2.7

Let us describe briefly how the preceding proof of Theorem 2.1 can be easily extended to the general non-linear functional $F$. We only consider $F\left(x_{0}, \ldots, x_{l}\right)=F\left(x_{0}\right)$ and it is real-valued (for simplicity).

Since $F^{\prime}$ is Lipschitz continuous, we get for some positive $L$, and for all $N$

$$
\left|F\left(X_{k}^{N}\right)\right| \leqslant L\left(1+\left|X_{k}^{N}\right|^{2}\right) \leqslant 2 L(N+1)\left(1+\sum_{j=-N}^{N} a_{j}^{2} \xi_{k+j}^{2}\right)
$$

so that, setting $\delta^{\prime}=\delta /\left(2 L(N+1)^{2} \sup _{j} a_{j}^{2}\right)$ where $\delta$ is given in (2.2), by the assumption on the validity of the LSI, we get

$$
\mathbb{E}\left(\mathrm{e}^{\delta^{\prime}\left|F\left(X_{k}^{N}\right)\right|}\right) \leqslant \mathrm{e}^{\delta^{\prime} L(N+1)} \mathbb{E}\left(\mathrm{e}^{\delta \xi_{0}^{2}}\right)<\infty .
$$

Hence for every $N$ fixed, by Lemma 4.7, $(1 / n) \sum_{k=1}^{n} F\left(X_{k}^{N}\right)$ satisfies the MDP as in Step 1 in the proof of Theorem 2.1.

Thus by the argument in Step 3 of Theorem 2.1, it remains to prove that $\forall \lambda \in \mathbb{R}$,

$$
\limsup _{N \rightarrow \infty} \limsup _{n \rightarrow \infty} \frac{1}{b_{n}^{2}} \log \mathbb{E} \exp \left(\lambda \frac{b_{n}}{\sqrt{n}} \sum_{k=1}^{n}\left[F\left(X_{k}\right)-F\left(X_{k}^{N}\right)-\mathbb{E}\left(F\left(X_{k}\right)-F\left(X_{k}^{N}\right)\right)\right]\right)=0 .
$$

We apply again (5.3) to

$$
G_{n}^{N}\left(\left(\xi_{l}\right)_{l \in \mathbb{Z}}\right)=\sum_{k=1}^{n}\left(F\left(X_{k}\right)-F\left(X_{k}^{N}\right)\right)
$$

Writing $F^{\prime}(X):.=\left(F^{\prime}\left(X_{1}\right), \ldots, F^{\prime}\left(X_{n}\right)\right)^{*}$ and similarly $F^{\prime}(X$.$) , we have$

$$
\begin{aligned}
\left|\nabla_{\xi} G_{n}^{N}\right|^{2} & =\sum_{i \in \mathbb{Z}}\left(\sum_{k=1}^{n} a_{i-k} F^{\prime}\left(X_{k}\right)-a_{i-k}^{N} F^{\prime}\left(X_{k}^{N}\right)\right)^{2} \\
& \leqslant 2 \sum_{i \in \mathbb{Z}}\left(\sum_{k=1}^{n}\left(a_{i-k}-a_{i-k}^{N}\right) F^{\prime}\left(X_{k}\right)\right)^{2}+2 \sum_{i \in \mathbb{Z}}\left(\sum_{k=1}^{n} a_{i-k}^{N}\left(F^{\prime}\left(X_{k}^{N}\right)-F^{\prime}\left(X_{k}\right)\right)\right)^{2} \\
& =2\left|\sqrt{T_{n}\left(\left|g-g^{N}\right|^{2}\right)} F^{\prime}(X .)\right|^{2}+2\left|\sqrt{T_{n}\left(\left|g^{N}\right|^{2}\right)}\left(F^{\prime}(X .)-F^{\prime}\left(X_{.}^{N}\right)\right)\right|^{2}
\end{aligned}
$$

By the fact that the derivative of $F$ is Lipschitz and the spectral density is bounded, the last term above is bounded by

$$
2 L\left\|g-g^{N}\right\|_{\infty}^{2}\left(n+\sum_{k=1}^{n} X_{k}^{2}\right)+2\left\|g^{N}\right\|_{\infty}^{2} \sum_{k=1}^{n}\left(X_{k}^{N}-X_{k}\right)^{2}
$$

Finally as $\lambda^{2}\left(b_{n}^{2} / n\right)\left\|g-g^{N}\right\|_{\infty}^{2}$ can be chosen arbitrary small for large $n$, we have by Lemma 4.3

$$
\begin{aligned}
\frac{1}{b_{n}^{2}} \log \mathbb{E} \exp \left(\lambda \frac{b_{n}}{\sqrt{n}}\left(G_{n}^{N}-\mathbb{E} G_{n}^{N}\right)\right) \leqslant & L C \lambda^{2}\left\|g-g^{N}\right\|_{\infty}^{2}-\frac{n}{4 b_{n}^{2}} \log \left(1-4 C L K^{2} \lambda^{2} \frac{b_{n}^{2}}{n}\left\|g-g^{N}\right\|_{\infty}^{2}\|g\|_{\infty}^{2}\right) \\
& -\frac{n}{4 b_{n}^{2}} \log \left(1-4 C L K^{2} \lambda^{2} \frac{b_{n}^{2}}{n}\left\|g^{N}\right\|_{\infty}^{2}\left\|g-g^{N}\right\|_{\infty}^{2}\right)
\end{aligned}
$$

and the r.h.s. of this last inequality is easily seen to behave as $n \rightarrow \infty$ as

$$
\left\|g-g^{N}\right\|_{\infty}^{2}\left(L C \lambda^{2}+2 C L K^{2} \lambda^{2}\|g\|_{\infty}^{2}\right) .
$$

By the famous Fejèr Theorem (Lemma 4.6), under the assumption of continuity of $g$, we get that

$$
\lim _{N \rightarrow \infty}\left\|g-g^{N}\right\|_{\infty}^{2}=0
$$

which yields to the desired negligibility.

### 5.5. Proof of Corollary 2.8

Under assumption (2.2), the crucial inequality (5.3), as a consequence of the LSI, may not be used. However, we may encompass this difficulty by noting that integrability (2.2) is, by Djellout and al. [13, Theorem 2.3], equivalent to a Transport inequality in $L_{1}$-Wasserstein distance which is itself equivalent to the inequality (5.3) with the Lipschitz norm instead of the gradient in the right hand side, but for this particular linear case, the gradient and Lipschitz norm are equal so that the same proof works.

### 5.6. Proof of statistical results in Section 3

### 5.6.1. Proof of Proposition 3.1

Considering $X_{n} / \sigma$ if necessary, we can assume without loss of generality that $\sigma=1$. Let us introduce

$$
r_{n}:=\frac{\sqrt{n}}{b_{n}}\left(\hat{\theta}_{n}-\theta\right) \quad \text { and } \quad R_{n}=\frac{1-\theta^{2}}{\sqrt{n} b_{n}} \sum_{i=1}^{n}\left(X_{i} X_{i-1}-\theta X_{i-1}^{2}\right) .
$$

By Theorem 2.1, $R_{n}$ satisfies the MDP. Before identifying its rate function let us first show that $r_{n}-R_{n}$ is negligible w.r.t. the MDP. To that end, note

$$
r_{n}=\frac{\sqrt{n}}{b_{n}} \frac{\sum_{i=1}^{n}\left(X_{i} X_{i-1}-\theta X_{i-1}^{2}\right)}{\sum_{i=1}^{n} X_{i-1}^{2}}=\frac{1}{\sqrt{n} b_{n}} \sum_{i=1}^{n}\left(X_{i} X_{i-1}-\theta X_{i-1}^{2}\right) \times \frac{n}{\sum_{i=1}^{n} X_{i-1}^{2}} .
$$

So

$$
r_{n}-R_{n}=\frac{1}{\sqrt{n} b_{n}} \sum_{i=1}^{n}\left(X_{i} X_{i-1}-\theta X_{i-1}^{2}\right) \times \frac{\frac{1}{n} \sum_{i=1}^{n} X_{i-1}^{2}-\left(1-\theta^{2}\right)^{-1}}{\frac{1}{n} \sum_{i=1}^{n} X_{i-1}^{2}} \times\left(1-\theta^{2}\right) .
$$

For $\epsilon>0, L>0$ and $\delta>0$, we have

$$
\begin{aligned}
\mathbb{P}\left(\left|r_{n}-R_{n}\right| \geqslant \epsilon\right)= & \mathbb{P}\left(\left|\frac{\sqrt{n}}{b_{n}} \sum_{i=1}^{n}\left(X_{i} X_{i-1}-\theta X_{i-1}^{2}\right)\right| \geqslant \frac{1}{1-\theta^{2}} L \sqrt{\delta \epsilon}\right) \\
& +\mathbb{P}\left(\left|\frac{1}{n} \sum_{i=1}^{n} X_{i-1}^{2}-\frac{1}{1-\theta^{2}}\right| \geqslant \frac{\sqrt{\delta \epsilon}}{L}\right)+\mathbb{P}\left(\frac{1}{n} \sum_{i=1}^{n} X_{i-1}^{2}<\delta\right) .
\end{aligned}
$$

For $\delta, \varepsilon$ sufficiently small but fixed, the two last terms are bounded by (for $n$ large enough)

$$
\mathbb{P}\left(\left|\frac{1}{\sqrt{n} b_{n}} \sum_{i=1}^{n}\left(X_{i-1}^{2}-\frac{1}{1-\theta^{2}}\right)\right| \geqslant \frac{\sqrt{n}}{b_{n}}\right)
$$

which is clearly negligible as $n \rightarrow+\infty$ by the MDP of

$$
\frac{1}{\sqrt{n} b_{n}} \sum_{i=1}^{n}\left(X_{i-1}^{2}-\frac{1}{1-\theta^{2}}\right) .
$$

The first one is negligible by the MDP of $R_{n}$ by letting $L \rightarrow \infty$. So $r_{n}$ satisfies the same MDP as $R_{n}$. It remains to identify the rate function governing the MDP of $R_{n}$. By Theorem 2.1, the rate function governing the MDP of $R_{n}$ is given by

$$
\begin{equation*}
I(x)=\frac{x^{2}}{2\left(1-\theta^{2}\right)^{2} A^{2}} \tag{5.20}
\end{equation*}
$$

with $A^{2}:=\theta^{2} \Sigma_{00}^{2}-2 \theta \Sigma_{01}^{2}+\Sigma_{11}^{2}$, where

$$
\begin{aligned}
& \Sigma_{00}^{2}=\frac{1}{2 \pi} \int_{\mathbb{T}} 2 f^{2}(u) \mathrm{d} u+\kappa_{4}\left(\frac{1}{2 \pi} \int_{\mathbb{T}} f(u) \mathrm{d} u\right)^{2}, \\
& \Sigma_{01}^{2}=\frac{1}{2 \pi} \int_{\mathbb{T}} 2 \cos (u) f^{2}(u) \mathrm{d} u+\kappa_{4}\left(\frac{1}{2 \pi} \int_{\mathbb{T}} f(u) \mathrm{d} u\right)\left(\frac{1}{2 \pi} \int_{\mathbb{T}} f(u) \cos (u) \mathrm{d} u\right), \\
& \Sigma_{11}^{2}=\frac{1}{2 \pi} \int_{\mathbb{T}}(1+\cos 2 u) f^{2}(u) \mathrm{d} u+\kappa_{4}\left(\frac{1}{2 \pi} \int_{\mathbb{T}} f(u) \cos (u) \mathrm{d} u\right)^{2} .
\end{aligned}
$$

Hence

$$
A^{2}=2 \theta^{2} \hat{r}_{0}\left(f^{2}\right)-2 \theta \hat{r}_{1}\left(f^{2}\right)+\hat{r}_{0}\left(f^{2}\right)+\hat{r}_{2}\left(f^{2}\right)+\kappa_{4}\left(\frac{\theta}{2 \pi} \int_{\mathbb{T}} f(u) \mathrm{d} u-\frac{1}{2 \pi} \int_{\mathbb{T}} f(u) \cos u \mathrm{~d} u\right)^{2}
$$

where $\hat{r}_{k}\left(f^{2}\right)=\frac{1}{2 \pi} \int_{\mathbb{T}} f^{2}(u) \mathrm{e}^{-\mathrm{i} k u} \mathrm{~d} u=\frac{1}{2 \pi} \int_{\mathbb{T}} f^{2}(u) \cos (k u) \mathrm{d} u$ is the $k$ th Fourier coefficient. The last term with coefficient $\kappa_{4}$ is zero for

$$
\frac{\theta}{2 \pi} \int_{\mathbb{T}} f(u) \mathrm{d} u-\frac{1}{2 \pi} \int_{\mathbb{T}} f(u) \cos u \mathrm{~d} u=\theta \operatorname{Var}\left(X_{0}\right)-\operatorname{Cov}\left(X_{0}, X_{1}\right)=0 .
$$

Furthermore (recalling that $\sigma=1$ and $\mathbb{E} X_{0} X_{n}=\theta^{n}\left(1-\theta^{2}\right)^{-1}$ for $n \geqslant 0$ ),

$$
f^{2}(u)=\frac{1}{\left(1-\theta^{2}\right)^{2}}\left(\sum_{n \in \mathbb{Z}} \theta^{|n|} \mathrm{e}^{\mathrm{i} n u}\right)^{2}=\frac{1}{\left(1-\theta^{2}\right)^{2}} \sum_{k \in \mathbb{Z}} \mathrm{e}^{\mathrm{i} k u} \sum_{n \in \mathbb{Z}} \theta^{|n|+|k-n|}
$$

where it follows that $\hat{r}_{k}\left(f^{2}\right)=\frac{1}{\left(1-\theta^{2}\right)^{2}} \sum_{n \in \mathbb{Z}} \theta^{|n|+|k-n|}$. From this last relation we deduce easily

$$
\hat{r}_{0}\left(f^{2}\right)=\frac{1+\theta^{2}}{\left(1-\theta^{2}\right)^{3}}, \quad \hat{r}_{1}\left(f^{2}\right)=\frac{2 \theta}{\left(1-\theta^{2}\right)^{3}}, \quad \hat{r}_{2}\left(f^{2}\right)=\frac{3 \theta^{2}-\theta^{4}}{\left(1-\theta^{2}\right)^{3}} .
$$

Substituting to the expression of $A^{2}$, we get

$$
A^{2}=\frac{1}{1-\theta^{2}} .
$$

Substituting in (5.20), we obtain the claimed rate function.

### 5.6.2. Proof of Proposition 3.2

We need the following stronger result in the centered Gaussian case, inspired by [2]
Lemma 5.2. Assume that $\left(\xi_{i}\right)$ are Gaussian $\mathcal{N}(0,1)$. Let $X^{(n)}=\left(X_{1}, \ldots, X_{n}\right)^{*}$ and $M_{n}$ be a $n \times n$ order symmetric matrix. Denote by $\left(\lambda_{j}^{n}\right)_{1 \leqslant j \leqslant n}$ the eigenvalues (counting up to the multiplicity) of $M_{n} T_{n}(f)$. Assume that $\sup _{n} \max _{j}\left|\lambda_{j}^{n}\right|<+\infty$ and for some measurable function $m$ on $\mathbb{T}$ such that fm $\in L^{\infty}(\mathbb{T})$,

$$
\begin{equation*}
\frac{1}{n} \sum_{j=1}^{n}\left(\lambda_{j}^{n}\right)^{2} \rightarrow \frac{1}{2 \pi} \int_{\mathbb{T}}(f(\theta) m(\theta))^{2} \mathrm{~d} \theta \tag{5.21}
\end{equation*}
$$

Then for every moderate deviation scale $1 \ll b_{n} \ll \sqrt{n}$, we have for all $\lambda$,

$$
\lim _{n \rightarrow+\infty} \frac{1}{b_{n}^{2}} \log \mathbb{E} \exp \left(\frac{\lambda b_{n}}{\sqrt{n}}\left(\left\langle X^{(n)}, M_{n} X^{(n)}\right\rangle-\mathbb{E}\left\langle X^{(n)}, M_{n} X^{(n)}\right\rangle\right)\right)=\frac{\lambda^{2}}{2 \pi} \int_{\mathbb{T}} f^{2}(\theta) m^{2}(\theta) \mathrm{d} \theta .
$$

Proof. Denote $T_{n}=\frac{1}{\sqrt{n} b_{n}}\left(\left\langle X^{(n)}, M_{n} X^{(n)}\right\rangle-\mathbb{E}\left\langle X^{(n)}, M_{n} X^{(n)}\right\rangle\right)$ we have

$$
\log \mathbb{E}\left(\mathrm{e}^{b_{n}^{2} \lambda T_{n}}\right)=-\lambda \frac{b_{n}}{\sqrt{n}} \mathbb{E}\left\langle X^{(n)}, M_{n} X^{(n)}\right\rangle-\frac{1}{2} \sum_{j=1}^{n} \log \left(1-2 \lambda \frac{b_{n}}{\sqrt{n}} \lambda_{j}^{n}\right) .
$$

Notice that by Taylor's Theorem for $|z|<1$

$$
\log (1-z)=-z-\frac{1}{2} z^{2}(1-t z)^{-2},
$$

where $t=t(z) \in[0,1]$. This applied here to $z_{j}^{n}=2 \lambda\left(b_{n} / \sqrt{n}\right) \lambda_{j}^{n}$, which satisfies $\sup _{1 \leqslant j \leqslant n}\left|z_{j}^{n}\right| \rightarrow 0$ as $n \rightarrow \infty$, and hence $\left|1-t\left(z_{j}^{n}\right) z_{j}^{n}\right| \rightarrow 1$ uniformly in $1 \leqslant j \leqslant n$. Since the Gaussian process $\left(X_{k}\right)$ is assumed centered we have

$$
\sum_{j=1}^{n} \lambda_{j}^{n}=\operatorname{tr}\left(M_{n} T_{n}(f)\right)=\mathbb{E}\left\langle X^{(n)}, M_{n} X^{(n)}\right\rangle .
$$

Thus

$$
\lim _{n \rightarrow \infty} \frac{1}{b_{n}^{2}} \log \mathbb{E}\left(\mathrm{e}^{b_{n}^{2} \lambda T_{n}}\right)=\lambda^{2} \lim _{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^{n}\left(\lambda_{j}^{n}\right)^{2} .
$$

The conclusion follows by our hypothesis.
Proof of Proposition 3.2. We have

$$
\frac{\sqrt{n}}{b_{n}}\left(L_{n}-\mathbb{E}\left(L_{n}\right)\right)=\frac{1}{2 \sqrt{n} b_{n}}\left(\left\langle X^{(n)},\left[T_{n}\left(f_{0}\right)^{-1}-T_{n}\left(f_{1}\right)^{-1}\right] X^{(n)}\right\rangle-\mathbb{E}\left(\left\langle X^{(n)}\left[T_{n}\left(f_{0}\right)^{-1}-T_{n}\left(f_{1}\right)^{-1}\right] X^{(n)}\right\rangle\right)\right) .
$$

We want to apply Lemma 5.2 with $f=f_{0}, M_{n}=\frac{1}{2}\left(T_{n}\left(f_{0}\right)^{-1}-T_{n}\left(f_{1}\right)^{-1}\right)$ and $m=\frac{1}{2}\left(f_{0}^{-1}-f^{-1}\right)$. The boundness of the eigenvalues ( $\lambda_{j}^{n}$ ) of $M_{n} T_{n}\left(f_{0}\right), n \geqslant 1$ is given by Lemma 10 in [2], and it is proved in [2, Proof of Proposition 7] that $(1 / n) \sum_{j=1}^{n} \delta_{\lambda_{j}^{n}}$ converges weakly to the image measure of the normalized Lebesgue measure $\mathrm{d} t /(2 \pi)$ by $f m=$ $\frac{1}{2}\left(1-f_{0} / f_{1}\right)$ (the factor $1 / 2$ is missed in [2]). Thus condition (5.21) is satisfied for $\left(\lambda_{j}^{n}\right)_{n, j}$ is bounded. Now the desired MDP follows by Lemma 5.2 and Ellis-Gärtner's theorem.

## Acknowledgement

We are grateful to M. Bouaziz and W. Bryc for fruitful discussions and valuable comments. We also thank the referees whose remarks and comments lead to significant improvement both in the presentation of the paper and in statistical applications.

## References

[1] F. Avram, On bilinear forms in Gaussian random variables and Toeplitz matrices, Probab. Theory Related Fields 79 (1988) $37-45$.
[2] B. Bercu, F. Gamboa, A. Rouault, Large deviations for quadratic forms of Gaussian stationary processes, Stochastic Process. Appl. 71 (1997) 75-90.
[3] S. Bobkov, F. Götze, Exponential integrability and transportation cost related to logarithmic Sobolev inequalities, J. Funct. Anal. 163 (1999) 1-28.
[4] P.J. Brockwell, R.A. Davis, Time Series: Theory and Methods, Springer-Verlag, New York, 1991.
[5] W. Bryc, A. Dembo, On large deviations of empirical measures for stationary Gaussian processes, Stochastic Process. Appl. 58 (1995) 23-34.
[6] W. Bryc, A. Dembo, Large deviations for quadratic functionals of Gaussian functionals, J. Theoret. Probab. 10 (1997) $307-332$.
[7] R.M. Burton, H. Dehling, Large deviations for some weakly dependent random processes, Statist. Probab. Lett. 9 (1990) $397-401$.
[8] P.L. Butzer, R.J. Nessel, Fourier Analysis and Approximation, vol. I, Birkhäuser, 1971.
[9] X. Chen, Moderate deviations for $m$-dependent random variables with Banach space values, Statist. Probab. Lett. 35 (1997) $123-134$.
[10] A. Dembo, O. Zeitouni, Large Deviations Techniques and their Applications, Jones and Bartlett, Boston, MA, 1993.
[11] J.D. Deuschel, D.W. Stroock, Large Deviations, Academic Press, Boston, 1989.
[12] H. Djellout, A. Guillin, Large deviations and moderate deviations for moving average processes, Ann. Math. Fac. Toulouse 10 (2001) $23-31$.
[13] H. Djellout, A. Guillin, L. Wu, Transportation cost-information inequalities for random dynamical systems and diffusions, Ann. Probab. 32 (2004) 2702-2732.
[14] M.D. Donsker, S.R.S. Varadhan, Large deviations for stationary Gaussian processes, Commun. Math. Phys. 97 (1985) 187-210.
[15] R. Fox, M. Taqqu, Central limit theorems for quadratic forms in random variables having long-range dependence, Probab. Theory Related Fields 74 (1987) 213-240.
[16] L. Giraitis, D. Surgailis, A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotical normality of Whittle's estimate, Probab. Theory Related Fields 86 (1990) 87-104.
[17] P. Hall, C.C. Heyde, Martingale Limit Theory and its Application, Academic Press, New York, 1980.
[18] T. Jiang, M.B. Rao, X. Wang, Moderate deviations for some weakly dependent random processes, Statist. Probab. Lett. 15 (1992) $71-76$.
[19] T. Jiang, M.B. Rao, X. Wang, Large deviations for moving average processes, Stochastic Process. Appl. 59 (1995) 309-320.
[20] M. Ledoux, Concentration of measure and logarithmic Sobolev inequalities, in: Séminaire de probabilités XXXIII, in: Lecture Notes in Math., vol. 1709, Springer, 1999, pp. 120-216.
[21] M. Rosenblatt, Gaussian and Non-Gaussian Linear Time Series and Random Fields, Springer-Verlag, New York, 2000.
[22] G. Samorodnitsky, M.S. Taqqu, Stable Non-Gaussian Random Processes, Chapman and Hall, New York, 1994.
[23] L. Wu, An introduction to large deviations, in: J.A. Yan, S. Peng, S. Fang, L. Wu (Eds.), Several Topics in Stochastic Analysis, Academic Press of China, Beijing, 1997, pp. 225-336 (in Chinese).
[24] L. Wu, On large deviations for moving average processes, in: T.L. Lai, H.L. Yang, S.P. Yung (Eds.), Probability, Finance and Insurance, Proceeding of a Workshop at the University of Hong-Kong, 15-17 July 2002, World Scientific, Singapore, 2004, pp. 15-49.


[^0]:    * Corresponding author.

    E-mail addresses: djellout@math.univ-bpclermont.fr (H. Djellout), guillin@ceremade.dauphine.fr (A. Guillin), Li-Ming.Wu@math.univ-bpclermont.fr (L. Wu).

