Probability, Uncertainty and Quantitative Risk
Vol. 8, No. 2, 2023, 177–212

https://www.aimsciences.org/journal/2095-9672
DOI:10.3934/puqr.2023008

Explicit bivariate rate functions for large deviations in AR(1) and MA(1) processes with Gaussian innovations

Maicon J. Karling, Artur O. Lopes, Sílvia R. C. Lopes*

Mathematics and Statistics Institute, Federal University of Rio Grande do Sul,
91500-900 *Porto Alegre, RS, Brazil*

Email: maiconkarling@gmail.com, arturoscar.lopes@gmail.com, silviarc.lopes@gmail.com

Abstract We investigate the large deviations properties for centered stationary AR(1) and MA(1) processes with independent Gaussian innovations, by giving the explicit bivariate rate functions for the sequence of two-dimensional random vectors $(\boldsymbol{S}_n)_{n\in\mathbb{N}} = \big(n^{-1}(\sum_{k=1}^n X_k, \sum_{k=1}^n X_k^2)\big)_{n\in\mathbb{N}}$. Via the Contraction Principle, we provide the explicit rate functions for the sample mean and the sample second moment. In the AR(1) case, we also give the explicit rate function for the sequence of two-dimensional random vectors $(\boldsymbol{\mathcal{W}}_n)_{n\geqslant 2} = \big(n^{-1}(\sum_{k=1}^n X_k^2, \sum_{k=2}^n X_k X_{k-1})\big)_{n\geqslant 2}$, but we obtain an analytic rate function that gives different values for the upper and lower bounds, depending on the evaluated set and its intersection with the respective set of exposed points. A careful analysis of the properties of a certain family of Toeplitz matrices is necessary. The large deviations properties of three particular sequences of one-dimensional random variables will follow after we show how to apply a weaker version of the Contraction Principle for our setting, providing new proofs for two already known results on the explicit deviation function for the sample second moment and Yule-Walker estimators. We exhibit the properties of the large deviations of the first-order empirical autocovariance, its explicit deviation function and this is also a new result.

Keywords Autoregressive processes, Empirical autocovariance, Large deviations, Moving average processes, Sample moments, Toeplitz matrices, Yule-Walker estimator

2020 Mathematics Subject Classification 11E25, 60F10, 60G10, 60G15, 62F12, 62M10

1. Introduction

Since the first establishment of the Large Deviations theory, there has been a great expansion of the number of surveys on Large Deviations Principles (LDP). Nowadays, a variety of examples applied to the time series analysis and stochastic processes are available; for instance, LDPs for stable laws (see, e.g., Heyde [25], Rozovskii [38, 39], and Zaigraev [43]), stationary Gaussian processes (see, e.g., Bercu *et al.* [4, 5], Bryc and Dembo [10], Donsker and Varadhan [18], and Zani [44]), autoregressive and moving average processes (see, e.g., Bercu [3], Bryc and

Received 20 December 2021; Accepted 21 December 2022; Early access 14 February 2023
*Corresponding author

Smolenski [11], Burton and Dehling [13], Djellout and Guillin [17], Macci and Trapani [30], Mas and Menneteau [32], Miao [34], and Wu [42]) and continuous processes (see, e.g., Bercu and Richou [6, 7]).

When considering the *empirical autocovariance function*

$$\tilde{\gamma}_n(h) = \frac{1}{n} \sum_{k=h+1}^{n} X_k X_{k-h}, \quad \text{for } 0 \leqslant h < n,$$

of a centered process $(X_j)_{j\in\mathbb{N}}$ at 0, based on a sample of size $n \in \mathbb{N}$, few results on LDP are known. Regarding Gaussian distributions, one of the first studies in the literature is the one from Bryc and Smolenski [11], concerning the LDP for the *sample second moment*

$$\tilde{\gamma}_n(0) = \frac{1}{n} \sum_{k=1}^{n} X_k^2. \tag{1.1}$$

Bryc and Dembo [10] showed that, for a fixed lag $l \geqslant 1$, the LDP for the vector $(\tilde{\gamma}_n(0), \ldots, \tilde{\gamma}_n(l))$ is available when $(X_j)_{j\in\mathbb{N}}$ is an independent and identically distributed (i.i.d.) process, with $X_j \sim \mathcal{N}(0,1)$. It is well known that most of the relevant stochastic processes are not independent and, as the authors have claimed, their approach needs some adjustments when trying to show that a similar LDP works, for instance, when dealing with the classical centered stationary Gaussian AR(1) process (Bryc and Dembo [10], example 1).

At the same time, Bercu *et al.* [5] proved the LDP for Toeplitz quadratic forms of centered stationary Gaussian processes in a univariate setting. Their survey eliminated the need for the variables of $(X_j)_{j\in\mathbb{N}}$ to be independent, extending the result in Bryc and Dembo [10] by including the AR(1) process. However, it is not clear if the LDP is available even for the bivariate random vector $(\tilde{\gamma}_n(0), \tilde{\gamma}_n(1))$, once it has only been proved for each one of the components separately. More precisely, the results in Bercu *et al.* [5] only cover the LDP of the sequence of random variables

$$W_n = \frac{1}{n} X^{(n)^*} M_n X^{(n)},$$

where $X^{(n)} = (X_1, \ldots, X_n)$ is the column vector with components $X_1, \ldots, X_n$, $(M_n)_{n\in\mathbb{N}}$ is a sequence of $n \times n$ Hermitian matrices, and $X^{(n)^*}$ denotes the conjugate transpose of $X^{(n)}$.

In a more general setting, Carmona *et al.* [14] present a level-1 LDP for the empirical autocovariance function of order h for any innovation processes, that encompasses the AR(d) process with Gaussian innovations. In this paper, the authors used the level-2 LDP together with the Contraction Principle. The process itself is obtained from iterations of a continuous uniquely ergodic transformation, preserving the Lebesgue measure on the circle. In Carmona and Lopes [15], the authors considered a similar problem where the dynamics are given by an expanding transformation on the circle. In the same line of research, Wu [42] proved the LDP for $(\tilde{\gamma}_n(0), \ldots, \tilde{\gamma}_n(l))$ under the assumption that $\mathbb{E}(\exp(\lambda \varepsilon_n^2))$ is finite, for $\lambda > 0$, where $(\varepsilon_j)_{j\in\mathbb{N}}$ is the white noise of an AR(d) process, excluding in turn the Gaussian case.

In this work, we shall consider $(X_j)_{j\in\mathbb{N}}$ as the *centered stationary Gaussian AR(1) process* defined by the equation

$$X_{j+1} = \phi X_j + \varepsilon_{j+1}, \quad \text{for } |\phi| < 1 \text{ and } j \in \mathbb{N}, \tag{1.2}$$

with the additional assumptions that:

- $(\varepsilon_j)_{j\geqslant 2}$ is a sequence of i.i.d. random variables, with $\varepsilon_j \sim \mathcal{N}(0,1)$, for $j \geqslant 2$;
- X_k is independent of $(\varepsilon_j)_{j\geqslant k+1}$, for any $k \in \mathbb{N}$;
- $X_1 \sim \mathcal{N}(0, 1/(1-\phi^2))$.

Thus, $(X_j)_{j\in\mathbb{N}}$ has a (positive) spectral density function given by (see Brockwell and Davis [9])

$$g_\phi(\omega) = \frac{1}{1+\phi^2 - 2\phi\cos(\omega)}, \quad \omega \in \mathbb{T} = [-\pi, \pi). \tag{1.3}$$

Our main goal is to extend the results of Bercu *et al.* [5] and Bryc and Dembo [10] in the bivariate case. More specifically, we investigate the large deviations properties associated with $(\boldsymbol{\mathcal{W}}_n)_{n\geqslant 2}$, where

$$\boldsymbol{\mathcal{W}}_n = (\tilde{\gamma}_n(0), \tilde{\gamma}_n(1)) = \frac{1}{n}\left(\sum_{k=1}^{n} X_k^2, \sum_{k=2}^{n} X_k X_{k-1}\right), \quad \text{for } n \geqslant 2, \tag{1.4}$$

and its explicit bivariate rate function is presented. The asymptotical behavior of $(\boldsymbol{\mathcal{W}}_n)_{n\geqslant 2}$ is well known, that is

$$\boldsymbol{\mathcal{W}}_n \xrightarrow{n\to+\infty} \left(\frac{1}{1-\phi^2}, \frac{\phi}{1-\phi^2}\right), \quad \text{almost surely}.$$

By definition of almost sure convergence, as $n \to +\infty$, the sequence of probabilities

$$\mathbb{P}\left(\left\| \boldsymbol{\mathcal{W}}_n - \left(\frac{1}{1-\phi^2}, \frac{\phi}{1-\phi^2}\right)\right\| > \delta \right) \tag{1.5}$$

converges to zero, for all $\delta > 0$. However, if this convergence is very slow, even for large n, we have a certainly reasonable chance of choosing a bad sample $X_1, \ldots, X_n$ from $(X_j)_{j\in\mathbb{N}}$, such that $\boldsymbol{\mathcal{W}}_n$ is distant from the true value $\left(\frac{1}{1-\phi^2}, \frac{\phi}{1-\phi^2}\right)$.

The Large Deviations theory considers the asymptotic behavior of the probabilities presented in (1.5), ensuring that they converge to zero approximately at an exponential rate (see Bucklew [12], chapter 1). The use of LDP is also a natural tool for the study of statistical properties related to risk (see, for instance, Ferreira *et al.* [21]). A classical definition of the Large Deviation Principle is given as follows (see Dembo and Zeitouni [16]).

Extending LDP properties from the univariate to the bivariate case is not simple. We believe this will be transparent to the reader in what will be done next.

Definition 1.1 *A sequence of random vectors $(\boldsymbol{V}_n)_{n\in\mathbb{N}}$ in $\mathbb{R}^d$, for $d \in \mathbb{N}$, satisfies a Large Deviation Principle (LDP) with speed n and rate function $J(\cdot)$, if $J(\cdot) : \mathbb{R}^d \to [0, +\infty]$ is a lower semi-continuous function such that,*

- *Upper bound: for any closed set $F \subset \mathbb{R}^d$,*

$$\limsup_{n\to+\infty} \frac{1}{n} \log \mathbb{P}\left(\boldsymbol{V}_n \in F\right) \leqslant -\inf_{\boldsymbol{x}\in F} J(\boldsymbol{x}); \tag{1.6}$$

- *Lower bound: for any open set $G \subset \mathbb{R}^d$,*

$$-\inf_{\boldsymbol{x}\in G} J(\boldsymbol{x}) \leqslant \liminf_{n\to+\infty} \frac{1}{n} \log \mathbb{P}\left(\boldsymbol{V}_n \in G\right). \tag{1.7}$$

Moreover, $J(\cdot)$ is said to be a good rate function if its level sets $J^{-1}([0,b])$ are compact, for all $b \in \mathbb{R}$.

Remark 1.1 (1) *In this work, we only deal with good rate functions, but in short, we sometimes write rate functions instead.*

(2) *The infimum over the empty set is taken to be* $+\infty$*, while the supremum is taken to be* $-\infty$*, the latter being of particular interest when the Fenchel-Legendre transform is considered in Section 2.2.*

(3) *In our setting, the large deviation version as described by Definition 1.1, for the deviation rate function* $J(\cdot)$*, is not suitable. In the reasoning that we follow to obtain our main result (see Theorem 2.1), we will be able to prove a version of the large deviation which is different from the one described in Definition 1.1. Among other things, we present an explicit expression for the rate function that depends on two functions* $J_1(\cdot)$*,* $J_2(\cdot)$*, given in (2.17) and (2.18), respectively (see Proposition 2.2). Besides that, a detailed analysis of when a certain family of Toeplitz matrices is positive definite will be necessary (see Section 2.1). We will also have to consider in Section 2.3 the concepts of steep function, exposed points, exposed hyperplane, and the Gärtner-Ellis' theorem (Dembo and Zeitouni [16], theorem 2.3.6). We show that there are sets such that the upper bound is different from the lower bound in the LDP estimate (see Example 2.2). More precisely, while the upper bound holds as in (1.6), the lower bound is given by the infimum over the intersection of any given open set* G *under consideration and the set of exposed points defined by (2.23), thus generating different upper and lower bounds in some cases.*

In general, it is not easy to prove that an arbitrary sequence of random vectors satisfies the LDP stated in Definition 1.1 (see, e.g., Bercu and Richou [7], Bryc and Dembo [10], Dembo and Zeitouni [16], Ellis [19], Macci and Trapani [30], and Mas and Menneteau [32]). An elegant way of proving such a principle is to verify the validity of the Gärtner-Ellis' theorem's conditions, as little use of the dependency structure is made and the focus mainly rests on the behavior of the *limiting cumulant generating function*, defined by

$$L(\boldsymbol{\lambda}) = \lim_{n\to+\infty} L_n(\boldsymbol{\lambda}), \quad \text{for all} \quad \boldsymbol{\lambda} \in \mathbb{R}^2, \tag{1.8}$$

where $L_n(\cdot): \mathbb{R}^2 \to \mathbb{R} \cup \{+\infty\}$ denotes the *normalized cumulant generating function* of $\boldsymbol{\mathcal{W}}_n$,

$$L_n(\boldsymbol{\lambda}) = \frac{1}{n} \log \mathbb{E}\left[\exp\left(n\langle \boldsymbol{\lambda}, \boldsymbol{\mathcal{W}}_n\rangle\right)\right].$$

We shall present an explicit expression for $L_n(\cdot)$ and $L(\cdot)$ in the case $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$ depends on two variables (the bivariate case). Subsequently, we obtain the explicit rate function through the Fenchel-Legendre transform of $L(\cdot)$ in Section 2.

As a second and distinguished part of our study, we analyze the LDP of the sequence of bivariate random vectors $(\boldsymbol{S}_n)_{n\in\mathbb{N}}$, where

$$\boldsymbol{S}_n = \frac{1}{n}\left(\sum_{k=1}^n X_k,\ \sum_{k=1}^n X_k^2\right).$$

We call $\boldsymbol{S}_n$ the S^2*-mean* since its first and second components are the sample mean and the sample second moment. We dedicate our efforts to the particular cases when $(X_j)_{j\in\mathbb{N}}$ is an AR(1) or an MA(1) process. This study is based on a particular result presented in Bryc and Dembo [10] and which has a very interesting application when the Contraction Principle can be applied.

Our study is organized as follows. Section 2 is dedicated to the analysis of the large deviations properties and computations of the explicit rate function for the random sequence $(\boldsymbol{\mathcal{W}}_n)_{n\geqslant 2}$, under the assumption that $(X_j)_{j\in\mathbb{N}}$ follows a centered stationary AR(1) process with Gaussian innovations. In Section 3, we obtain the rate functions for some particular cases, namely, the sample second moment, the first-order empirical autocovariance, and the Yule-Walker estimator of this process. As an independent analysis of the studies in Sections 2 and 3, we dedicate Section 4 to show that the LDP for the S^2-mean of the AR(1) process is available. Next, we give the details of the LDP for the sample second moment of an MA(1) process and, as a consequence, the LDP for its S^2-mean. The proofs of some of our main results are given in Section 5. Section 6 concludes the manuscript.

2. LDP and the centered stationary Gaussian AR(1) process

The purpose of this section is to study the large deviation properties of the sequence $(\boldsymbol{\mathcal{W}}_n)_{n\geqslant 2}$ introduced in (1.4). First, we explicitly express this sequence's normalized cumulant generating function. Next, we compute the Fenchel-Legendre transform of the limiting cumulant generating function. Finally, as an extension of this reasoning, we recall the Gärtner-Ellis's theorem and show how it can be used to give a lower and upper bound for the sequence of probabilities in (1.5), with the rate function given by the Fenchel-Legendre transform of the limiting cumulant generating function associated to $(\boldsymbol{\mathcal{W}}_n)_{n\geqslant 2}$.

2.1 Analysis of the normalized cumulant generating function

Consider $\boldsymbol{\lambda} = (\lambda_1, \lambda_2) \in \mathbb{R}^2$. Let $L_n(\cdot,\cdot) : \mathbb{R}^2 \to \mathbb{R}$ represent the normalized cumulant generating function associated to the sequence $(\boldsymbol{\mathcal{W}}_n)_{n\geqslant 2}$, given by

$$L_n(\lambda_1, \lambda_2) = \frac{1}{n} \log \mathbb{E}\left(e^{n\langle(\lambda_1,\lambda_2),\boldsymbol{\mathcal{W}}_n\rangle}\right), \quad \text{for } n \geqslant 2, \tag{2.1}$$

where $\langle(x_1, y_1), (x_2, y_2)\rangle := x_1x_2 + y_1y_2$ denotes the usual inner product in $\mathbb{R}^2$. If $\boldsymbol{X}_n := (X_1, \ldots, X_n) \in \mathbb{R}^n$ and $\boldsymbol{X}_n^\top \in \mathbb{R}^{1\times n}$ denotes the transpose of $\boldsymbol{X}_n$, then one can rewrite (1.4) as

$$\boldsymbol{\mathcal{W}}_n = \frac{1}{n}\left(\boldsymbol{X}_n^\top T_n(\varphi_1)\, \boldsymbol{X}_n,\ \boldsymbol{X}_n^\top T_n(\varphi_2)\, \boldsymbol{X}_n\right), \tag{2.2}$$

where $\varphi_1 : \mathbb{T} \to \{1\}$ and $\varphi_2 : \mathbb{T} \to [-1, 1]$ are the real-valued functions given, respectively, by

$$\varphi_1(\omega) = 1 \qquad \text{and} \qquad \varphi_2(\omega) = \cos(\omega).$$

Here $T_n(\varphi)$ represents the Toeplitz matrix associated to a function $\varphi : \mathbb{T} \to \mathbb{R}$, defined by

$$T_n(\varphi) = \left[\frac{1}{2\pi}\int_{\mathbb{T}} e^{i\,(j-k)\omega}\varphi(\omega)\ d\omega\right]_{1\leqslant j,k\leqslant n}.$$

Remark 2.1 *A vast literature comprehending Toeplitz matrices has emerged in the last century and one of the most famous and referenced works is given by Grenander and Szegö [24]. A modern treatment of this subject can be found in the works by Gray [23] and Nikolski [35].*

Inserting (2.2) into (2.1) and using the linearity property of Toeplitz matrices we get

$$L_n(\lambda_1, \lambda_2) = \frac{1}{n} \log \mathbb{E}\left(e^{\boldsymbol{X}_n^\top T_n(\varphi_{\boldsymbol{\lambda}})\, \boldsymbol{X}_n}\right), \tag{2.3}$$

where $\varphi_{\boldsymbol{\lambda}} : \mathbb{T} \to \mathbb{R}$ is defined by

$$\varphi_{\boldsymbol{\lambda}}(\omega) = \lambda_1\varphi_1(\omega) + \lambda_2\varphi_2(\omega) = \lambda_1 + \lambda_2 \cos(\omega). \tag{2.4}$$

Observe that $\varphi_{\boldsymbol{\lambda}}(\cdot)$ depends on the choice of $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$ and that

$$T_n(\varphi_{\boldsymbol{\lambda}}) = \frac{1}{2}\begin{bmatrix} 2\lambda_1 & \lambda_2 & 0 & 0 & \cdots & 0 \\ \lambda_2 & 2\lambda_1 & \lambda_2 & 0 & \ddots & \vdots \\ 0 & \lambda_2 & 2\lambda_1 & \lambda_2 & \ddots & 0 \\ 0 & 0 & \lambda_2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 2\lambda_1 & \lambda_2 \\ 0 & \cdots & 0 & 0 & \lambda_2 & 2\lambda_1 \end{bmatrix}.$$

The fact that $\boldsymbol{X}_n$ has a multivariate Gaussian distribution gives us some advantage here. A known result from Probability theory (Bickel and Doksum [8], section B.6) shows that there is always a standard multivariate Gaussian random vector $\boldsymbol{Y}_n = (Y_{n,1}, \dots, Y_{n,n})$ with independent components, such that

$$\boldsymbol{X}_n = T_n(g_\phi)^{1/2}\,\boldsymbol{Y}_n, \tag{2.5}$$

where $g_\phi(\cdot)$ is given in (1.3) and $T_n(g_\phi)^{1/2}$ is the square root matrix of $T_n(g_\phi)$. We also note that $(T_j(g_\phi))_{j\in\mathbb{N}}$ is the sequence of autocovariance matrices associated to the process $(X_j)_{j\in\mathbb{N}}$ and, since $T_j(g_\phi)$ is a positive definite matrix for each $j \in \mathbb{N}$, the sequence of square-root matrices $(T_j(g_\phi)^{1/2})_{j\in\mathbb{N}}$ is well defined. Additionally, from (2.5) we obtain

$$\boldsymbol{X}_n^\top\, T_n(\varphi_{\boldsymbol{\lambda}})\,\boldsymbol{X}_n = \boldsymbol{Y}_n^\top\, T_n(g_\phi)^{1/2}\, T_n(\varphi_{\boldsymbol{\lambda}})\, T_n(g_\phi)^{1/2}\,\boldsymbol{Y}_n. \tag{2.6}$$

Let $(\alpha^{\boldsymbol{\lambda}}_{n,k})_{k=1}^n$, with $\alpha^{\boldsymbol{\lambda}}_{n,1} \leqslant \cdots \leqslant \alpha^{\boldsymbol{\lambda}}_{n,n}$, denote the eigenvalues of $T_n(g_\phi)^{1/2}\, T_n(\varphi_{\boldsymbol{\lambda}})\, T_n(g_\phi)^{1/2}$. Since this matrix is real and symmetric, there exists a sequence of orthogonal matrices $(P_j)_{j\in\mathbb{N}}$ such that

$$T_n(g_\phi)^{1/2}\, T_n(\varphi_{\boldsymbol{\lambda}})\, T_n(g_\phi)^{1/2} = P_n\,\Lambda_n\,P_n^\top, \tag{2.7}$$

with Λ_n denoting the $n \times n$ diagonal matrix $\mathrm{Diag}(\alpha^{\boldsymbol{\lambda}}_{n,1}, \dots, \alpha^{\boldsymbol{\lambda}}_{n,n})$. Therefore, from (2.6) and (2.7) we obtain

$$\boldsymbol{Y}_n^\top\, T_n(g_\phi)^{1/2}\, T_n(\varphi_{\boldsymbol{\lambda}})\, T_n(g_\phi)^{1/2}\,\boldsymbol{Y}_n = \boldsymbol{Y}_n^\top\, P_n\,\Lambda_n\,P_n^\top\,\boldsymbol{Y}_n. \tag{2.8}$$

As P_n is orthogonal, the product $P_n^\top\,\boldsymbol{Y}_n$ has a multivariate Gaussian distribution with independent components. Hence, from (2.6) and (2.8) it follows that

$$\boldsymbol{X}_n^\top\, T_n(\varphi_{\boldsymbol{\lambda}})\,\boldsymbol{X}_n = \sum_{k=1}^n \alpha^{\boldsymbol{\lambda}}_{n,k} Z_{n,k}, \tag{2.9}$$

where $Z_{n,1}, \dots, Z_{n,n}$ are i.i.d. random variables, each one having a χ^2_1 distribution with moment generating function given by

$$M_{Z_{n,k}}(t) = \mathbb{E}(e^{tZ_{n,k}}) = \begin{cases} (1-2t)^{-1/2}, & t < \dfrac{1}{2}, \\ +\infty, & t \geqslant \dfrac{1}{2}, \end{cases} \tag{2.10}$$

for $k = 1, \dots, n$.

Returning to the analysis of (2.3) and considering (2.9), as $Z_{1,n}, \dots, Z_{n,n}$ are mutually independent, we conclude that

$$L_n(\lambda_1, \lambda_2) = \frac{1}{n}\log\mathbb{E}\left(e^{\sum_{k=1}^n \alpha^{\boldsymbol{\lambda}}_{n,k} Z_{n,k}}\right) = \frac{1}{n}\log\left(\prod_{k=1}^n \mathbb{E}\left(e^{\alpha^{\boldsymbol{\lambda}}_{n,k} Z_{n,k}}\right)\right). \tag{2.11}$$

From (2.10), we observe that $\mathbb{E}\left(e^{\alpha^{\boldsymbol{\lambda}}_{n,k} Z_{n,k}}\right)$ is only defined if each one of the $\alpha^{\boldsymbol{\lambda}}_{n,k} < 1/2$. In other words, (2.11) is finite if

$$0 < 1 - 2\alpha^{\boldsymbol{\lambda}}_{n,k}, \quad \text{for all } k \text{ such that } 1 \leqslant k \leqslant n. \tag{2.12}$$

Note that, (2.12) is equivalent to requiring that $I_n - 2\,T_n(g_\phi)^{1/2}\,T_n(\varphi_{\boldsymbol{\lambda}})\,T_n(g_\phi)^{1/2}$ must be positive definite, where I_n represents the $n \times n$ identity matrix. Since $T_n(g_\phi)$ is positive definite and

$$I_n - 2\,T_n(g_\phi)^{1/2}\,T_n(\varphi_{\boldsymbol{\lambda}})\,T_n(g_\phi)^{1/2} = T_n(g_\phi)^{1/2}\,\left(T_n^{-1}(g_\phi) - 2T_n(\varphi_{\boldsymbol{\lambda}})\right)\,T_n(g_\phi)^{1/2},$$

it is sufficient to show that (Horn and Johnson [26], section 7)

$$D_{n,\boldsymbol{\lambda}} = T_n^{-1}(g_\phi) - 2T_n(\varphi_{\boldsymbol{\lambda}}) = \begin{pmatrix} r_1 & q & 0 & \cdots & 0 \\ q & p & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & p & q \\ 0 & \cdots & 0 & q & r_1 \end{pmatrix} \tag{2.13}$$

is positive definite, where $r_1 = 1 - 2\lambda_1$, $p = 1 + \phi^2 - 2\lambda_1$ and $q = -\phi - \lambda_2$. The domain $\mathcal{D} \subseteq \mathbb{R}^2$, where $D_{n,\boldsymbol{\lambda}}$ (and so $I_n - 2\,T_n(g_\phi)^{1/2}\,T_n(\varphi_{\boldsymbol{\lambda}})\,T_n(g_\phi)^{1/2}$) is positive definite, is given in the following lemma.

Lemma 2.1 *Given $|\phi| < 1$, consider $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$, the matrix $D_{n,\boldsymbol{\lambda}}$ in (2.13), and the two sets*

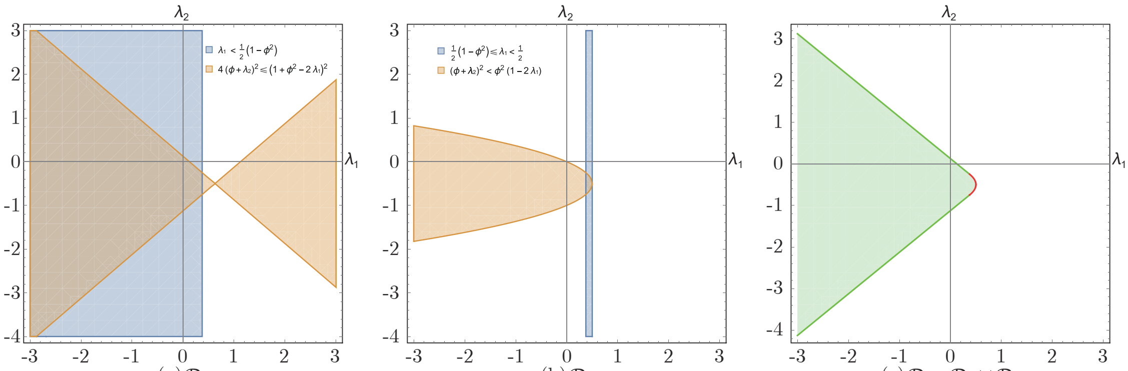
$$\begin{aligned} \mathcal{D}_1 &= \left\{(\lambda_1, \lambda_2) \in \mathbb{R}^2 \,\middle|\, \lambda_1 < \frac{1-\phi^2}{2},\ 4(\phi + \lambda_2)^2 \leqslant (1 + \phi^2 - 2\lambda_1)^2\right\}, \\ \mathcal{D}_2 &= \left\{(\lambda_1, \lambda_2) \in \mathbb{R}^2 \,\middle|\, \frac{1-\phi^2}{2} \leqslant \lambda_1 < \frac{1}{2},\ (\phi + \lambda_2)^2 < \phi^2(1 - 2\lambda_1)\right\}. \end{aligned} \tag{2.14}$$

If $\boldsymbol{\lambda} \in \mathcal{D}_1$, then $D_{n,\boldsymbol{\lambda}}$ is positive definite for all $n \geqslant 2$. Additionally, if $\boldsymbol{\lambda} \in \mathcal{D}_2$, then there exists N (dependent on λ_1 and λ_2) such that $D_{n,\boldsymbol{\lambda}}$ is positive definite for all $n \geqslant N$. Otherwise, $D_{n,\boldsymbol{\lambda}}$ has at least one non-positive eigenvalue for all $n \in \mathbb{N}$, so that it is not positive definite.

Proof See Subsection 5.1. □

To illustrate the domains presented in (2.14) in a particular case, Figure 2.1 shows the graphs of $\mathcal{D}_1$, $\mathcal{D}_2$ and $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$ when $\phi = 0.5$. In the left-hand side panel, the rectangular region represents the points for which $\lambda_1 < \dfrac{1-\phi^2}{2}$, the triangular region represents the points for which $4(\phi + \lambda_2)^2 \leqslant (1 + \phi^2 - 2\lambda_1)^2$, and $\mathcal{D}_1$ is, therefore, the intersection between these two regions; a similar representation for $\mathcal{D}_2$ is given on the middle side panel, where this set is represented by the intersection between the interior of the parabola and the rectangular region. The union $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$ is plotted in the right-hand side panel. Notice that the origin $(0, 0)$ is an interior point of $\mathcal{D}$.

The knowledge of the domain on which the matrix $D_{n,\boldsymbol{\lambda}}$ is positive definite allows us to continue our reasoning in the direction of obtaining the explicit expression of $L_n(\cdot, \cdot)$ and its limiting function when $n \to +\infty$. We note that on the boundary of $\mathcal{D}$, $D_{n,\boldsymbol{\lambda}}$ can be either positive definite or non-negative definite (in the presence of null eigenvalues, which contradicts (2.12)). As the behavior of the limiting cumulant generating function over the boundary of $\mathcal{D}$ does not affect the large deviations properties, we shall not go into too much detail, but to



(a) $\mathcal{D}_1$ (b) $\mathcal{D}_2$ (c) $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$

Figure 2.1 Regions $\mathcal{D}_1$ (left-hand side panel), $\mathcal{D}_2$ (middle side panel) and $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$ (right-hand side panel), for $\mathcal{D}_1$ and $\mathcal{D}_2$ defined in (2.14) in the particular case when $\phi = 0.5$. In the third draw, we plotted the lines $\lambda_2 = -\phi \pm (1 + \phi^2 - 2\lambda_1)/2$, for $\lambda_1 < (1-\phi^2)/2$ in thick green color to illustrate that this border belongs to the set $\mathcal{D}$, whereas the curve $(\phi + \lambda_2)^2 = \phi^2(1 - 2\lambda_1)$, for $\dfrac{1-\phi^2}{2} \leqslant \lambda_1 \leqslant \dfrac{1}{2}$, is plotted in red color to indicate that it does not make part of the set $\mathcal{D}$.

whom it may concern, a similar discussion when the centered stationary Gaussian MA(1) process is under consideration is presented in Karling *et al.* [28].

Even though representing a particular degenerate case, it is important to note the following: if $\phi = 0$, the process $(X_j)_{j\in\mathbb{N}}$ in (1.2) reduces itself to an i.i.d. sequence of random variables with standard Gaussian distribution, and it is shown in Bryc and Dembo [10] (pg. 330) that

$$\lim_{n\to+\infty} L_n(\lambda_1,\lambda_2) = \begin{cases} -\dfrac{1}{2}\log\left(\dfrac{1-2\lambda_1+\sqrt{(1-2\lambda_1)^2-4\lambda_2^2}}{2}\right), & \text{if } \lambda_1 < \dfrac{1}{2} \text{ and } 4\lambda_2^2 \leqslant (1-2\lambda_1)^2, \\ +\infty, & \text{otherwise.} \end{cases}$$

We obtain the following lemma when $\phi \neq 0$, which is a more interesting case.

Lemma 2.2 *Given* $0 < |\phi| < 1$ *, let* $L_n(\cdot,\cdot)$ *denote the normalized cumulant generating function associated to* $(\mathcal{W}_n)_{n\geqslant 2}$ *and* $\mathcal{D}$ *be the set given in Lemma 2.1, then*

$$\lim_{n\to+\infty} L_n(\lambda_1,\lambda_2) = L(\lambda_1,\lambda_2),$$

where $L : \mathbb{R}^2 \to \mathbb{R} \cup \{+\infty\}$ *is defined by*

$$L(\lambda_1,\lambda_2) = \begin{cases} -\dfrac{1}{2}\log(\Phi(\lambda_1,\lambda_2,\phi)), & if\ (\lambda_1,\lambda_2) \in \mathcal{D}, \\ +\infty, & if (\lambda_1,\lambda_2) \in \mathbb{R}^2 \setminus \mathcal{D}, \end{cases} \tag{2.15}$$

with

$$\Phi(\lambda_1,\lambda_2,\phi) = \frac{1+\phi^2-2\lambda_1+\sqrt{(1+\phi^2-2\lambda_1)^2-4(\phi+\lambda_2)^2}}{2}.$$

Proof See Subsection 5.2. □

2.2 Rate function associated to the random sequence $(\mathcal{W}_n)_{n\geqslant 2}$

Given $0 < |\phi| < 1$, consider the curve

$$\mathcal{C}_\phi = \{(x,y) \in \mathbb{R}^2 : |y| < x\,,\ y^2\phi^2 = x(x\phi^2 - 1)\} \tag{2.16}$$

and the sets

$$\mathcal{A}_\phi = \{(x, y) \in \mathbb{R}^2 : |y| < x \,,\ y^2\phi^2 \geqslant x(x\phi^2 - 1)\}$$

and

$$\mathcal{B}_\phi = \{(x, y) \in \mathbb{R}^2 : |y| < x \,,\ y^2\phi^2 \leqslant x(x\phi^2 - 1)\}.$$

Then, $\mathcal{C}_\phi = \mathcal{A}_\phi \cap \mathcal{B}_\phi$.

Let us define

$$J_1(x, y) = \frac{1}{2}\left[x(1+\phi^2) - 1 - 2y\phi + \log\left(\frac{x}{x^2 - y^2}\right)\right], \quad \text{for } (x, y) \in \mathcal{A}_\phi, \tag{2.17}$$

$$\text{and} \qquad J_2(x, y) = \frac{(\phi y - x)^2}{2x} + \log|\phi|, \quad \text{for } (x, y) \in \mathcal{B}_\phi. \tag{2.18}$$

The values of $J_1(x, y)$ converge to $+\infty$ as (x, y) converges to the set $\{(x, y) \in \mathbb{R}^2 : x \geqslant 0, |y| = x\}$, which is the external boundary of $\mathcal{A}_\phi$ (see Figure 2.2). The next proposition expresses three remarkable properties of these two functions, showing that they are convex and that they match continuously and in a differentiable way over the curve $\mathcal{C}_\phi$.

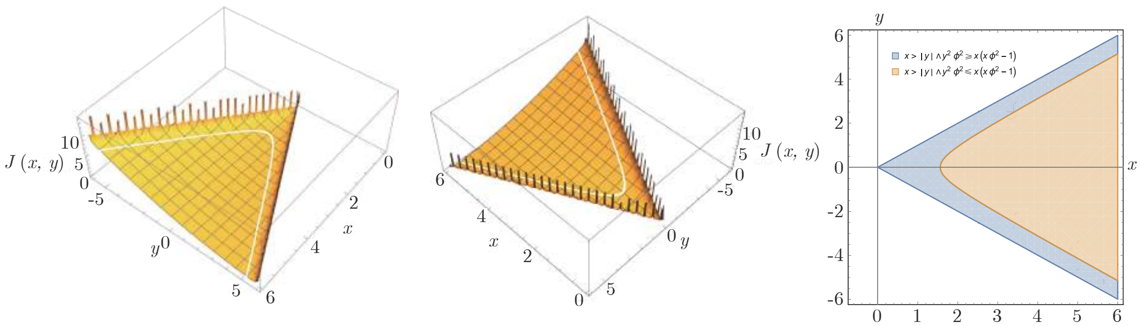


Figure 2.2 The two left-hand side panels show the graph of the function $J(x, y)$, given in (2.19), from two different points of view, with $\phi = 0.8$, $x \in (0, 6]$ and $y \in [-6, 6]$. The white line represents the curve $\mathcal{C}_{0.8}$, defined in (2.16), where the domains of the two functions $J_1(\cdot,\cdot)$ and $J_2(\cdot,\cdot)$ intersect, changing roles in the definition of $J(\cdot,\cdot)$. Notice that, $J_1(\cdot,\cdot)$ converges to $+\infty$ as (x, y) approaches the curves $\{(x, y) \in \mathbb{R}^2 : x \geqslant 0, |y| = x\}$; this behavior is inherited by $J(\cdot,\cdot)$. The right-hand side panel displays the domains $\mathcal{A}_\phi$ (in blue color) and $\mathcal{B}_\phi$ (in orange color).

Proposition 2.1 *Consider the functions* $J_1(\cdot,\cdot)$ *and* $J_2(\cdot,\cdot)$ *given, respectively, by* (2.17) *and* (2.18). *Then,*

(1) $J_1(\cdot,\cdot)$ *and* $J_2(\cdot,\cdot)$ *match in a continuous way at* $\mathcal{C}_\phi$ *;*

(2) *the gradients of* $J_1(\cdot,\cdot)$ *and* $J_2(\cdot,\cdot)$ *match at* $\mathcal{C}_\phi$ *;*

(3) $J_1(\cdot,\cdot)$ *is a strictly convex function, whereas* $J_2(\cdot,\cdot)$ *is just a convex function.*

Proof See Subsection 5.3. □

Remark 2.2 *Note that, the restriction* $|y| < x$ *in the next expression (2.19) implies that* $0 < x$. *These two restrictions are related, respectively, to the almost sure inequalities* $|\sum_{k=2}^{n} X_k X_{k-1}| < \sum_{k=1}^{n} X_k^2$ *and* $0 < \sum_{k=1}^{n} X_k^2$ *(see, e.g., McLeod and Jiménez [33]).*

Proposition 2.2 *Given* $0 < |\phi| < 1$, *the extended real function* $J(\cdot,\cdot) : \mathbb{R}^2 \to \mathbb{R} \cup \{+\infty\}$, *defined by*

$$J(x,y) = \begin{cases} J_1(x,y) = \dfrac{1}{2}\left[x(1+\phi^2) - 1 - 2y\phi + \log\left(\dfrac{x}{x^2-y^2}\right)\right], & if (x,y) \in \mathcal{A}_\phi, \\ J_2(x,y) = \dfrac{(\phi y - x)^2}{2x} + \log|\phi|, & if (x,y) \in \mathcal{B}_\phi, \\ +\infty, & otherwise, \end{cases} \tag{2.19}$$

is the explicit Fenchel-Legendre transform of $L(\cdot,\cdot)$, *given in expression* (2.15).

Proof See Subsection 5.4. □

A graph of the function $J(\cdot,\cdot)$, in (2.19), is shown in Figure 2.2 when $\phi = 0.8$. Since $L(\cdot,\cdot)$ is a convex function, $J(\cdot,\cdot)$ must also be a convex function (Ellis [19], section VI.5). In the following subsection, we show that $J(\cdot,\cdot)$ is related to the large deviations properties associated to the sequence $(\mathcal{W}_n)_{n\geqslant 2}$, defined in (1.4), and that it is a good rate function.

2.3 The Gärtner-Ellis' theorem and LDP for $(\mathcal{W}_n)_{n\in\mathbb{N}}$

Two main conditions must be satisfied to apply the general version of the Gärtner-Ellis' theorem described by Dembo and Zeitouni [16] on pages 43-44.

- **Condition A** for each $(\lambda_1,\lambda_2) \in \mathbb{R}^2$, the limiting cumulant generating function $L(\cdot,\cdot)$, defined as the limit in (1.8), exists as an extended real number. Moreover, if

$$\mathcal{D}_L = \{(\lambda_1,\lambda_2) \in \mathbb{R}^2 \mid L(\lambda_1,\lambda_2) < +\infty\}$$

denotes the *effective domain* of $L(\cdot,\cdot)$, the origin must belong to $\mathcal{D}_L^\circ$ (the interior of $\mathcal{D}_L$).

- **Condition B** $L(\cdot,\cdot)$ is lower semicontinuous and an *essentially smooth* function, that is,

(1) $\mathcal{D}_L^\circ$ is non-empty;

(2) $L(\cdot,\cdot)$ is differentiable throughout $\mathcal{D}_L^\circ$;

(3) $L(\cdot,\cdot)$ is *steep*, i.e., $\lim_{n\to+\infty} ||\nabla L(\lambda_{1,n},\lambda_{2,n})|| = +\infty$ in the case $(\lambda_{1,n},\lambda_{2,n})_{n\in\mathbb{N}}$ is a sequence in $\mathcal{D}_L^\circ$ converging to a boundary point of $\mathcal{D}_L^\circ$, where $||(x,y)|| = \sqrt{x^2+y^2}$ denotes the usual Euclidean norm in $\mathbb{R}^2$.

Let $L(\cdot,\cdot)$ denote the function in (2.15) and $\mathcal{D}$ the domain defined in Lemma 2.1. The effective domain of $L(\cdot,\cdot)$ is given by the set $\mathcal{D}$, i.e., $\mathcal{D}_L = \mathcal{D}$. Furthermore, if $(\lambda_1,\lambda_2) = (0,0)$, then

$$0 < (1-\phi^2)^2 \Rightarrow 0 < 1 - 2\phi^2 + \phi^4 \Rightarrow 4\phi^2 < (1+\phi^2)^2$$

and, since $|\phi| < 1$, it follows that $0 < \dfrac{1-\phi^2}{2}$. Whence, the origin $(0,0) \in \mathbb{R}^2$ belongs to the interior of $\mathcal{D}_1$ and, consequently, to $\mathcal{D}$ as well, proving that Condition A above is fulfilled.

If $\phi = 0$, then $L(\cdot,\cdot)$ is lower semicontinuous and an essentially smooth function (see section 3.6 of Bryc and Dembo [10]), so that Condition B is verified and the Gärtner-Ellis' theorem is fully applicable. However, when $\phi \neq 0$, the sequence $\{(\lambda_{1,n},\lambda_{2,n}) = (\frac{2^n-1}{2^{n+1}}, -\phi), n \geqslant 1\} \subset \mathcal{D}_2^\circ$ is such that $(\lambda_{1,n},\lambda_{2,n}) \to (1/2,-\phi) \notin \mathcal{D}$, but $\lim_{n\to+\infty} ||\nabla L(\lambda_{1,n},\lambda_{2,n})|| = \phi^{-2} < +\infty$. Therefore, although Conditions B.1 and B.2 are true, Condition B.3 is not, implying that $L(\cdot,\cdot)$ is not essentially smooth. By a similar argument, $L(\cdot,\cdot)$ is not lower semicontinuous. Since $L(\cdot,\cdot)$ does not satisfy the Condition B above when $\phi \neq 0$, we will be able to show only a weaker LDP version of the Gärtner-Ellis' theorem.

Definition 2.1 *The point* $(a,b) \in \mathbb{R}^2$ *is an* exposed point *of* $J(\cdot,\cdot)$ *if, for some* $\boldsymbol{z} \in \mathbb{R}^2$ *and all* $(x,y) \neq (a,b)$,

$$\langle \boldsymbol{z}, (a-x, b-y) \rangle > J(a,b) - J(x,y). \tag{2.20}$$

The vector $\boldsymbol{z}$ *is called an* exposing hyperplane *associated to* (a,b). *Given* ϕ, *we denote by* $\mathcal{F}_\phi$ *the set of exposed points whose exposing hyperplane belongs to* $\mathcal{D}_L^\circ$.

Remark 2.3 *If* $J(\cdot,\cdot)$ *is a strictly convex function, then every point is exposed.*

The next theorem describes the large deviation properties and is one of our main results. The set of exposed points $\mathcal{F}_\phi$ is explicitly determined in Subsection 5.5.

Theorem 2.1 *Let* $(\boldsymbol{\mathcal{W}}_n)_{n\geqslant 2}$ *be the sequence defined by* (1.4) *and* $J(\cdot,\cdot)$ *the function in* (2.19). *Then,* $J(\cdot,\cdot)$ *is a good rate function and:*

(1) *for any closed set* $F \subset \mathbb{R}^2$

$$\limsup_{n\to+\infty} \frac{1}{n} \log \mathbb{P}(\boldsymbol{\mathcal{W}}_n \in F) \leqslant - \inf_{(x,y)\in F} J(x,y); \tag{2.21}$$

(2) *for any open set* $G \subset \mathbb{R}^2$

$$\liminf_{n\to+\infty} \frac{1}{n} \log \mathbb{P}(\boldsymbol{\mathcal{W}}_n \in G) \geqslant - \inf_{(x,y)\in G\cap\mathcal{F}_\phi} J(x,y), \tag{2.22}$$

where $\mathcal{F}_\phi$ *is the set of exposed points of* $J(\cdot,\cdot)$ *whose exposing hyperplane belongs to* $\mathcal{D}_L^\circ$, *given by*

$$\mathcal{F}_\phi = \{(x,y) \in \mathbb{R}^2 : \ |y| < x \ \text{ and } \ y^2\phi^2 > x(x\phi^2 - 1)\}. \tag{2.23}$$

That is, $\mathcal{F}_\phi = \mathcal{A}_\phi^\circ$.

Proof See Subsection 5.5. □

Theorem 2.1 does not provide a full LDP in the sense of Definition 1.1 because we have to deal with exposed points and with the concept of steepness. Anyway, it states the precise lower and upper rates for the probabilities in (1.5). Note that, the upper bounds in (1.6) and (2.21) are identical, but the lower bounds in (1.7) and (2.22) are not, the former being much stronger than the latter one. We also observe that, if G is an open set such that $G \cap \mathcal{F}_\phi = \varnothing$, then, by *Remark* 1.1, the lower bound in (2.22) is equal to $-\infty$ and not of great interest. On the other hand, if $G \subset \mathcal{F}_\phi$ is open or closed, then the upper and lower bounds coincide. However, if $G \cap \mathcal{F}_\phi \neq \varnothing$ and G is not entirely contained in $\mathcal{F}_\phi$, then the upper bounds and lower bounds in (2.21) and (2.22) might be different (see Example 2.2 below).

Let us give two examples to enlighten Theorem 2.1.

Example 2.1 *Let us fix* $\phi = 0.5$. *Consider the open set* $G = (2,3) \times (-1,1)$ *and the closed set* $F = \overline{G} = [2,3] \times [-1,1]$. *As Figure 2.3 shows,* G *is the interior of the rectangular region with vertices at the points* $(2,-1), (2,1), (3,1)$ *and* $(3,-1)$, *while* F *is the closure of* G. *Note that, since* $J(\cdot,\cdot)$ *is a continuous function, taking infimum (or supremum) in* F *or* G *gives the same value. Moreover, both* G *and* F *lie in the interior of* $\mathcal{F}_{0.5}$ *and, therefore, in the domain of* $J_1(\cdot,\cdot)$. *Since*

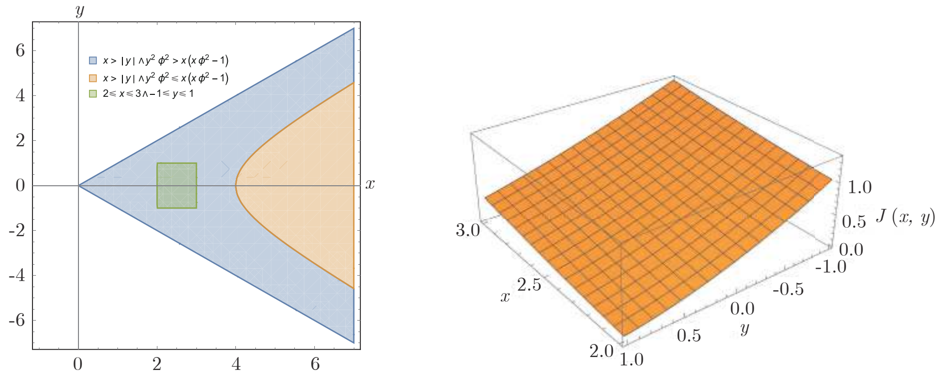


Figure 2.3 In the left-hand side panel, a plot of the region $F = [2,3] \times [-1,1]$ is presented, showing that it is contained in the set of exposed points $\mathcal{F}_{0.5}$. In the right-hand side panel, a graph of the function $J(x,y)$, for $(x,\, y) \in F$.

$$\inf_{(x,y)\in G\cap\mathcal{F}_{0.5}} J(x,y) = \inf_{(x,y)\in G} J(x,y) = \inf_{(x,y)\in F} J(x,y) = J_1(2,1) = \frac{1}{4} + \frac{\log(2/3)}{2},$$

it follows from Theorem 2.1 (particularly from the expressions (2.21) and (2.22)) that

$$-\frac{1}{4} - \frac{\log(2/3)}{2} \leqslant \liminf_{n\to+\infty} \frac{1}{n} \log \mathbb{P}(\boldsymbol{\mathcal{W}}_n \in G) \leqslant \limsup_{n\to+\infty} \frac{1}{n} \log \mathbb{P}(\boldsymbol{\mathcal{W}}_n \in F) \leqslant -\frac{1}{4} - \frac{\log(2/3)}{2},$$

implying that

$$\lim_{n\to+\infty} \frac{1}{n} \log \mathbb{P}(\boldsymbol{\mathcal{W}}_n \in G) = \lim_{n\to+\infty} \frac{1}{n} \log \mathbb{P}(\boldsymbol{\mathcal{W}}_n \in F) = -\frac{1}{4} - \frac{\log(2/3)}{2} \approx -0.047267.$$

Here, the upper and lower bounds, given respectively in (2.21) and (2.22), coincide.

Example 2.2 *Let us consider again* $\phi = 0.5$. *But now, consider the case when* $G = (5,6) \times (-4,-1)$ *and* $F = \overline{G} = [5,6] \times [-4,-1]$, *so that* G *and* F *lie no more entirely in the set of exposed points* $\mathcal{F}_{0.5}$ *(see Figure 2.4). In this case, the lower and upper bounds in (2.22) and (2.21) are not equal, as*

$$\inf_{(x,y)\in F} J(x,y) = J_2(5,-1) = \frac{121}{40} - \log(2) \approx 2.33185$$

and

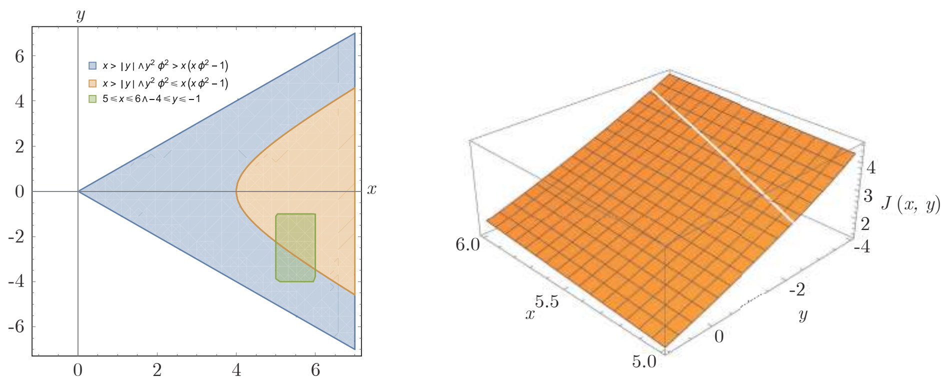


Figure 2.4 In the left-hand side panel, a plot of the region $F = [5,6] \times [-4,-1]$ is presented, showing that it is not entirely contained in the set of exposed points $\mathcal{F}_{0.5}$. In the right-hand side panel, a graph of the function $J(x,y)$ for $(x,\, y) \in F$.

$$\inf_{(x,y)\in G\cap\mathcal{F}_{0.5}} J(x,y) = \inf_{\left\{5<x<6,-4<y<-\sqrt{\frac{x(x\phi^2-1)}{\phi^2}}\right\}} J(x,y) = J_1(5,-\sqrt{5}) \approx 3.04377.$$

Hence, from Theorem 2.1 it follows that

$$-3.04377 \leqslant \liminf_{n\to+\infty} \frac{1}{n}\log\mathbb{P}(\boldsymbol{\mathcal{W}}_n \in G) \leqslant \limsup_{n\to+\infty}\frac{1}{n}\log\mathbb{P}(\boldsymbol{\mathcal{W}}_n \in F) \leqslant -2.33185.$$

Here, the upper and lower bounds do not coincide.

When $\phi = 0$, if we set $\mathcal{F}_0 = \{(x,y)\in\mathbb{R}^2 : |y| < x\}$ then $(\boldsymbol{\mathcal{W}}_n)_{n\in\mathbb{N}}$ actually satisfies the LDP with rate function $J(x,y) = J_1(x,y)$, for $(x,y)\in\mathcal{F}_0$ and $J(x,y) = +\infty$, otherwise (see Bryc and Dembo [10]). In other words, the lower bounds in (1.7) and (2.22) are identical for this particular case. Furthermore, it is important to remark that the sets of exposed points form a sequence of enclosing sets, in the sense that, if $|\phi_1| < |\phi_2|$, then $\mathcal{F}_{\phi_2} \subset \mathcal{F}_{\phi_1}$; we also have $\mathcal{F}_\phi \subset \mathcal{F}_0$ and $\mathcal{F}_{-\phi} = \mathcal{F}_\phi$, for any $0 < |\phi| < 1$.

3. LDP for one-dimensional processes via contraction

We dedicate this section to showing three particular examples where the reasoning of the last section can be applied, via the usage of a weaker version of the Contraction Principle, to get explicit rate functions for univariate random sequences. The LDP for two of these examples were already known from Bercu *et al.* [5] and Bryc and Smolenski [11]. We show that they can be obtained from Theorem 2.1 and two additional results, namely Theorem 3.1 and Remark 3.1 stated below. In Subsection 3.2, we demonstrate a result that we believe is new in the literature.

Since we proved only a weak version of the LPD for $\boldsymbol{\mathcal{W}}_n$, defined in (1.4), the Contraction Principle as presented in Dembo and Zeitouni [16] is not applicable. However, some versions in the literature make it possible to show weaker versions of the LDP. For instance, Lewis and Pfister [29] show a variation of the Contraction Principle for Vague Large Deviation Principles (see their theorem 3.3), and for Narrow Large Deviation Principles (see their theorem 5.2). Another thorough discussion about transformations of weak LDPs is provided by Fayolle and de La Fortelle [20].

The reasoning used in the proof of the Contraction Principle (see theorem 4.2.1 in Dembo and Zeitouni [16]) is not appropriate for our needs. In subsection 5.2 of Robertson and Almost [36], the authors proved the Contraction Principle first for the lower bound and next for the upper bound, which is more suitable for our setting. In the proof of the main results in this section, we will need the following theorem.

Theorem 3.1 (Weak Contraction Principle). *Consider two sets $E \subseteq \mathbb{R}^d$ and $E' \subseteq \mathbb{R}$, a continuous function $f(\cdot) : E \to E'$, and a good rate function $J(\cdot) : E \to [0,+\infty]$.*

(1) *Define for each $c \in E'$ the function*

$$I(c) = \inf_{\boldsymbol{x}\in E}\left\{J(\boldsymbol{x}) \mid \textit{ with } \boldsymbol{x} \textit{ such that } f(\boldsymbol{x}) = c\right\}.$$

Then $I : E' \to [0,+\infty]$ is a good rate function.

(2) *If $J(\cdot)$ controls the LDP associated to a sequence $(\boldsymbol{Z}_n)_{n\in\mathbb{N}}$ on E with upper bound*

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}(\boldsymbol{Z}_n \in F) \leqslant -\inf_{\boldsymbol{x}\in F} J(\boldsymbol{x}), \quad \textit{for any closed set } F \subset E, \tag{3.1}$$

and lower bound

$$\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}(\boldsymbol{Z}_n\in G)\geqslant -\inf_{\boldsymbol{x}\in G\cap\mathcal{F}} J(\boldsymbol{x}),\qquad \textit{for any open set } G\subset E, \tag{3.2}$$

where $\mathcal{F}$ *is the set of exposed points, then* $(f(\boldsymbol{Z}_n))_{n\in\mathbb{N}}$, *defined on* E', *satisfies a weak LDP with upper bound*

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}(f(\boldsymbol{Z}_n)\in F')\leqslant -\inf_{\boldsymbol{x}\in f^{-1}(F')} J(\boldsymbol{x}),\qquad \textit{for any closed set } F'\subset E', \tag{3.3}$$

and lower bound

$$\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}(f(\boldsymbol{Z}_n)\in G')\geqslant -\inf_{\boldsymbol{x}\in f^{-1}(G')\cap\mathcal{F}} J(\boldsymbol{x}),\qquad \textit{for any open set } G'\subset E'. \tag{3.4}$$

Proof The proof of Statement 1 is the same as the one given in Dembo and Zeitouni [16], page 127. For the proof of Statement 2, to show the upper bound, we observe that F' being closed in E', together with the hypothesis of $f(\cdot)$ being continuous, implies that $f^{-1}(F')$ is closed in E. Hence, the upper bound in (3.3) follows as a consequence of (3.1), i.e.,

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}(f(\boldsymbol{Z}_n)\in\mathcal{F}')=\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}(\boldsymbol{Z}_n\in f^{-1}(\mathcal{F}'))\leqslant -\inf_{\boldsymbol{x}\in f^{-1}(F')} J(\boldsymbol{x}).$$

Similarly, G' being open in E' implies that $f^{-1}(G')$ is open in E. Therefore, using (3.2) one obtains

$$\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}(f(\boldsymbol{Z}_n)\in G')=\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}(\boldsymbol{Z}_n\in f^{-1}(G'))\geqslant -\inf_{\boldsymbol{x}\in f^{-1}(G')\cap\mathcal{F}} J(\boldsymbol{x}).$$

The above proof was adapted from subsection 5.2 in Robertson and Almost [36]. □

Remark 3.1 *From the lower bound in (3.4), under particular circumstances, if one can show that*

$$\inf_{\boldsymbol{x}\in f^{-1}(G')\cap\mathcal{F}} J(\boldsymbol{x})=\inf_{\boldsymbol{x}\in f^{-1}(G')} J(\boldsymbol{x}), \tag{3.5}$$

for any open set $G'\subset E'$, *then a full LDP for* $(f(\boldsymbol{Z}_n))_{n\in\mathbb{N}}$ *holds, in the sense of Definition 1.1.*

We show here three particular cases in which Theorem 3.1 is applicable, but it just gives us a weak type of LDP. In despite of that, we can get a full LDP (in the sense of Definition 1.1) by proving that (3.5) holds in these three cases. Note that, for the lower bound (3.2) it is necessary to consider the set $G\cap\mathcal{F}$, and the Theorem 3.1 takes this into account.

Since the sequence of random vectors $(\boldsymbol{\mathcal{W}}_n)_{n\geqslant 2}$ has $J(\cdot,\cdot)$, given in (2.19), as a good rate function, Theorem 3.1, statement 1, ensures that any sequence of vectors $(f(\boldsymbol{\mathcal{W}}_n))_{n\geqslant 2}$, for $f:\mathbb{R}^2\to\mathbb{R}$ continuous, has a good rate function

$$I(c)=\inf_{(x,y)\in\mathbb{R}^2}\big\{J(x,y)\,|\text{ with }(x,y)\text{ such that } f(x,y)=c\big\},\qquad\text{for all } c\in\mathbb{R}. \tag{3.6}$$

There is a standard procedure involving Calculus techniques for computing the infimum in (3.6), namely, checking for the critical points of the derivatives from $J(\cdot,\cdot)$. In the examples considered below, the Wolfram Mathematica software (version 11.2.0.0) was used in the calculations.

Note that, $f(\boldsymbol{\mathcal{W}}_n)=f\left(\frac{1}{n}\left(\sum_{k=1}^n X_k^2,\sum_{k=2}^n X_kX_{k-1}\right)\right)$ is a continuous function involving only the components $\frac{1}{n}\sum_{k=1}^n X_k^2$ and $\frac{1}{n}\sum_{k=2}^n X_kX_{k-1}$. Any statistic that can be written in terms of these components, as a continuous transformation of $\boldsymbol{\mathcal{W}}_n$, is suitable for the method presented below. Our focus will be for the functions $f_i:\mathbb{R}^2\to\mathbb{R}$, for $i=1,2,3$, respectively defined by

(1) $f_1(x,y)=x$,
(2) $f_2(x,y)=y$,

(3) $f_3(x,y) = \dfrac{y}{x}$, for $x > 0$,

to be considered in Sections 3.1-3.3. Other continuous functions $f : \mathbb{R}^2 \to \mathbb{R}$ could also be considered. However, in the present work, we restrict our attention to these three cases because of the following: to apply Theorem 3.1 it is necessary to analyze the values of the deviation function only along certain straight lines on the plane $\mathbb{R}^2$; we will show that the exposed points do not interfere in the estimation of the rate function that we seek.

3.1 LDP for the sample second moment

Consider the sample second moment $\tilde{\gamma}_n(0) = \dfrac{1}{n}\sum_{k=1}^n X_k^2$ of a random sample $X_1, \ldots, X_n$, extracted from the process $(X_j)_{j\in\mathbb{N}}$ which satisfies (1.2). Bryc and Smolenski [11] proved that the sequence $(\tilde{\gamma}_n(0))_{n\in\mathbb{N}}$ satisfies the LDP (Definition 1.1) with rate function given by

$$\mathbb{I}(c) = \begin{cases} \dfrac{1}{2}\left[c\left(1+\phi^2\right) - \sqrt{1+4\phi^2c^2} - \log\left(\dfrac{2c}{1+\sqrt{1+4\phi^2c^2}}\right)\right], & \text{if } c > 0, \\ +\infty, & \text{if } c \leqslant 0. \end{cases} \tag{3.7}$$

Here we will show how to obtain this rate function as a particular case, by using Proposition 2.2 and Theorem 3.1.

Note that $\tilde{\gamma}_n(0)$ is equal to the first coordinate of the vector $\boldsymbol{\mathcal{W}}_n$, given in (1.4). Consider the continuous function $f_1 : \mathbb{R}^2 \to \mathbb{R}$ given by $f_1(x,y) = x$. Then $f_1(\boldsymbol{\mathcal{W}}_n) = \tilde{\gamma}_n(0)$ and, since $(\boldsymbol{\mathcal{W}}_n)_{n\geqslant 2}$ has good rate function $J(\cdot,\cdot)$, given in (2.19), we use the result of Theorem 3.1 to prove that $(\tilde{\gamma}_n(0))_{n\in\mathbb{N}}$ satisfies an upper and lower bound of the types (3.3) and (3.4), respectively. To apply Theorem 3.1 it is necessary to analyze the values of the deviation function along straight lines parallel to the y axis. Next, we will show that (3.5) holds, proving that $(\tilde{\gamma}_n(0))_{n\in\mathbb{N}}$ satisfies the LDP with associated rate function which we denote by $I_1 : \mathbb{R} \to [0,+\infty]$. Finally, we will conclude that $I_1 = \mathbb{I}$.

Take $c > 0$, then $\{(x,y) \in \mathbb{R}^2 | f_1(x,y) = c\}$ is the straight line with equation $x = c$, parallel to the y axis. Note that, if $0 < c \leqslant 1/\phi^2$, then $J(c,y) = J_1(c,y) = \dfrac{1}{2}\big[c(1+\phi^2) - 1 - 2y\phi + \log\left(\dfrac{c}{c^2-y^2}\right)\big]$, for $|y| < c$; whereas if $c > 1/\phi^2$, then

$$J(c,y) = \begin{cases} J_1(c,y), & \sqrt{\dfrac{c(c\phi^2-1)}{\phi^2}} < |y| < c, \\ J_2(c,y), & |y| \leqslant \alpha_c, \end{cases}$$

with $J_2(c,y) = \dfrac{(\phi y - c)^2}{2c} + \log|\phi|$ and $\alpha_c = \sqrt{\dfrac{c(c\phi^2-1)}{\phi^2}}$. To simplify the reading, let us assume that $\phi > 0$ (if $\phi < 0$, we can treat this case by symmetry, considering $y \mapsto J(c,-y)$). Also, let us concentrate on the case $c > 1/\phi^2$. The first-order derivatives of $J_1(\cdot)$ and $J_2(\cdot)$ are

$$\frac{\mathrm{d}}{\mathrm{d}y}J_1(c,y) = \frac{y}{c^2-y^2} - \phi \qquad \text{and} \qquad \frac{\mathrm{d}}{\mathrm{d}y}J_2(c,y) = \frac{\phi(y\phi - c)}{c}. \tag{3.8}$$

Hence, the derivative of $J_1(c,\cdot)$ is negative in $(-c, y_c)$ and positive in (y_c, c), where $y_c = \dfrac{-1+\sqrt{1+4c^2\phi^2}}{2\phi}$. The global minimum of $J_1(c,\cdot)$ is attained at y_c. Moreover, given that $\phi > 0$, it follows that $\alpha_c < y_c$. On the other hand, since $y\phi - c < 0$, $J_2(c,\cdot)$ is a decreasing function in terms of y. Figure 3.1 illustrates the case when $c = 4$ and $\phi = 0.8$.

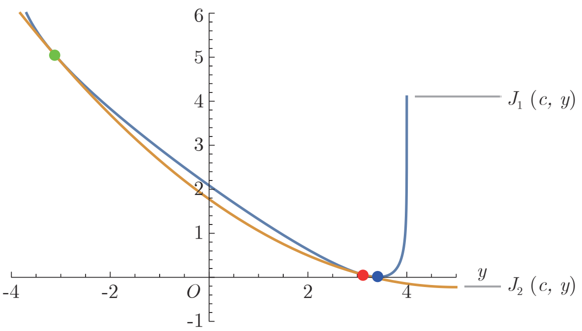


Figure 3.1 Graphs of $J_1(c,y)$ and $J_2(c,y)$ with $\phi=0.8$ and $c=4$. The green and red dots represent the points $-\alpha_c$ and α_c, respectively, where $J_1(c,\cdot)$ and $J_2(c,\cdot)$ change roles in the law of $J(\cdot,\cdot)$. The blue dot is the point $(y_c, J(c,y_c))$. As this graph shows us, given that $\phi>0$, $J(c,y)$ is equal to $J_1(c,y)$ if $y\in(c,-\alpha_c)\cup(\alpha_c,c)$, and equal to $J_2(c,y)$ if $y\in[-\alpha_c,\alpha_c]$. When $J(c,y)$ changes from $J_2(c,y)$ to $J_1(c,y)$ on the red dot α_c, the derivative of $J_1(c,y)$ being negative for $y\in(\alpha_c,y_c)$ shows that this function keeps decreasing from left to right until it reaches the global minimum at y_c (remember from Proposition 2.1 that J_1 and J_2 are convex), so that every point on its right and its left is greater than $J_1(c,y_c)$, including the points in $[-\alpha_c,\alpha_c]$ where $J_2(c,y)$ rules. A similar analysis follows by symmetry for the case $\phi<0$.

In particular, it holds that

$$\inf_{(x,y)\in f_1^{-1}(c)\cap\mathcal{F}_\phi} J(x,y) = \inf_{(x,y)\in f_1^{-1}(c)} J_1(x,y) = \inf_{(x,y)\in f_1^{-1}(c)} J(x,y), \tag{3.9}$$

for any $c>0$. So the infimum in (3.9) is always attained in the interior of $\mathcal{F}_\phi$ and it is independent of the pre-image set $f_1^{-1}(c)$, for any given $c>0$. Therefore, since $f_1^{-1}(G') = f_1^{-1}(\cup_{c\in G'}\{c\}) = \cup_{c\in G'} f_1^{-1}(c)$ for any open set $G'\subset(0,\infty)$, the LDP for $\tilde{\gamma}_n(0)$ follows from Theorem 3.1 and Remark 3.1.

Thus, the rate function we are looking for can be computed via

$$\begin{aligned} I_1(c) &= \inf_{|y|<x}\{J(x,y)\,|\,f_1(x,y)=c\} = \inf_{|y|<x}\{J(x,y)\,|\,x=c\} = J_1(c,y_c) \\ &= \frac{1}{2}\left[c(1+\phi^2) - 1 - 2y_c\phi + \log\left(\frac{c}{c^2-y_c^2}\right)\right] \\ &= \frac{1}{2}\left[c(1+\phi^2) - \sqrt{1+4c^2\phi^2} + \log\left(\frac{2c\phi^2}{\sqrt{1+4c^2\phi^2}-1}\right)\right]. \end{aligned}$$

Then, after some algebraic computations, we obtain

$$\log\left(\frac{2c\phi^2}{\sqrt{1+4c^2\phi^2}-1}\right) = \log\left(\frac{2c\phi^2(\sqrt{1+4c^2\phi^2}+1)}{(\sqrt{1+4c^2\phi^2}-1)(\sqrt{1+4c^2\phi^2}+1)}\right)$$

$$= \log\left(\frac{2c\phi^2(\sqrt{1+4c^2\phi^2}+1)}{1+4c^2\phi^2-1}\right) = \log\left(\frac{\sqrt{1+4c^2\phi^2}+1}{2c}\right) = -\log\left(\frac{2c}{\sqrt{1+4c^2\phi^2}+1}\right),$$

ending up with

$$I_1(c) = \frac{1}{2}\left[c(1+\phi^2) - \sqrt{1+4c^2\phi^2} - \log\left(\frac{2c}{\sqrt{1+4c^2\phi^2}+1}\right)\right]. \tag{3.10}$$

Considering that $I_1(c)=+\infty$, for $c\leqslant 0$, we conclude that $I_1(c)=\mathbb{I}(c)$, for all $c\in\mathbb{R}$, with $\mathbb{I}(\cdot)$ defined in (3.7). If $\phi=0$, an application of the Contraction Principle shows that $I_1(c)=\dfrac{c-1-\log(c)}{2}$, since, in this case, $(\boldsymbol{\mathcal{W}}_n)_{n\geqslant 2}$ satisfies the full LDP.

To conclude, we get the same rate function as in expression (1.2) in Bryc and Smolenski [11]. The graphs of $I_1(\cdot)$ are illustrated in Figure 3.2 for four different values of ϕ. Notice that $I_1(\cdot)$ is symmetric with respect to the values of ϕ, i.e., $I_1(\cdot)$ is the same function for ϕ and $-\phi$, given that $\phi \in (0,1)$.

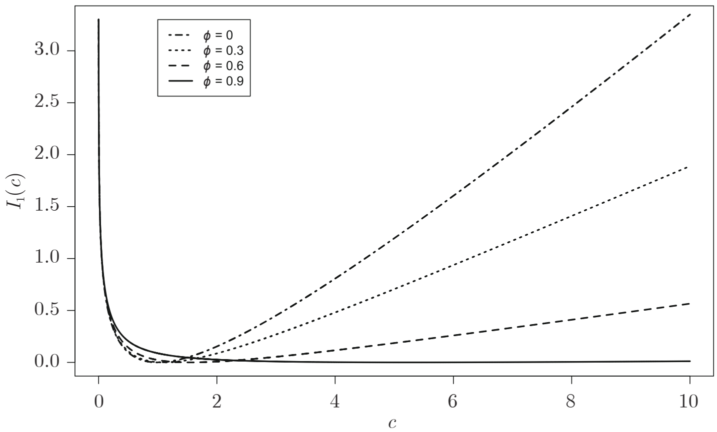


Figure 3.2 Graphs of $I_1(c)$ for $\phi \in \{0,\, 0.3,\, 0.6,\, 0.9\}$ and $c \in (0, 10]$.

3.2 LDP for the first-order empirical autocovariance

Consider now the first-order empirical autocovariance of $X_1, \ldots, X_n$, defined below as

$$\tilde{\gamma}_n(1) = \frac{1}{n}\sum_{k=2}^{n} X_k X_{k-1}. \tag{3.11}$$

Following the same reasoning as the one used for the sample second moment in Subsection 3.1, below we compute the explicit rate function associated with the sequence $(\tilde{\gamma}_n(1))_{n\geqslant 2}$, under the assumption that $(X_j)_{j\in\mathbb{N}}$ follows an AR(1) process. We believe that this result is new in the literature.

Proposition 3.1 *The sequence* $(\tilde{\gamma}_n(1))_{n\geqslant 2}$, *with* $\tilde{\gamma}_n(1)$ *defined in* (3.11) *and* $(X_j)_{j\in\mathbb{N}}$ *following an AR(1) process, as defined in* (1.2), *satisfies an LDP with good rate function* $I_2 : \mathbb{R} \to [0, +\infty]$, *where*

$$I_2(c) = \frac{1+3c^2(1+\phi^2)^2+\left(-2-6c\phi+3\log\left[\dfrac{1+3c^2(1+\phi^2)^2+A(c,\phi)+A(c,\phi)^2}{3(1+\phi^2)A(c,\phi)\left(\frac{\left(1+3c^2(1+\phi^2)^2+A(c,\phi)+A(c,\phi)^2\right)^2}{9(1+\phi^2)^2A(c,\phi)^2}-c^2\right)}\right]\right)A(c,\phi)+A(c,\phi)^2}{6A(c,\phi)}, \tag{3.12}$$

for any $c \in \mathbb{R}$, *and*

$$A(c,\phi) = \sqrt[3]{1+18c^2\left(1+\phi^2\right)^2+3\sqrt{3}\sqrt{-c^2\left(1+\phi^2\right)^2\left(c^4\left(1+\phi^2\right)^4-11c^2\left(1+\phi^2\right)^2-1\right)}}. \tag{3.13}$$

Proof Consider the continuous function $f_2 : \mathbb{R}^2 \to \mathbb{R}$, with law $f_2(x,y) = y$. Note that $f_2(\boldsymbol{\mathcal{W}}_n) = \tilde{\gamma}_n(1)$. So, Theorem 3.1 is applicable and $(\tilde{\gamma}_n(1))_{n\geqslant 2}$ has a good rate function $I_2 : \mathbb{R} \to [0, +\infty]$. To obtain a sharper lower bound than the one given by (3.4), it is necessary to analyze the values of the deviation function along straight lines parallel to the x axis. We will show that $(\tilde{\gamma}_n(1))_{n\geqslant 2}$ satisfies (3.5), hence proving that the LDP is satisfied in the sense of Definition 1.1. To achieve our goal, consider an arbitrary $c \in \mathbb{R}$. The set $\{(x,y) \in \mathbb{R}^2 | f_2(x,y) = c\}$ is the straight line with equation $y = c$, parallel to the x axis. Then, for $|c| < x$ we have

$$J_1(x,c) = \frac{1}{2}\left[x(1+\phi^2) - 1 - 2c\phi + \log\left(\frac{x}{x^2-c^2}\right)\right] \quad \text{and} \quad J_2(x,c) = \frac{(\phi c - x)^2}{2x} + \log|\phi|.$$

The derivatives with respect to x are

$$\frac{\mathrm{d}}{\mathrm{d}x}J_1(x,c) = \frac{x^2(x\phi^2 + x - 1) - c^2(x\phi^2 + x + 1)}{2(x^3 - c^2 x)} \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}x}J_2(x,c) = \frac{1}{2} - \frac{c^2\phi^2}{2x^2}.$$

Now consider $A(c,\phi)$ as in (3.13) and

$$x_c = \frac{1 + 3c^2\left(1+\phi^2\right)^2 + A(c,\phi) + A\left(c,\phi\right)^2}{3\left(1+\phi^2\right)A(c,\phi)}. \tag{3.14}$$

Then $J_1(x,c)$ is decreasing for $x \in (0, x_c)$ and increasing for $x \in (x_c, +\infty)$, attaining a global minimum at x_c. Although it is not obvious, we have $x_c > |c|$, for all $c \in \mathbb{R}$. Similarly, $J_2(x,c)$ is decreasing for $x \in (0, c\phi)$ and increasing for $(c\phi, +\infty)$. On the line $y = c$, we have

$$J(x,c) = \begin{cases} J_1(x,c), & |c| < x < \beta_c, \\ J_2(x,c), & x \geqslant \beta_c, \end{cases} \tag{3.15}$$

where $\beta_c = \dfrac{1+\sqrt{1+4c^2\phi^4}}{2\phi^2}$. Note that $\beta_c > c\phi$, because $0 < |\phi| < 1$ and

$$\beta_c = \frac{1+\sqrt{1+4c^2\phi^4}}{2\phi^2} \geqslant \frac{\sqrt{4c^2\phi^4}}{2\phi^2} \geqslant |c| > |c\phi| \geqslant c\phi.$$

We plotted the graphs of $J_1(x,c)$ and $J_2(x,c)$ side-by-side in Figure 3.3 in the particular case when $\phi = 0.8$ and $c = 2$. Notice that the minimum of $J(x,c)$ is again attained on the domain of $J_1(x,c)$. Since this is verified for every $c \in \mathbb{R}$ and the domain of $J_1(x,c)$ is precisely $\mathcal{F}_\phi$, we conclude that

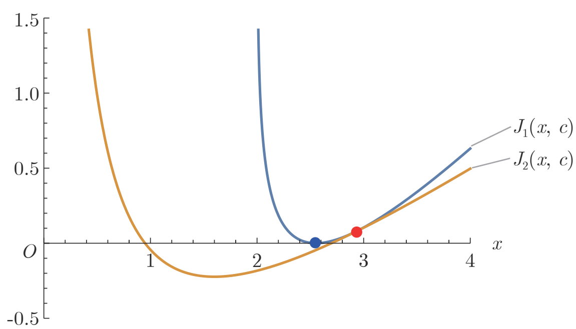


Figure 3.3 Graphs of $J_1(x,c)$ and $J_2(x,c)$, with $\phi = 0.8$. Here $c = 2$ is fixed. The red dot represents the point β_c where $J_1(\cdot,c)$ and $J_2(\cdot,c)$ change roles in the law of $J(\cdot,\cdot)$. The blue dot is the point $(x_c, J(x_c,c))$. As the graph shows us, $J(c,y)$ is equal to $J_1(c,y)$ when $x \in (|c|, \beta_c)$, and equal to $J_2(c,y)$ when $y \in [\beta_c, +\infty)$. When $J(c,y)$ changes from $J_2(x,c)$ to $J_1(x,c)$ on the red dot β_c, the derivative of $J_1(c,y)$ being positive for $y \in (x_c, +\infty)$ shows that this function keeps decreasing from right to left until it reaches the global minimum at x_c (remember from Proposition 2.1 that J_1 and J_2 are convex), so that every point on its right and its left is greater than $J_1(x_c,c)$, including the points in $[\beta_c, +\infty)$ where $J_2(x,c)$ rules.

$$\inf_{(x,y)\in f_2^{-1}(c)\cap\mathcal{F}_\phi} J(x,y) = \inf_{(x,y)\in f_2^{-1}(c)} J_1(x,y) = \inf_{(x,y)\in f_2^{-1}(c)} J(x,y), \quad \text{for all} \quad c\in\mathbb{R}.$$

Then, for the same reason as for the sample second moment, (3.5) follows and the LDP is proved. It only remains to find the explicit rate function, which can be computed as

$$\begin{aligned} I_2(c) &= \inf_{x>|y|}\{J(x,y)\,|\,f_2(x,y)=c\} = J_1(x_c,c)\\ &= \frac{1}{2}\left[x_c(1+\phi^2)-1-2c\phi+\log\left(\frac{x_c}{x_c^2-c^2}\right)\right]\end{aligned}$$

from where the expression in (3.12) follows. □

The graph of $I_2(\cdot)$ is illustrated in Figure 3.4 for five different values of ϕ. We believe that this rate function was never exhibited in the literature and the LDP of $(\tilde{\gamma}_n(1))_{n\geqslant 2}$ likewise, which gives a new and interesting result that has been derived from the general theory presented in this work.

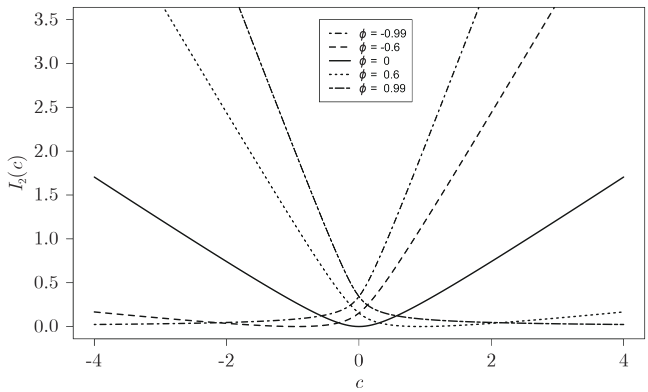


Figure 3.4 Graphs of $I_2(c)$ for $\phi\in\{-0.99,-0.6,\ 0,\ 0.6,\ 0.99\}$ and $c\in[-4,4]$.

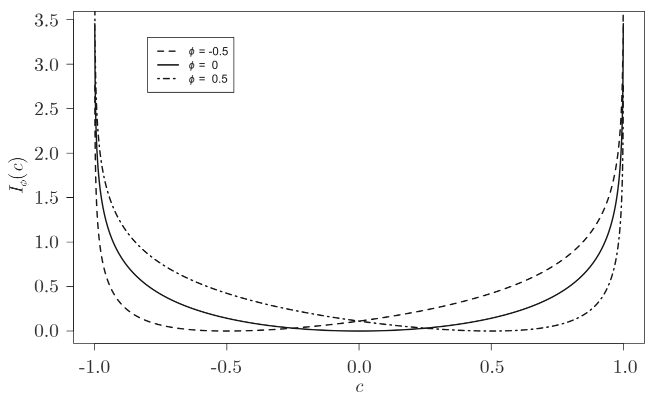


Figure 3.5 Graph of $I_\phi(c)$ for $\phi\in\{-0.5,0,0.5\}$ and $c\in(-1,1)$.

3.3 LDP for the Yule-Walker estimates

Consider the Yule-Walker estimator

$$\tilde{\phi}_n = \frac{\sum_{k=2}^{n} X_k X_{k-1}}{\sum_{k=1}^{n} X_k^2} \tag{3.16}$$

of the parameter ϕ for the AR(1) processes given in (1.2). The asymptotical behavior of (3.16) is well known, so that (Brockwell and Davis [9]) $\sqrt{n}(\tilde{\phi}_n - \phi) \Rightarrow \mathcal{N}(0, 1-\phi^2)$ and also that (Mann and Wald [31]) $\tilde{\phi}_n \xrightarrow{n\to+\infty} \phi$, almost surely.

In Bercu *et al.* [5] it was proved that the Yule-Walker estimator satisfies the LDP (Definition 1.1) with the rate function given by

$$S(c) = \begin{cases} \frac{1}{2}\log\left(\frac{1+\phi^2-2\phi c}{1-c^2}\right), & \text{if } |c|<1, \\ +\infty, & \text{otherwise.} \end{cases}$$

Since the rate function can be related to the sequence of probabilities $\mathbb{P}\left(\tilde{\phi}_n \geqslant c\right)$, for $|\tilde{\phi}_n| < 1$ and $n \geqslant 2$, note that $S(c)$ is finite, for $|c|<1$, and infinite when $|c| \geqslant 1$. Later on, Bercu *et al.* [4] provided a Sharp Large Deviation Principle (SLDP) for Hermitian quadratic forms of stationary Gaussian processes, obtaining the Yule-Walker's SLDP as a particular case. In Bercu [3], the study on LDP of the Yule-Walker estimator in AR(1) processes was extended to the unstable ($|\phi|=1$) and explosive ($|\phi|>1$) cases.

Here we obtain the result from Bercu *et al.* [5] by using the results of Proposition 2.2, Theorem 3.1, and Remark 3.1. From (3.16) note that $\tilde{\phi}_n = f_3(\boldsymbol{\mathcal{W}}_n)$, where $\boldsymbol{\mathcal{W}}_n$ is the random vector given in (1.4) and $f_3 : \mathbb{R}^2 \to \mathbb{R}$ is the continuous function defined by

$$f_3(x,y) = \frac{y}{x}, \quad \text{for } |y| < x. \tag{3.17}$$

Hence, Theorem 3.1 is applicable and $(\tilde{\phi}_n)_{n\geqslant 2}$ must satisfy an upper and lower bounds of the type (3.3) and (3.4), respectively. Moreover, by the same reasoning, as shown in Subsections 3.1 and 3.2, we can prove that the lower bound is a lower bound of the type (1.7). Indeed, consider $c \in (-1,1)$. Then the set $\{(x,y)\in\mathbb{R}^2 | f_3(x,y)=c\}$ is the straight line that passes through the origin and has slope c. Additionally, on this line we have

$$J_1(x,cx) = \frac{1}{2}\left[x(1+\phi^2) - 1 - 2cx\phi + \log\left(\frac{x}{x^2-c^2x^2}\right)\right] \quad \text{and} \quad J_2(x,cx) = \frac{1}{2x}(\phi cx - x)^2 + \log|\phi|,$$

with derivatives given respectively by

$$\frac{\mathrm{d}}{\mathrm{d}x}J_1(x,cx) = \frac{x(1-2c\phi+\phi^2)-1}{2x} \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}x}J_2(x,cx) = \frac{1}{2}(c\phi-1)^2.$$

Since $|c|<1$ and $|\phi|<1$, it immediately follows that the derivative of $J_2(x,cx)$ is positive, for all c and ϕ, proving that this function is increasing. On the other hand, the function $J_1(x,cx)$ has a global minimum at $x_c = \frac{1}{1-2c\phi+\phi^2}$. By definition,

$$J(x,cx) = \begin{cases} J_1(x,cx), & 0 < x < \delta_c, \\ J_2(x,c,x), & x \geqslant \delta_c, \end{cases} \tag{3.18}$$

where $\delta_c = \frac{1}{(1-c^2)\phi^2}$. Since $|c\phi|<1$, it is easy to prove that $x_c < \delta_c$ for all $c\in(-1,1)$ and any given $\theta \in (-1,1)$. Figure 3.6 illustrates a particular case when $\phi = 0.8$ and $c = 0.5$.

We have that

$$\inf_{(x,y)\in f_3^{-1}(c)\cap\mathcal{F}_\phi} J(x,y) = \inf_{(x,y)\in f_3^{-1}(c)} J_1(x,y) = \inf_{(x,y)\in f_3^{-1}(c)} J(x,y), \quad \text{for all } c\in(-1,1),$$

implying that (3.5) is valid and that the LDP for the sequence $(\tilde{\phi}_n)_{n\geqslant 2}$ holds. The associated rate function $I_\phi(\cdot) : \mathbb{R} \to [0,+\infty]$ can be computed as follows.

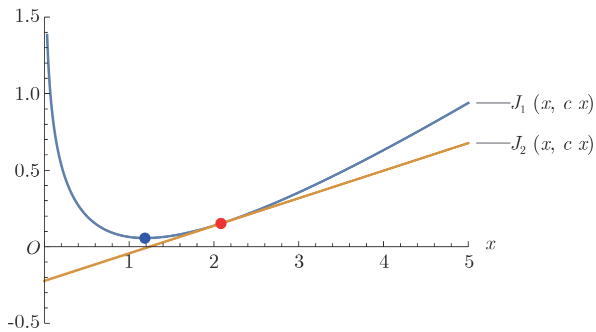


Figure 3.6 Graphs of $J_1(x, cx)$ and $J_2(x, cx)$, with $\phi = 0.8$. Here $c = 0.5$ is fixed. The red dot represents the point δ_c where $J_1(x, cx)$ and $J_2(x, cx)$ change roles in the law of $J(\cdot, \cdot)$. The blue dot is the point $(x_c, J(x_c, cx_c))$. Notice that $J(x, cx)$ is equal to $J_1(x, cx)$ when $x \in (0, \delta_c)$, and equal to $J_2(x, cx)$ when $x \in [\delta_c, +\infty)$. When $J(x, cx)$ changes from $J_2(x, cx)$ to $J_1(x, cx)$ on the red dot δ_c, the derivative of $J_1(x, cx)$ being positive for $x \in (x_c, +\infty)$ shows that this function keeps decreasing from right to left until it reaches the global minimum at x_c (remember from Proposition 2.1 that J_1 and J_2 are convex) so that every point on its right and its left is greater than $J_1(x_c, cx_c)$, including the points in $[\delta_c, +\infty)$ where $J_2(x, cx)$ rules.

$$
\begin{aligned}
I_\phi(c) &= \inf_{|y|<x} \{J(x, y) \,|\, f_3(x, y) = c\} = J_1(x_c, c, x_c) \\
&= \frac{1}{2}\left[x_c(1 - 2c\phi + \phi^2) - 1 + \log\left(\frac{1}{x_c(1 - c^2)}\right)\right] = \frac{1}{2}\log\left(\frac{1 + \phi^2 - 2\phi c}{1 - c^2}\right).
\end{aligned}
$$

Considering that $I_\phi(c) = +\infty$, for $|c| \geqslant 1$, we obtain $I_\phi(c) = S(c), \forall c \in \mathbb{R}$. Therefore, we get the same result as in expression (4.6) in Bercu *et al.* [5]. The graph of $I_\phi(\cdot)$ is illustrated in Figure 3.5 for three different values of ϕ.

4. Large deviations for the S^2-mean

Given that the rate function for the sample second moment $\frac{1}{n}\sum_{k=1}^n X_k^2$, defined in (1.1), is known (see the functional in (3.7)), we will show in Subsection 4.1 that there exists a simple approach leading to the LDP, and also its deviation function, for the sequence of bivariate S^2-mean $(\boldsymbol{S}_n)_{n\in\mathbb{N}}$, where

$$
\boldsymbol{S}_n = \frac{1}{n}\left(\sum_{k=1}^n X_k, \ \sum_{k=1}^n X_k^2\right), \tag{4.1}
$$

in the case of an AR(1) process. As the main auxiliary tool, we use a proposition proved in Bryc and Dembo [10], enunciated below for completeness. In Section 4.2 we follow the same reasoning for getting similar results for the MA(1) process.

Proposition 4.1 *Let* $(X_j)_{j\in\mathbb{N}}$ *be a real-valued centered stationary Gaussian process whose spectral density* $f(\cdot)$ *is differentiable. Then,* $(\boldsymbol{S}_n)_{n\in\mathbb{N}}$*, for* $\boldsymbol{S}_n$ *given in* (4.1)*, satisfies the LDP (in* $\mathbb{R}^2$*) with good rate function*

$$
H(x, y) = I(y - x^2) + \frac{x^2}{2f(0)}, \tag{4.2}
$$

where $0/0 := 0$ *in* (4.2) *and* $I(\cdot)$ *is the rate function associated to* $(n^{-1}\sum_{k=1}^n X_k^2)$.

Proof See section 3.5 in Bryc and Dembo [10]. □

We dedicate the next two subsections to the particular study of the LDP of the S^2-mean when $(X_j)_{j\in\mathbb{N}}$ is an AR(1) process (Subsection 4.1), and when it is an MA(1) process

(Subsection 4.2). Since the LDP for the sample second moment was already established for the AR(1) process, it is straightforward to show such a property in this case. For the MA(1) process, however, we must first derive the LDP of the sample second moment to apply Proposition 4.1 and to provide the LDP for the S^2-mean, likewise.

4.1 AR(1) process

In this subsection we apply expression (3.7), established by Bryc and Smolenski [11], to obtain an explicit expression for the rate function of the sample mean of the AR(1) process $(X_j)_{j\in\mathbb{N}}$, defined in (1.2). Since this process is a real-valued centered stationary Gaussian process, it follows from Proposition 4.1 that $(\boldsymbol{S}_n)_{n\in\mathbb{N}}$ satisfies the LDP with rate function

$$H_1(x,y) = \mathbb{I}(y-x^2) + \frac{x^2}{2g_\phi(0)},$$

where $\mathbb{I}(\cdot)$ is defined by (3.7) and $g_\phi(\cdot)$ denotes the spectral density function given in (1.3). Note that $g_\phi(\cdot)$ is differentiable. The explicit rate function is given by

$$H_1(x,y)=\begin{cases}\frac{1}{2}\left[y(1+\phi^2)-2x^2\phi-\sqrt{1+4\phi^2(y-x^2)^2}-\log\left(\frac{2(y-x^2)}{1+\sqrt{1+4\phi^2(y-x^2)^2}}\right)\right], & \text{if } y>x^2,\\ +\infty, & \text{if } y\leqslant x^2.\end{cases}$$

As a consequence, by an application of the Contraction Principle with the auxiliary continuous function $f_1(x,y)=x$, we can obtain the rate function for the AR(1) sample mean $\overline{X}_n = n^{-1}\sum_{k=1}^{n} X_k$. Following the same steps from Section 3.1, notice that the infimum

$$\begin{aligned} I_{\overline{X}}(c) &= \inf_{y>x^2}\{H_1(x,y)\,|\,f_1(x,y)=c\} = \inf_{y>c^2} H_1(c,y)\\ &= \inf_{y>c^2}\left\{\frac{1}{2}\left[y(1+\phi^2)-2c^2\phi-\sqrt{1+4\phi^2(y-c^2)^2}-\log\left(\frac{2(y-c^2)}{1+\sqrt{1+4\phi^2(y-c^2)^2}}\right)\right]\right\}\end{aligned}$$

is attained at $y_c = \dfrac{1+c^2(1-\phi^2)}{1-\phi^2}$. Hence, the sequence $(\overline{X}_n)_{n\in\mathbb{N}}$ satisfies the LDP with the rate function

$$I_{\overline{X}}(c) = H_1(c,y_c) = \frac{c^2(1-\phi)^2}{2}, \qquad \text{for } c\in\mathbb{R}.$$

The graphs of $I_{\overline{X}}(\cdot)$ are depicted in Figure 4.1 for three different values of ϕ. Notice that, $I_{\overline{X}}(\cdot)$ has the shape of a parabola.

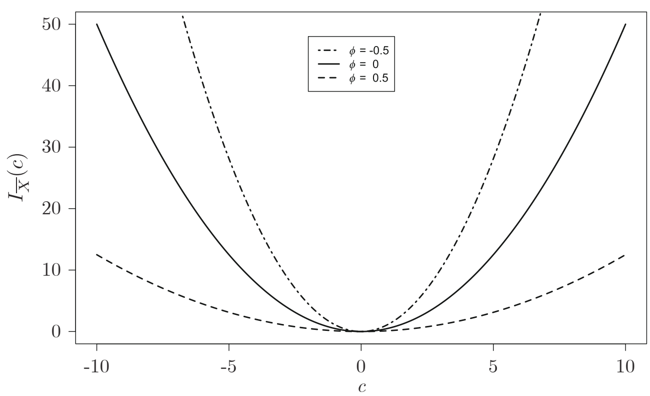


Figure 4.1 Graph of $I_{\overline{X}}(c)$ for $\phi\in\{-0.5,0,0.5\}$ and $c\in[-10,10]$.

4.2 MA(1) process

In this subsection, we present an explicit expression for the rate function of the sample second moment of a random sample following the MA(1) process, defined by the equation

$$Y_j = \varepsilon_j + \theta\, \varepsilon_{j-1}, \quad \text{with } |\theta| < 1 \text{ and } j \in \mathbb{N}. \tag{4.3}$$

Throughout this section, we assume that the innovations $(\varepsilon_j)_{j\geqslant 0}$ are i.i.d., with $\varepsilon_j \sim \mathcal{N}(0,1)$. Then $Y_j \sim \mathcal{N}(0, 1+\theta^2)$, for each $j \in \mathbb{N}$, and the spectral density function associated to $(Y_j)_{j\in\mathbb{N}}$ is given by

$$h_\theta(\omega) = 1 + \theta^2 + 2\theta\cos(\omega), \qquad \text{for } \omega \in \mathbb{T} = [-\pi, \pi).$$

The process $(Y_j)_{j\in\mathbb{N}}$ is stationary for any $\theta \in \mathbb{R}$ (Shumway and Stoffer [40], definition 3.4). In addition to that, the assumption $|\theta| < 1$ in (4.3) ensures that the process is also invertible and that $h_\theta(\cdot)$ is positive for all $\omega \in \mathbb{T}$.

Let us consider $\tilde{\gamma}_n(0) = \dfrac{1}{n}\sum_{k=1}^n Y_k^2$, the sample second moment of a random sample following the MA(1) process described in (4.3). Since the autocovariance function of $(Y_j)_{j\in\mathbb{N}}$ is equal to

$$\gamma_Y(k) = \begin{cases} 1+\theta^2, & \text{if } k = 0, \\ \theta, & \text{if } |k| = 1, \\ 0, & \text{if } |k| > 1, \end{cases}$$

it is known that (Brockwell and Davis [9], section 7.3) $\tilde{\gamma}_n(0) \xrightarrow{n\to+\infty} \gamma_Y(0) = 1+\theta^2$, almost surely. We show here that the sequence $(\tilde{\gamma}_n(0))_{n\in\mathbb{N}}$ satisfies the LDP and we exhibit its explicit rate function, as for this case no explicit rate function has been presented in the literature.

Consider the normalized cumulant generating function $L_n(\lambda) = \dfrac{1}{n}\log \mathbb{E}(e^{n\lambda\tilde{\gamma}_n(0)})$. In this case, the asymptotic distribution of $L_n(\cdot)$ is known (Grenander and Szegö [24]) and we immediately obtain the convergence

$$\lim_{n\to+\infty} L_n(\lambda) = L(\lambda) = \begin{cases} -\dfrac{1}{4\pi}\displaystyle\int_{\mathbb{T}} \log[1 - 2\lambda h_\theta(\omega)]\,\mathrm{d}\omega, & \text{if } \lambda \in \left(-\infty, \dfrac{1}{2M_{h_\theta}}\right), \\ +\infty, & \text{otherwise}, \end{cases}$$

where M_{h_θ} denotes the essential supremum of $h_\theta(\cdot)$, given by $M_{h_\theta} = (1+|\theta|)^2$. As presented in Bercu *et al.* [5] and corollary 1 in Bryc and Dembo [10], $(\tilde{\gamma}_n(0))_{n\in\mathbb{N}}$ satisfies the LDP whose good rate function is the Fenchel-Legendre dual of $L(\cdot)$, given by

$$K_\theta(x) = \begin{cases} \sup_{\lambda < \frac{1}{2M_{h_\theta}}} \left\{ x\lambda + \dfrac{1}{4\pi}\displaystyle\int_{\mathbb{T}} \log[1 - 2\lambda h_\theta(\omega)]\,\mathrm{d}\omega \right\}, & \text{for } x > 0, \\ +\infty, & \text{for } x \leqslant 0. \end{cases} \tag{4.4}$$

Since

$$\begin{aligned} \int_{\mathbb{T}} \log[1 - 2\lambda h_\theta(\omega)]\,\mathrm{d}\omega &= \int_{\mathbb{T}} \log[1 - 2\lambda(1+\theta^2) - 4\lambda\theta\cos(\omega)]\,\mathrm{d}\omega \\ &= 2\pi \log\left[\frac{1 - 2\lambda(1+\theta^2) + \sqrt{(1-2\lambda(1+\theta^2))^2 - 16\lambda^2\theta^2}}{2}\right], \end{aligned}$$

the supremum in (4.4) is attained at

$$\lambda_\theta(x) = \frac{A_\theta(x) + \dfrac{B_\theta(x)}{C_\theta(x)} + C_\theta(x)}{12x^2\left(\theta^2 - 1\right)^2}, \tag{4.5}$$

where

$$A_\theta(x) = 4x\left(x\left(\theta^2+1\right) - \left(\theta^2-1\right)^2\right),$$

$$B_\theta(x) = 4x^2\left(x^2\left(\theta^4+14\theta^2+1\right) + 4x\left(\theta^2+1\right)\left(\theta^2-1\right)^2 + \left(\theta^2-1\right)^4\right),$$

and

$$C_\theta(x) = -(1+i\sqrt{3})\Big[-x^6\left(\theta^6-33\theta^4-33\theta^2+1\right) - 6x^5\left(\theta^2-1\right)^2\left(\theta^4-10\theta^2+1\right) + 6x^4\left(\theta^2-1\right)^4\left(\theta^2+1\right) + x^3\left(\theta^2-1\right)^6 + 3\sqrt{3}\sqrt{c_\theta(x)}\,\Big]^{1/3},$$

with

$$c_\theta(x) = -x^8\left(\theta^2-1\right)^4\Big(4x^4\theta^2 + 32x^3\left(\theta^4+\theta^2\right) + x^2\left(\theta^4+46\theta^2+1\right)\left(\theta^2-1\right)^2 + 6x\left(\theta^2+1\right)\left(\theta^2-1\right)^4 + \left(\theta^2-1\right)^6\Big).$$

Remark 4.1 *Although* $C_\theta(\cdot)$ *appears in a complex form, it can be proved that* $B_\theta(x)/C_\theta(x) + C_\theta(x) \in \mathbb{R}$*, for any* $x > 0$*. In fact,* $\lambda_\theta(x)$ *in (4.5) is one of the solutions from the polynomial equation*

$$\lambda^3\left(4x^2\theta^4 - 8x^2\theta^2 + 4x^2\right) + \lambda^2\left(-4x^2\theta^2 - 4x^2 + 4x\theta^4 - 8x\theta^2 + 4x\right) + \lambda\left(x^2 - 4x\theta^2 - 4x + \theta^4 - 2\theta^2 + 1\right) + x - \theta^2 - 1 = 0,$$

which has three real roots if $x > 0$*. Moreover, we have* $\lambda_\theta(x) < \dfrac{1}{2M_{h_\theta}}$*.*

We conclude that $(\tilde{\gamma}_n(0))_{n\in\mathbb{N}}$ satisfies the LDP with rate function given by

$$K_\theta(x) = x\lambda_\theta(x) + \frac{1}{2}\log\left[\frac{1-2\lambda_\theta(x)(1+\theta^2) + \sqrt{(1-2\lambda_\theta(x)(1+\theta^2))^2 - 16\lambda_\theta(x)^2\theta^2}}{2}\right]$$

$$= \frac{f_\theta(x)}{12x\left(\theta^2-1\right)^2} + \frac{1}{2}\log\left(\frac{1}{2} - \frac{\left(\theta^2+1\right)f_\theta(x)}{12x^2\left(\theta^2-1\right)^2} + \sqrt{\left(\frac{1}{2} - \frac{\left(\theta^2+1\right)f_\theta(x)}{12x^2\left(\theta^2-1\right)^2}\right)^2 - \frac{\theta^2 f_\theta(x)^2}{36x^4\left(\theta^2-1\right)^4}}\right), \tag{4.6}$$

for all $x > 0$ and $K_\theta(x) = +\infty$, for $x \leqslant 0$, with $f_\theta(x) = A_\theta(x) + \dfrac{B_\theta(x)}{C_\theta(x)} + C_\theta(x)$. The graph of $K_\theta(\cdot)$ is illustrated in Figure 4.2 for four different values of θ and $x \in (0, 5]$.

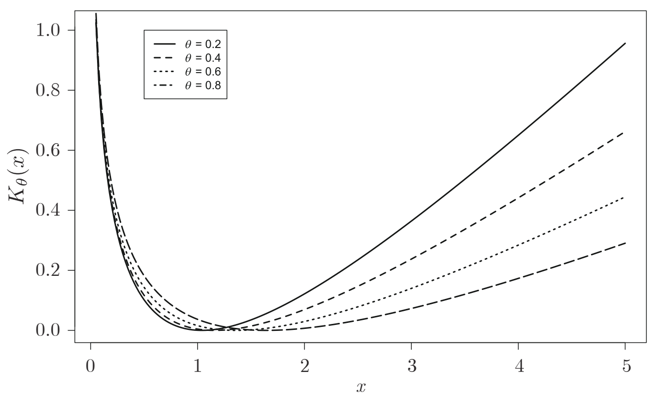


Figure 4.2 Graphs of $K_\theta(x)$ for $\theta \in \{0.2, 0.4, 0.6, 0.8\}$ and $x \in (0, 5]$.

By Proposition 4.1, we may now conclude that $(\boldsymbol{S}_n)_{n\in\mathbb{N}}$ satisfies the LDP with rate function

$$H_2(x,y) = \begin{cases} K_\theta(y-x^2) + \dfrac{x^2}{2(1+\theta)^2}, & \text{if } y > x^2, \\ +\infty, & \text{if } y \leqslant x^2, \end{cases}$$

where $K_\theta(\cdot)$ is given in (4.6). Note that, by the Contraction Principle, the sequence $(f_1(\boldsymbol{S_n}))_{n\in\mathbb{N}} = \left(n^{-1}\sum_{k=1}^n Y_k\right)_{n\in\mathbb{N}}$, where $f_1(x,y)=x$ and $\boldsymbol{S}_n = n^{-1}\left(\sum_{k=1}^n Y_k, \sum_{k=1}^n Y_k^2\right)$, must satisfy the LDP with rate function

$$\begin{aligned} I_{\overline{Y}}(c) &= \inf_{y>x^2}\{H_2(x,y) \mid f_1(x,y)=c\} \\ &= \inf_{y>c^2} H_2(c,y) = \inf_{y>c^2}\left\{K_\theta(y-c^2) + \frac{c^2}{2(1+\theta)^2}\right\}. \end{aligned} \tag{4.7}$$

However, when trying to compute the infimum in (4.7), we face a non-trivial problem.

Fortunately, the LDP for the sample mean of the moving average process has already been given in Burton and Dehling [13]. More generally, the authors considered the sequence

$$X_k = \sum_{i\in\mathbb{Z}} a_{i+k}\,\xi_i, \quad \text{for } k\in\mathbb{Z}, \tag{4.8}$$

with $(\xi_i)_{i\in\mathbb{Z}}$ a sequence of i.i.d. random variables. They proved the LDP under the hypotheses that $(a_i)_{i\in\mathbb{Z}}$ is an absolutely summable sequence and that the moment generating function $\mathbb{E}(e^{t\,\xi_1})$ is finite for all $t\in\mathbb{R}$. A similar approach has been given in Djellout and Guillin [17]. In this paper, the authors proved an analogous result under the hypotheses that the sequence $(\xi_i)_{i\in\mathbb{Z}}$ is bounded and that $\sum_{i\in\mathbb{Z}} a_i^2 < +\infty$.

If we set $a_0=1$, $a_1=\theta$, and $a_i=0$ for $i\in\mathbb{Z}\setminus\{0,1\}$, then (4.8) resumes to $X_k = \xi_{-k} + \theta\,\xi_{-k+1}$. By setting $\xi_{-j}=\varepsilon_j$ for $j\geqslant 0$, it is plain that $X_j = Y_j$ for $j\in\mathbb{N}$, where Y_j is the MA(1) process given in (4.3). Then by theorem 2.1 in Burton and Dehling [13], the Sample Mean $(\overline{Y}_n)_{n\in\mathbb{N}} = \left(n^{-1}\sum_{k=1}^n Y_k\right)_{n\in\mathbb{N}}$ satisfies the LDP with rate function

$$I_{\overline{Y}}(c) = \sup_{\lambda\in\mathbb{R}}\left\{\frac{c\,\lambda}{1+\theta} - \frac{\lambda^2}{2}\right\} = \frac{c^2}{2(1+\theta)^2}, \quad \text{for } c\in\mathbb{R}. \tag{4.9}$$

Comparing (4.9) with (4.7), we note that the infimum on (4.7) is attained at a root of the transcendental equation $K_\theta(y-c^2)=0$. Notice that the graphs that are shown in Figure 4.2 support this claim.

5. Proofs

In this section, we give the proofs of the main results of this work.

5.1 Proof of Lemma 2.1

The proof of Lemma 2.1 is based on the techniques given on page 270 in Jensen [27]. In summary, we use Sylvester's Criterion (Horn and Johnson [26], theorem 7.2.5) to check for the positive definiteness of each leading principal minor of $D_{n,\boldsymbol{\lambda}}$, resorting to the use of an auxiliary function with its corresponding iterates. By Sylvester's Criterion, $D_{n,\boldsymbol{\lambda}}$ is positive definite if, and only if, the leading principal minors of $D_{n,\boldsymbol{\lambda}}$ are positive. Hence, we analyze each one of the leading principal minors of $D_{n,\boldsymbol{\lambda}}$ as follows:

- *1-st Step:* since the first leading principal minor of $D_{n,\boldsymbol{\lambda}}$ is $r_1 = 1-2\lambda_1$, we require that $r_1>0$. As a consequence, since $p = r_1+\phi^2$, we obtain $0<r_1<p \Rightarrow 0<p$.

- *2-nd Step:* the second leading principal minor of $D_{n,\boldsymbol{\lambda}}$ is defined as the determinant

$$\begin{vmatrix} r_1 & q \\ q & p \end{vmatrix} = p\,r_1 - q^2 = \left(p - \frac{q^2}{r_1}\right) r_1. \tag{5.1}$$

Since we already restricted our analysis for $r_1 > 0$, (5.1) requires in addition that $r_2 := p - \dfrac{q^2}{r_1} > 0$.

- *3-rd Step:* the third leading principal minor of $D_{n,\boldsymbol{\lambda}}$ is the determinant

$$\begin{vmatrix} r_1 & q & 0 \\ q & p & q \\ 0 & q & p \end{vmatrix} = p^2\,r_1 - q^2\,r_1 - q^2\,p = \left(p - \cfrac{q^2}{p - \cfrac{q^2}{r_1}}\right)\left(p - \frac{q^2}{r_1}\right) r_1. \tag{5.2}$$

Since we already restricted our analysis for $r_1 > 0$ and $p - \dfrac{q^2}{r_1} > 0$, (5.2) requires that $r_3 := \left(p - \cfrac{q^2}{p - \cfrac{q^2}{r_1}}\right) > 0$.

- *k-th Step:* by induction, the k-th leading principal minor of $D_{n,\boldsymbol{\lambda}}$, for $1 \leqslant k \leqslant n-1$, is the determinant

$$\begin{vmatrix} r_1 & q & 0 & \cdots & 0 \\ q & p & \ddots & \ddots & \vdots \\ 0 & q & \ddots & q & 0 \\ \vdots & \ddots & \ddots & p & q \\ 0 & \cdots & 0 & q & p \end{vmatrix} = r_k\,r_{k-1}\,\cdots\,r_2\,r_1,$$

for $r_2 = p - \dfrac{q^2}{r_1}$, $r_3 = \left(p - \cfrac{q^2}{p - \cfrac{q^2}{r_1}}\right)$ and $r_k = G^{k-1}(r_1)$, where G^k denotes the k-th iterate of $G : (0, +\infty) \to (-\infty, p)$, given by

$$G(a) = p - \frac{q^2}{a}.$$

Since $n \in \mathbb{N}$ is arbitrary, we must require that $G^k(r_1) > 0$, for all $k \in \mathbb{N}$. Without loss of generality, we may assume that $q \neq 0$ (if $q = 0$, then $D_{n,\boldsymbol{\lambda}}$ is a diagonal matrix; this happens if, and only if, $\lambda_2 = -\phi$). Note that $G(\cdot)$ is an increasing concave function that has the following two fixed points

$$R = \frac{1}{2}\left(p - \sqrt{p^2 - 4q^2}\right) \qquad \text{and} \qquad Q = \frac{1}{2}\left(p + \sqrt{p^2 - 4q^2}\right).$$

If $p^2 > 4q^2$, the point named Q is an attractor point and the point named R is a repulsor point. If $p^2 = 4q^2$, then $R = Q = p/2$ is neither an attractor, nor a repulsor point, but if $x > p/2$, then $G^k(x) > p/2$, for all $k \in \mathbb{N}$, and it converges towards $p/2$ as $k \to \infty$, whereas if $x < p/2$, $G^k(x)$ converges towards the region $(0, -\infty)$. If $p^2 < 4q^2$, then the entire graph of $G(\cdot)$ lies below the graph of the identity function in the $\mathbb{R}^2$ plane, so that $G^k(x)$ enters the region $(0, +\infty)$, for large k. For these reasons, let us consider henceforth $p^2 \geqslant 4q^2$.

The problem of knowing when $G^k(r_1) > 0$, for all $k \in \mathbb{N}$, then reduces to knowing whenever $r_1 \geqslant R$. Indeed, if $r_1 > R$, note that every point greater than R converges towards Q and, since $R > 0$, it holds that $G^k(r_1) > R > 0$, for all $k \in \mathbb{N}$. If $r_1 = R$, then $G^k(r_1) = R > 0$, for all $k \in \mathbb{N}$. However, if $r_1 < R$, then there exist $n_0 \in \mathbb{N}$ such that $G^{n_0}(r_1) < 0$. Since $r_1 = 1 - 2\lambda_1 = p - \phi^2$, we get

$$r_1 \geqslant R \Leftrightarrow r_1 \geqslant \frac{p - \sqrt{p^2 - 4q^2}}{2} \Leftrightarrow \sqrt{p^2 - 4q^2} \geqslant p - 2r_1 = p - 2(p - \phi^2) = 2\phi^2 - p. \tag{5.3}$$

If $p > 2\phi^2$, then the right-hand side of (5.3) is negative, implying that $r_1 \geqslant R$. But if $p \leqslant 2\phi^2$,

$$r_1 \geqslant R \Leftrightarrow p^2 - 4q^2 \geqslant \left(2\phi^2 - p\right)^2 \Leftrightarrow p^2 - 4q^2 \geqslant 4\phi^4 - 4\phi^2 p + p^2 \Leftrightarrow \phi^2(p - \phi^2) \geqslant q^2.$$

Therefore, we obtain the domain $\tilde{\mathcal{D}} = \mathcal{D}_1 \cup \tilde{\mathcal{D}}_2$, where

$$\mathcal{D}_1 = \{r_1 > 0,\ p^2 \geqslant 4q^2,\ p > 2\phi^2\} \quad \text{and}$$
$$\tilde{\mathcal{D}}_2 = \{r_1 > 0,\ p^2 \geqslant 4q^2,\ p \leqslant 2\phi^2,\ q^2 \leqslant \phi^2(p - \phi^2)\}.$$

Note that $r_1 > 0$ is equivalent to $p > \phi^2$. Moreover, from

$$0 \geqslant -\left(\phi^2 - \frac{p}{2}\right)^2 = -\phi^4 + 2\phi^2\frac{p}{2} - \frac{p^2}{4} = -\phi^4 + \phi^2 p - \frac{p^2}{4} = \phi^2(p - \phi^2) - \frac{p^2}{4},$$

we conclude that $\phi^2(p - \phi^2) \leqslant \frac{p^2}{4}$. Hence, if $q^2 \leqslant \phi^2(p - \phi^2)$, it follows that $4q^2 \leqslant p^2$. Therefore, if p and q belong to

$$\mathcal{D}_1 = \{p > 2\phi^2,\ p^2 \geqslant 4q^2\} \quad \text{or} \quad \tilde{\mathcal{D}}_2 = \{\phi^2 < p \leqslant 2\phi^2,\ q^2 \leqslant \phi^2(p - \phi^2)\},$$

then $G^k(r_1) > 0$, for all $k \in \mathbb{N}$.

- *n-th Step*: last but not least, the n-th leading principal minor (the determinant) of $D_{n,\boldsymbol{\lambda}}$ is

$$|D_{n,\boldsymbol{\lambda}}| = \begin{vmatrix} r_1 & q & 0 & \cdots & 0 \\ q & p & \ddots & \ddots & \vdots \\ 0 & q & \ddots & q & 0 \\ \vdots & \ddots & \ddots & p & q \\ 0 & \cdots & 0 & q & r_1 \end{vmatrix} = r_1(r_{n-1}\, r_{n-2} \cdots r_2\, r_1) - q^2(r_{n-2} \cdots r_2\, r_1)$$
$$= \left(r_1 - \frac{q^2}{r_{n-1}}\right)(r_{n-1}\, r_{n-2} \cdots r_2\, r_1) = (G^{n-1}(r_1) - \phi^2)(r_{n-1}\, r_{n-2} \cdots r_2\, r_1).$$

On the one hand, if $(\lambda_1, \lambda_2) \in \mathcal{D}_1$, then $p > 2\phi^2 \Rightarrow r_1 = p - \phi^2 > \phi^2$, and two cases are possible: the first is when $R \leqslant r_1 \leqslant Q$, and since $G(x) \geqslant x$, for any point $x \in [R, Q]$, it immediately follows that $G(r_1) \geqslant r_1 > \phi^2$; the second case is when $r_1 > Q$, but then $Q \geqslant p/2 > \phi^2$ and, since $G(\cdot)$ is an increasing function, it follows that $G(r_1) > G(Q) = Q > \phi^2$. Moreover, in the former case, it follows that $G(r_1) \in [r_1, Q]$, whereas in the latter case, $G(r_1) \in (Q, r_1)$. Hence, by induction, it follows that $G^n(r_1) > \phi^2$, for all $n \in \mathbb{N}$.

On the other hand, if $(\lambda_1, \lambda_2) \in \tilde{\mathcal{D}}_2$, it is not always true that $G^{n-1}(r_1) > \phi^2$ for every $n \in \mathbb{N}$ (take, for instance, $\lambda_1 = 1/3$, $\lambda_2 = -1/2$ and $\theta = 99/100$). Additionally, if $q^2 = \phi^2(p - \phi^2)$ and $p \leqslant 2\phi^2$, we have

$$R = \frac{p - |p - 2\phi^2|}{2} = r_1 = p - \phi^2 \leqslant 2\phi^2 - \phi^2 = \phi^2,$$

showing that $G^n(r_1) = G^n(R) = R = r_1 \leqslant \phi^2$, for all $n \in \mathbb{N}$. However, it can be shown that, for

n large enough, we eventually obtain $G^{n-1}(r_1) > \phi^2$, for $(\lambda_1, \lambda_2) \in \mathcal{D}_2$, where $\mathcal{D}_2$ is the set $\tilde{\mathcal{D}}_2$ minus the curve $\{(\lambda_1, \lambda_2) \in \mathbb{R}^2 \mid q^2 = \phi^2(p - \phi^2),\ \phi^2 < p \leqslant 2\phi^2\}$, i.e, the set in (2.14). Indeed, if $(\lambda_1, \lambda_2) \in \mathcal{D}_2$, since $\phi^2 < p \leqslant 2\phi^2$, we obtain

$$\begin{aligned} q^2 < \phi^2(p-\phi^2) &\Leftrightarrow p^2 - 4q^2 > p^2 - 4\phi^2 p + 4\phi^4 = (2\phi^2 - p)^2 \\ &\Rightarrow \sqrt{p^2 - 4q^2} > |2\phi^2 - p| = 2\phi^2 - p \\ &\Leftrightarrow Q = \frac{p + \sqrt{p^2-4q^2}}{2} > \frac{p + 2\phi^2 - p}{2} = \phi^2 \Rightarrow \lim_{n\to+\infty} G^n(r_1) = Q > \phi^2, \end{aligned}$$

so that $\exists\, N \in \mathbb{N}; n \geqslant N \Rightarrow G^n(r_1) > \phi^2$.

Thus, the set $\mathcal{D}_1 \cup \mathcal{D}_2$ is, therefore, the domain where all leading principal minors of $D_{n,\boldsymbol{\lambda}}$ are positive for all $n \geqslant N$, and, consequently, where the matrix $D_{n,\boldsymbol{\lambda}}$ is positive definite. Converting the domains $\mathcal{D}_1$ and $\mathcal{D}_2$ to the (λ_1, λ_2) notation, we obtain the desired expressions given in (2.14).

5.2 Proof of Lemma 2.2

For this proof, consider the measure space $L^\infty(\mathbb{T}) := L^\infty(\mathbb{T}, \mathcal{B}(\mathbb{T}), \nu)$ equipped with the usual norm (Bartle [2], chapter 6)

$$||f||_\infty = \inf\{S^f(N) \,|\, N \in \mathcal{B}(\mathbb{T}),\, \nu(N) = 0\}, \quad \text{where } S^f(N) = \sup\{|h(x)| \,:\, x \notin N\},$$

for any $f \in L^\infty(\mathbb{T})$, and $\nu(\cdot)$ denotes the Lebesgue measure acting on the Borel σ-algebra $\mathcal{B}(\mathbb{T})$ over $\mathbb{T}$. Let m_f and M_f also denote, respectively, the *essential infimum* and *essential supremum* of the function $f(\cdot)$. We recall that the essential supremum of a real-valued function $f(\cdot)$ is defined as the smallest number a for which $f(x) \leqslant a$ except on a set of total length or measure 0; a similar definition holds for the essential infimum (Gray [23], pg. 193).

Let $(\alpha^{\boldsymbol{\lambda}}_{n,k})_{k=1}^n$ represent the sequence of eigenvalues of $T_n(g_\phi)^{1/2}\, T_n(\varphi_{\boldsymbol{\lambda}})\, T_n(g_\phi)^{1/2}$, with $g_\phi(\cdot)$ denoting the spectral density function defined in (1.3), and $\varphi_{\boldsymbol{\lambda}}(\cdot)$ the function given in (2.4). If $(\lambda_1, \lambda_2) \in \mathcal{D}$, then Lemma 2.1 guarantees that $\alpha^{\boldsymbol{\lambda}}_{n,k} < 1/2$, for all $1 \leqslant k \leqslant n$ and $n \geqslant N$, for some $N \in \mathbb{N}$. Hence, from (2.10) and (2.11) it follows that

$$L_n(\lambda_1, \lambda_2) = -\frac{1}{2n}\sum_{k=1}^{n} \log\left(1 - 2\alpha^{\boldsymbol{\lambda}}_{n,k}\right), \quad \text{for } (\lambda_1, \lambda_2) \in \mathcal{D} \text{ and } n \text{ large enough.} \tag{5.4}$$

If $(\lambda_1, \lambda_2) \notin \mathcal{D}$, we have immediately $L_n(\lambda_1, \lambda_2) = +\infty$. Therefore, we only need to consider the case when (λ_1, λ_2) belongs to $\mathcal{D}$, because then $L_n(\lambda_1, \lambda_2)$ is finite for n large enough and it is given by (5.4).

Let $\varphi_{\boldsymbol{\lambda}}\, g_\phi : \mathbb{T} \to \mathbb{R}$ be the function defined by

$$(\varphi_{\boldsymbol{\lambda}}\, g_\phi)(\omega) = \varphi_{\boldsymbol{\lambda}}(\omega)\, g_\phi(\omega) = \frac{\lambda_1 + \lambda_2\cos(\omega)}{1 + \phi^2 - 2\phi\cos(\omega)}. \tag{5.5}$$

Since $(\varphi_{\boldsymbol{\lambda}}\, g_\phi)(\cdot)$ is continuous and bounded in $\mathbb{T}$, it attains a maximum and a minimum in that interval, and by this reason, it follows that $m_{\varphi_{\boldsymbol{\lambda}}\, g_\phi} = \min_{\omega\in\mathbb{T}}\{(\varphi_{\boldsymbol{\lambda}}\, g_\phi)(\omega)\}$ and $M_{\varphi_{\boldsymbol{\lambda}}\, g_\phi} = \max_{\omega\in\mathbb{T}}\{(\varphi_{\boldsymbol{\lambda}}\, g_\phi)(\omega)\}$. Furthermore, as $\varphi_{\boldsymbol{\lambda}},\, g_\phi \in L^\infty(\mathbb{T})$, it holds that (see Avram [1])

$$|\alpha^{\boldsymbol{\lambda}}_{n,k}| \leqslant ||\varphi_{\boldsymbol{\lambda}}||_\infty ||g_\phi||_\infty, \quad \text{for all } 1 \leqslant k \leqslant n \text{ and } n \in \mathbb{N}. \tag{5.6}$$

As $\dfrac{\mathrm{d}}{\mathrm{d}\omega}(\varphi_{\boldsymbol{\lambda}}\, g_\phi)(\omega) = -\dfrac{\lambda_2\sin(\omega)}{1+\phi^2-2\phi\cos(\omega)} - \dfrac{2\phi(\lambda_1 + \lambda_2\cos(\omega))\sin(\omega)}{(1+\phi^2-2\phi\cos(\omega))^2}$, we note that $(\varphi_{\boldsymbol{\lambda}}\, g_\phi)(\omega)$ has two critical points, one at $\omega = -\pi$, and another at $\omega = 0$. Moreover,

- if $\lambda_2 < -2\phi\lambda_1/(1+\phi^2)$, then $\dfrac{d^2}{d\omega^2}(\varphi_{\boldsymbol\lambda}\, g_\phi)(-\pi) < 0$ and $\dfrac{d^2}{d\omega^2}(\varphi_{\boldsymbol\lambda}\, g_\phi)(0) > 0$;
- if $\lambda_2 > -2\phi\lambda_1/(1+\phi^2)$, then $\dfrac{d^2}{d\omega^2}(\varphi_{\boldsymbol\lambda}\, g_\phi)(-\pi) > 0$ and $\dfrac{d^2}{d\omega^2}(\varphi_{\boldsymbol\lambda}\, g_\phi)(0) < 0$;
- if $\lambda_2 = -2\phi\lambda_1/(1+\phi^2)$, then $(\varphi_{\boldsymbol\lambda}\, g_\phi)(\omega) = \lambda_1/(1+\phi^2)$ is constant.

Therefore, since $(\varphi_{\boldsymbol\lambda}\, g_\phi)(-\pi) = \dfrac{\lambda_1-\lambda_2}{1+\phi^2+2\phi}$ and $(\varphi_{\boldsymbol\lambda}\, g_\phi)(0) = \dfrac{\lambda_1+\lambda_2}{1+\phi^2-2\phi}$, we conclude that

$$m_{\varphi_{\boldsymbol\lambda}\, g_\phi} = \begin{cases} \dfrac{\lambda_1+\lambda_2}{1+\phi^2-2\phi}, & \text{if } \lambda_2 < -\dfrac{2\phi\lambda_1}{1+\phi^2}, \\ \dfrac{\lambda_1-\lambda_2}{1+\phi^2+2\phi}, & \text{if } \lambda_2 \geqslant -\dfrac{2\phi\lambda_1}{1+\phi^2}, \end{cases}$$

and

$$M_{\varphi_{\boldsymbol\lambda}\, g_\phi} = \begin{cases} \dfrac{\lambda_1-\lambda_2}{1+\phi^2+2\phi}, & \text{if } \lambda_2 < -\dfrac{2\phi\lambda_1}{1+\phi^2}, \\ \dfrac{\lambda_1+\lambda_2}{1+\phi^2-2\phi}, & \text{if } \lambda_2 \geqslant -\dfrac{2\phi\lambda_1}{1+\phi^2}. \end{cases}$$

Let us separate the rest of this proof in two cases: when $(\lambda_1,\lambda_2) \in \mathcal{D}^\circ$ and when $(\lambda_1,\lambda_2) \in \mathcal{D}_1 \cap \partial\mathcal{D}$.

- **Case 1** considering first the case in which $(\lambda_1,\lambda_2) \in \mathcal{D}^\circ$, it follows that $2|\phi+\lambda_2| < 1+\phi^2-2\lambda_1$, whence

$$-\left(\frac{1+\phi^2-2\lambda_1}{2}\right) < \phi+\lambda_2 < \frac{1+\phi^2-2\lambda_1}{2}. \tag{5.7}$$

From the left-hand side of (5.7), we get $\dfrac{\lambda_1-\lambda_2}{1+\phi^2+2\phi} < \dfrac{1}{2}$, while from the right-hand side of (5.7) we obtain $\dfrac{\lambda_1+\lambda_2}{1+\phi^2-2\phi} < \dfrac{1}{2}$. Hence, we conclude that $M_{\varphi_{\boldsymbol\lambda}\, g_\phi} < 1/2$. At the same time, from

$$||\varphi_{\boldsymbol\lambda}||_\infty ||g_\phi||_\infty \geqslant ||\varphi_{\boldsymbol\lambda}\, g_\phi||_\infty = \max\{|M_{\varphi_{\boldsymbol\lambda}\, g_\phi}|, |m_{\varphi_{\boldsymbol\lambda}\, g_\phi}|\} \geqslant -m_{\varphi_{\boldsymbol\lambda}\, g_\phi},$$

we conclude that $m_{\varphi_{\boldsymbol\lambda}\, g_\phi} \geqslant -||\varphi_{\boldsymbol\lambda}||_\infty ||g_\phi||_\infty$. Therefore, we just proved that

$$[m_{\varphi_{\boldsymbol\lambda}\, g_\phi}, M_{\varphi_{\boldsymbol\lambda}\, g_\phi}] \subseteq [-||\varphi_{\boldsymbol\lambda}||_\infty ||g_\phi||_\infty, 1/2). \tag{5.8}$$

The denominator in the left-hand side of (5.5) satisfies

$$\inf_{\omega\in\mathbb{T}} |1+\phi^2-2\phi\cos(\omega)| = \min\{1+\phi^2-2\phi, 1+\phi^2+2\phi\} > 0, \quad \text{for all} \quad \phi \in (-1,1).$$

Then, by theorem 5.1 in Tyrtyshnikov [41], if F is any arbitrary continuous function with bounded support (i.e., the set of those $x \in \mathbb{R}$ for which $F(x) \neq 0$ is bounded), it follows that

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{k=1}^{n} F(\alpha_{n,k}^{\boldsymbol\lambda}) = \frac{1}{2\pi} \int_{\mathbb{T}} (F \circ (\varphi_{\boldsymbol\lambda}\, g_\phi))(\omega)\, d\omega. \tag{5.9}$$

In particular, the latter convergence applies itself when considering the continuous function $F : [-||\varphi_{\boldsymbol\lambda}||_\infty ||g_\phi||_\infty, 1/2) \to \mathbb{R}$ defined by

$$F(x) = -\frac{\log(1-2x)}{2}.$$

Indeed, from (5.6) and (5.8), combined with the result of Lemma 2.1, we conclude that $F(\cdot)$ has bounded support and that $F(\alpha_{n,k}^{\boldsymbol\lambda})$ are finite, for every $1 \leqslant k \leqslant n$ and n large enough. Besides that, $(F \circ (\varphi_{\boldsymbol\lambda}\, g_\phi))(\omega) = \log[1-2\,(\varphi_{\boldsymbol\lambda}\, g_\phi)(\omega)]$ is finite, for every $\omega \in \mathbb{T}$, due to (5.8). Therefore, the two sides of (5.9) are well defined and such convergence holds, giving

$$\begin{aligned}\lim_{n\to+\infty} L_n(\lambda_1,\lambda_2) &= \lim_{n\to+\infty} -\frac{1}{2n}\sum_{k=1}^{n}\log(1-2\alpha_{n,k}^{\boldsymbol{\lambda}}) = \lim_{n\to+\infty}\frac{1}{n}\sum_{k=1}^{n} F(\alpha_{n,k}^{\boldsymbol{\lambda}})\\ &= \frac{1}{2\pi}\int_{\mathbb{T}}(F\circ(\varphi_{\boldsymbol{\lambda}}\, g_\phi))(\omega)\,d\omega = -\frac{1}{4\pi}\int_{\mathbb{T}}\log(1-2\,(\varphi_{\boldsymbol{\lambda}}\, g_\phi)(\omega))\,d\omega\\ &= -\frac{1}{2}\log\left(\frac{1+\phi^2-2\lambda_1+\sqrt{(1+\phi^2-2\lambda_1)^2-4(\phi+\lambda_2)^2}}{2}\right)\\ &= -\frac{1}{2}\log\Phi(\lambda_1,\lambda_2,\phi),\end{aligned}$$

where the penultimate equality was achieved using equation 4.224(9) in Gradshteyn and Ryzhik [22].

- **Case 2** note that, for any $(\lambda_1,\lambda_2)\in\mathcal{D}$ it holds that

$$\begin{aligned}L_n(\lambda_1,\lambda_2) &= -\frac{1}{2n}\sum_{k=1}^{n}\log(1-2\alpha_{n,k}^{\boldsymbol{\lambda}}) = -\frac{1}{2n}\log\left[\prod_{k=1}^{n}(1-2\alpha_{n,k}^{\boldsymbol{\lambda}})\right]\\ &= -\frac{1}{2n}\log\det(I_n-2\,T_n(g_\phi)^{1/2}\,T_n(\varphi_{\boldsymbol{\lambda}})\,T_n(g_\phi)^{1/2})\\ &= -\frac{1}{2n}\log[\det(D_{n,\boldsymbol{\lambda}})\det(T_n(g_\phi))].\end{aligned}$$

Since the inverse matrix of $T_n(g_\phi)$ is the tridiagonal matrix

$$T_n(g_\phi)^{-1} = \begin{bmatrix} 1 & -\phi & 0 & \cdots & 0\\ -\phi & 1+\phi^2 & \ddots & \ddots & \vdots\\ 0 & -\phi & \ddots & -\phi & 0\\ \vdots & \ddots & \ddots & 1+\phi^2 & -\phi\\ 0 & \cdots & 0 & -\phi & 1\end{bmatrix},$$

from theorem 3.6 in Karling *et al.* [28], we obtain $\det(T_n(g_\phi)) = \det(T_n(g_\phi)^{-1})^{-1} = (1-\phi^2)^{-1}$, which is independent of n. Moreover, if $(\lambda_1,\lambda_2)\in\mathcal{D}_1\cap\partial\mathcal{D}$, then using the notation from (2.13) and the lemmas 3.1 and 3.5 by Karling *et al.* [28], we have $p^2=4q^2$ and

$$\begin{aligned}\det(D_{n,\boldsymbol{\lambda}}) &= r_1\left[1+(n-1)\left(\frac{2r_1-p}{p}\right)\right]\left(\frac{p}{2}\right)^{n-1} - q^2\left[1+(n-2)\left(\frac{2r_1-p}{p}\right)\right]\left(\frac{p}{2}\right)^{n-2}\\ &= r_1\left[1+(n-1)\left(\frac{2r_1-p}{p}\right)\right]\left(\frac{p}{2}\right)^{n-1} - \frac{p}{2}\left[1+(n-2)\left(\frac{2r_1-p}{p}\right)\right]\left(\frac{p}{2}\right)^{n-1}\\ &= \left[\frac{(2r_1-p)(2r_1+p)}{2p}+(n-2)\frac{(2r_1-p)^2}{2p}\right]\left(\frac{p}{2}\right)^{n-1} = \rho_{n-2}\left(\frac{1+\phi^2-2\lambda_1}{2}\right)^{n-2},\end{aligned}$$

where

$$\rho_n = \frac{(1-\phi^2-2\lambda_1)(3+\phi^2-6\lambda_1)}{4} + \frac{(1-\phi^2-2\lambda_1)^2}{4}\,n \sim O(n).$$

Hence, it follows that

$$\lim_{n\to\infty} L_n(\lambda_1,\lambda_2) = -\frac{1}{2}\log\left(\frac{1+\phi^2-2\lambda_1}{2}\right) = -\frac{1}{2}\log\Phi(\lambda_1,\lambda_2,\phi).$$

5.3 Proof of Proposition 2.1

Consider the functions $J_1(\cdot,\cdot)$ and $J_2(\cdot,\cdot)$ given, respectively, by (2.17) and (2.18).

(1) If (x,y) is in $\mathcal{C}_\phi$, it holds that $y = \frac{\pm\sqrt{x(x\phi^2-1)}}{|\phi|}$. Hence, the first statement of Proposition 2.1 follows by making this substitution in (2.17) and (2.18) to check that the functions match at $\mathcal{C}_\phi$.

(2) The gradients of $J_1(\cdot,\cdot)$ and $J_2(\cdot,\cdot)$ are respectively given by

$$\nabla J_1(x,y) = \left(\frac{x^3\left(\phi^2+1\right) - x^2 - xy^2\left(\phi^2+1\right) - y^2}{2\left(x^3 - xy^2\right)}, \frac{y}{x^2-y^2} - \phi\right)$$

and

$$\nabla J_2(x,y) = \left(\frac{1}{2} - \frac{y^2\phi^2}{2x^2}, \frac{\phi(y\phi - x)}{x}\right).$$

Thus, if (x,y) is in $\mathcal{C}_\phi$, after some computations we end up with

$$\nabla J_1\left(x, \frac{\pm\sqrt{x(x\phi^2-1)}}{|\phi|}\right) = \nabla J_2\left(x, \frac{\pm\sqrt{x(x\phi^2-1)}}{|\phi|}\right) = \left(\frac{1+x(1-\phi^2)}{2x}, \frac{\pm\sqrt{x(x\phi^2-1)}}{|\phi|}\right).$$

(3) Let $H_1(\cdot,\cdot)$ and $H_2(\cdot,\cdot)$ represent the Hessian matrices associated to $J_1(\cdot,\cdot)$ and $J_2(\cdot,\cdot)$, respectively. Then, for a generic point (x,y) in the domain of these functions, we have

$$H_1(x,y) = \begin{bmatrix} \frac{x^4+4x^2y^2-y^4}{2(x^3-xy^2)^2} & -\frac{2xy}{(x^2-y^2)^2} \\ -\frac{2xy}{(x^2-y^2)^2} & \frac{x^2+y^2}{(x^2-y^2)^2} \end{bmatrix} \quad \text{and} \quad H_2(x,y) = \begin{bmatrix} \frac{y^2\phi^2}{x^3} & -\frac{y\phi^2}{x^2} \\ -\frac{y\phi^2}{x^2} & \frac{\phi^2}{x} \end{bmatrix}.$$

The eigenvalues of $H_1(x,y)$ are given by

$$\frac{3x^4+6x^2y^2-y^4 \pm \sqrt{x^8+60x^6y^2+6x^4y^4-4x^2y^6+y^8}}{4\left(x^3-xy^2\right)^2}, \tag{5.10}$$

whereas, the eigenvalues of $H_2(x,y)$ are 0 and $\frac{\phi^2\left(x^2+y^2\right)}{x^3}$. On the one hand, (5.10) is positive for all $(x,y) \in \mathbb{R}^2$ such that $|y| < x$. Hence, $J_1(\cdot,\cdot)$ is a strictly convex function. On the other hand, since one of the eigenvalues of $H_2(x,y)$ is equal to zero, only convexity is guaranteed for $J_2(x,y)$, provided that $|y| < x$.

5.4 Proof of Proposition 2.2

Let $J : \mathbb{R}^2 \to \mathbb{R}$ denote the Fenchel-Legendre dual of $L(\cdot,\cdot)$, defined by the supremum

$$\begin{aligned} J(x,y) &= \sup_{(\lambda_1,\lambda_2)\in\mathbb{R}^2} \left\{x\lambda_1 + y\lambda_2 - L(\lambda_1,\lambda_2)\right\} \\ &= \sup_{(\lambda_1,\lambda_2)\in\mathcal{D}} \left\{x\lambda_1 + y\lambda_2 + \frac{1}{2}\log\left(\frac{1+\phi^2-2\lambda_1+\sqrt{(1+\phi^2-2\lambda_1)^2-4(\phi+\lambda_2)^2}}{2}\right)\right\}. \end{aligned} \tag{5.11}$$

To explicitly compute $J(\cdot,\cdot)$, consider the auxiliary function $K : \mathcal{D} \to \mathbb{R}$, defined by

$$K(\lambda_1,\lambda_2) = x\lambda_1 + y\lambda_2 - L(\lambda_1,\lambda_2), \quad \text{with } (x,y) \in \mathbb{R}^2.$$

We first note that $K(\cdot,\cdot)$ is a concave function, as it is the sum of concave functions. Therefore, the supremum in (5.11) is attained at the boundary of $\mathcal{D}$ or an interior point of this set. The partial derivatives of $K(\cdot,\cdot)$ are

$$K_{\lambda_1}(\lambda_1,\lambda_2) = x - \frac{1}{\sqrt{1-4\lambda_1+4\lambda_1^2-4\lambda_2^2-8\phi\lambda_2-2\phi^2(2\lambda_1+1)+\phi^4}}$$

and

$$K_{\lambda_2}(\lambda_1, \lambda_2) = y - \frac{2(\phi + \lambda_2)}{\ell(\lambda_1, \lambda_2, \phi)},$$

where

$$\ell(\lambda_1, \lambda_2, \phi) = \left(1 + \phi^2 - 2\lambda_1 + \sqrt{(1 + \phi^2 - 2\lambda_1)^2 - 4(\phi + \lambda_2)^2}\right) \sqrt{(1 + \phi^2 - 2\lambda_1)^2 - 4(\phi + \lambda_2)^2}.$$

Provided that $x > |y|$, the only solution to the system of equations

$$\begin{cases} K_{\lambda_1}(\lambda_1, \lambda_2) = 0, \\ K_{\lambda_2}(\lambda_1, \lambda_2) = 0, \end{cases}$$

is given by $\lambda_1^* = \dfrac{1 + \phi^2}{2} - \dfrac{x^2 + y^2}{2x(x^2 - y^2)}$ and $\lambda_2^* = \dfrac{y}{x^2 - y^2} - \phi$. Let us divide the computation of $J(\cdot, \cdot)$ into three cases:

- **Case 1** if $(x, y) \in \mathcal{A}_\phi^\circ$, then two subcases are possible since $\mathcal{A}_\phi^\circ = \mathcal{A}_\phi^1 \cup \mathcal{A}_\phi^2$, where

$$\mathcal{A}_\phi^1 := \left\{(x, y) \in \mathbb{R}^2 \,:\, \frac{x^2(2x\phi^2 - 1)}{1 + 2x\phi^2} < y^2 < x^2\right\}$$

and

$$\mathcal{A}_\phi^2 := \left\{(x, y) \in \mathbb{R}^2 \,:\, \frac{x(x\phi^2 - 1)}{\phi^2} < y^2 \leqslant \frac{x^2(2x\phi^2 - 1)}{(1 + 2x\phi^2)}\right\}.$$

If $(x, y) \in \mathcal{A}_\phi^1$, then $(\lambda_1^*, \lambda_2^*) \in \mathcal{D}^1$, whereas if $(x, y) \in \mathcal{A}_\phi^2$, then $(\lambda_1^*, \lambda_2^*) \in \mathcal{D}^2$. In either of these cases,

$$J(x, y) = K(\lambda_1^*, \lambda_2^*) = \frac{1}{2}\left[x(1 + \phi^2) - 1 - 2y\phi + \log\left(\frac{x}{x^2 - y^2}\right)\right].$$

- **Case 2** if $(x, y) \in \mathcal{B}_\phi$, note that $|y| < x$ together with

$$y^2\phi^2 \leqslant x(x\phi^2 - 1) \Leftrightarrow y^2 - x^2 \leqslant -\frac{x}{\phi^2} \Rightarrow \frac{1}{x^2 - y^2} \leqslant \frac{\phi^2}{x} \Rightarrow \frac{x^2 + y^2}{x(x^2 - y^2)} \leqslant \frac{\phi^2(x^2 + y^2)}{x^2}$$

implies that

$$\lambda_1^* = \frac{1 + \phi^2}{2} - \frac{x^2 + y^2}{2x(x^2 - y^2)} \geqslant \frac{1 + \phi^2}{2} - \frac{\phi^2(x^2 + y^2)}{2x^2} \geqslant \frac{1 + \phi^2}{2} - \phi^2 = \frac{1 - \phi^2}{2}. \tag{5.12}$$

In particular, $(\lambda_1^*, \lambda_2^*)$ cannot belong to $\mathcal{D}_1$. Then note that

$$(\phi + \lambda_2^*)^2 - \phi^2(1 - 2\lambda_1^*) = \frac{(\phi^2 + x\phi^4)y^4 + (x - 2x^3\phi^4)y^2 + x^5\phi^4 - x^4\phi^2}{x\,(x^2 - y^2)^2}. \tag{5.13}$$

The denominator in (5.13) is positive in $\mathcal{B}_\phi$ and the numerator is a quadratic polynomial function in terms of y^2. Provided that $x \geqslant 1/\phi^2$, this polynomial has four real roots, namely, the solutions of

$$y^2 = \frac{x^3\phi^2}{1 + x\phi^2} \qquad \text{and} \qquad y^2 = x\left(x - \frac{1}{\phi^2}\right).$$

The software Wolfram Mathematica can handle such analytical expressions, assuring that indeed (5.13) is positive when $y^2 \leqslant x(x - 1/\phi^2)$, i.e., when $(x, y) \in \mathcal{B}_\phi$. This proves that $(\phi + \lambda_2^*)^2 \geqslant \phi^2(1 - 2\lambda_1^*)$, excluding the possibility of $(\lambda_1^*, \lambda_2^*)$ to belong to $\mathcal{D}_2$. In view of that, the supremum in (5.11) is attained at a point over the curve $\partial\mathcal{D}\cap\partial\mathcal{D}_2$. Since for any $\boldsymbol{\lambda}=(\lambda_1,\lambda_2)\in\partial\mathcal{D}\cap\partial\mathcal{D}_2$ it holds that $(\phi+\lambda_2)^2=\phi^2(1-2\lambda_1)$ and $(1-\phi^2)/2\leqslant\lambda_1\leqslant 1/2$, we have

$$\frac{1}{2}\log\left(\frac{1+\phi^2-2\lambda_1+\sqrt{(1+\phi^2-2\lambda_1)^2-4(\phi+\lambda_2)^2}}{2}\right)=\log|\phi|,$$

so that (5.11) is equal to

$$\begin{aligned}\sup_{(\lambda_1,\lambda_2)\in\mathcal{D}}K(\lambda_1,\lambda_2) &= \sup_{(\lambda_1,\lambda_2)\in\partial\mathcal{D}\cap\partial\mathcal{D}_2}\{x\lambda_1+y\lambda_2+\log|\phi|\}\\ &= \sup_{\lambda_2\in[-\phi-\phi^2,-\phi+\phi^2]}\left\{\lambda_2\left(y-\frac{x}{\phi}\right)-x\frac{\lambda_2^2}{2\phi^2}\right\}+\log|\phi|.\end{aligned}$$

Then we note that $\lambda_2\left(y-\frac{x}{\phi}\right)-x\frac{\lambda_2^2}{2\phi^2}$ is a quadratic function with respect to λ_2 and its maximum is attained when $\lambda_2^{**}=\frac{y\phi^2}{x}-\phi$. Since $|y|<x$, it follows that $\lambda_2^{**}\in[-\phi-\phi^2,$ $-\phi+\phi^2]$. Hence,

$$J(x,y)=\lambda_2^{**}\left(y-\frac{x}{\phi}\right)-x\frac{\lambda_2^{**2}}{2\phi^2}+\log|\phi|=\frac{\phi^2(y-\frac{x}{\phi})^2}{2x}+\log|\phi|=\frac{(\phi y-x)^2}{x}+\log|\phi|.$$

- **Case 3** If $|y|\geqslant x$, since $K(\lambda_1,\lambda_2)$ is unbounded, we have $J(x,y)=+\infty$.

5.5 Proof of Theorem 2.1

Since Condition A is valid for any $\phi\in(-1,1)$, the upper and lower bounds given in (2.21) and (2.22), respectively, are a direct consequence from the Gärtner-Ellis' theorem (items (a) and (b) from theorem 2.3.6 in Dembo and Zeitouni [16]). The statement that $J(\cdot,\cdot)$ is a good rate function follows from lemma 2.3.9 in Dembo and Zeitouni [16].

Since $J_1(\cdot,\cdot)$ is strictly convex, every point in its domain is exposed. However, the convexity of $J_2(\cdot,\cdot)$ is not sufficient for a similar conclusion. In fact, consider a fixed point $(a,b)\in\mathcal{B}_\phi$, then theorem 25.1 in Rockafellar [37] ensures that

$$\langle\nabla J_2(a,b),(a-x,b-y)\rangle\geqslant J_2(a,b)-J_2(x,y),\quad\forall\ (x,y),$$

the gradient of $J_2(\cdot,\cdot)$ is the only possible exposing hyperplane associated to (a,b). But, as

$$\langle\nabla J_2(a,b),(a-x,b-y)\rangle-(J_2(a,b)-J_2(x,y))=\frac{\phi^2(bx-ay)^2}{2a^2x}$$

vanishes for $y=\frac{bx}{a}$, the strict inequality in (2.20) does not hold, implying that (a,b) cannot be an exposed point. Geometrically, the above shows that for each point $(a,b)\in\mathcal{B}_\phi$, there exists a line passing through (a,b) such that (2.20) fails. This proves that the only exposed points of $J(\cdot,\cdot)$ are the ones defined by (2.23).

6. Conclusion

In this work, we have shown a weaker version of the LDP for the sequence $(\boldsymbol{\mathcal{W}}_n)_{n\geqslant 2}$, given in (1.4), for the bivariate case (see Theorem 2.1). We also have presented the explicit rate function for the upper and lower bounds (see Proposition 2.2). The same technique to find such properties is not restricted to the AR(1) process. There may exist other classes of processes that can be explored as well. If we take another process $(Z_j)_{j\in\mathbb{N}}$ which still has a multivariate

Gaussian distribution, equipped with another spectral density function, other than the one given in (1.3), the proposed technique may remain valid. As for the AR(1) case, the LDP is, however, not always guaranteed and in most cases, the rate function is hard to compute. This difficulty mainly arises when trying to compute a closed form for the Fenchel-Legendre transform. Besides that, obtaining a similar convergence result, as given in Lemma 2.2, for another class of Gaussian processes remains an intriguing problem. A remarkable class of processes that requires a more sophisticated approach is the class of MA(1) processes, which was not covered in this work when evaluating the LDP for the random vectors $(\boldsymbol{\mathcal{W}}_n)_{n\geqslant 2}$ (see Karling *et al.* [28]).

In Section 3, we have presented three important particular examples by using the previous reasoning from Section 2, together with a weak version of the Contraction Principle (see Theorem 3.1). Two of these examples were already known from Bercu *et al.* [5] and Bryc and Smolenski [11] for univariate sequences. Here we have obtained them as a continuous transformation of the random vector $\boldsymbol{\mathcal{W}}_n$, given in (1.4). In Subsection 3.2, we have presented a result that we believe is new in the literature. In Subsection 3.3, the LDP for the Yule-Walker estimator was obtained, via Theorem 3.1 and Remark 3.1, returning the same result as in Bercu *et al.* [5]. The approach that was used here, first proving the large deviations properties for bivariate random vectors and then particularizing to univariate random sequences, has recently been used with continuous stochastic processes by Bercu and Richou [6], where the authors investigated the LDP of the maximum likelihood estimates for the Ornstein-Uhlenbeck process with a shift. A similar approach was subsequently used by the same authors in Bercu and Richou [7], allowing them to circumvent the classical difficulty of non-steepness.

In Section 4, we have provided the LDP for the sequence of bivariate S^2-mean, for both AR(1) and MA(1) processes. For the AR(1) process, the computations were simple and the previous technique of proving the LDP for the bivariate random vector $\boldsymbol{\mathcal{W}}_n$ was extremely helpful. Nevertheless, we found some issues when dealing with the MA(1) process due to the complexity of the computations involved. The same technique explored above may perhaps be available for general AR(d) processes with Gaussian innovations. This is an important issue that remains to be explored in the future.

Acknowledgements

The authors would like to thank the Editor, Associate Editor, and the anonymous Referee for the numerous comments and suggestions that improved this work. They also would like to express their sincere thanks to Dr. Bernard Bercu for indicating valuable references from the Large Deviations theory.

M.J. Karling was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)-Brazil (Grant No. 1736629) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)-Brazil (Grant No. 170168/2018-2). A.O. Lopes' research was partially supported by CNPq-Brazil (Grant No. 304048/2016-0). S.R.C. Lopes' research was partially supported by CNPq-Brazil (Grant No. 303453/2018-4).

References

[1] Avram, F., On bilinear forms in Gaussian random variables and Toeplitz matrices, Probability Theory and Related Fields, 1988, 79(1): 37–45.

[2] Bartle, R. G., The elements of integration and Lebesgue measure, John Wiley & Sons, New York, 1995.

[3] Bercu, B., On large deviations in the Gaussian autoregressive process: Stable, unstable and explosive case, Bernoulli, 2001, 7(2): 299–316.

[4] Bercu, B., Gamboa, F. and Lavielle, M., Sharp large deviations for Gaussian quadratic forms with applications,

ESAIM: Probability and Statistics, 2000, 4(1): 1–24.

[5] Bercu, B., Gamboa, F. and Rouault, A., Large deviations for quadratic forms of stationary Gaussian processes, Stochastic Processes and their Applications, 1997, 71(1): 75–90.

[6] Bercu, B. and Richou, A., Large deviations for the Ornstein-Uhlenbeck with shift, Advances in Applied Probability, 2015, 47(3): 880–901.

[7] Bercu, B. and Richou, A., Large deviations for the Ornstein-Uhlenbeck process without tears, Statistics & Probability Letters, 2017, 123: 45–55.

[8] Bickel, P. J. and Doksum, K. A., Mathematical Statistics: Basic Ideas and Selected Topics, vol. 1, 2nd ed., Prentice-Hall, Upper Saddle River, 2001.

[9] Brockwell, P. J. and Davis, R. A., Time Series: Theory and Methods, 2nd ed., Springer, New York, 1991.

[10] Bryc, W. and Dembo, A., Large deviations for quadratic functionals of Gaussian processes, Journal of Theoretical Probability, 1997, 10(2): 307–332.

[11] Bryc, W. and Smolenski, W., On the large deviation principle for a quadratic functional of the autoregressive process, Statistics & Probability Letters, 1993, 17(4): 281–285.

[12] Bucklew, J. A., Large Deviation Techniques in Decision, Simulation, and Estimation, John Wiley & Sons, New York, 1990.

[13] Burton, R. M. and Dehling, H., Large deviations for some weakly dependent random processes, Statistics & Probability Letters, 1990, 9(5): 397–401.

[14] Carmona, S. C., Landim, C., Lopes, A. O. and Lopes, S. R. C., A level 1 large-deviation principle for the autocovariances of uniquely ergodic transformations with additive noise, Journal of Statistical Physics, 1998, 91: 395–421.

[15] Carmona, S. C. and Lopes, A. O., Large deviations for expanding transformations with additive white noise, Journal of Statistical Physics, 2000, 98: 1311–1333.

[16] Dembo, A. and Zeitouni, O., Large Deviations Techniques and Applications, 2nd ed., Springer-Verlag, New York, 2010.

[17] Djellout, H. and Guillin, A., Large and moderate deviations for moving average processes, Annales de la Faculté des Sciences de Toulouse, 2001, X(1): 23–31.

[18] Donsker, M. D. and Varadhan, S. R. S., Large deviations for stationary Gaussian processes, Communications in Mathematical Physics, 1985, 97: 187–210.

[19] Ellis, R. S., Entropy, Large Deviations, and Statistical Mechanics, 2nd ed., Springer-Verlag, New York, 1985.

[20] Fayolle, G. and De La Fortelle, A., Entropy and large deviations for discrete-time Markov chains, Problems of Information Transmission, 2002, 38(4): 354–367.

[21] Ferreira, H. H., Lopes, A. O. and Lopes, S. R. C., Decision theory and large deviations for dynamical hypotheses tests: The Neyman-Pearson Lemma, Min-max and Bayesian tests, Journal of Dynamics and Games, 2022, 9(2): 123–150.

[22] Gradshteyn, I. S. and Ryzhik, I. M., Table of Integrals, Series, and Products, 7th ed., Academic Press, San Diego, 2007.

[23] Gray, R. M., Toeplitz and circulant matrices: A review, Foundations and Trends in Communications and Information Theory, 2006, 2(3): 155–239.

[24] Grenander, U. and Szegö, G., Toeplitz Forms and their Applications, 2nd ed., Cambridge University Press, Cambridge, 1958.

[25] Heyde, C. C., A contribution to the theory of large deviations for sums of independent random variables, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, 1967, 7(5): 303–308.

[26] Horn, R. A. and Johnson, C. R., Matrix Analysis, 2nd ed., Cambridge University Press, New York, 2013.

[27] Jensen, J. L., Saddlepoint Approximations, Oxford University Press, New York, 1995.

[28] Karling, M. J., Lopes, A. O. and Lopes, S. R. C., Pentadiagonal matrices and an application to the centered MA(1) stationary Gaussian process, International Journal of Applied Mathematics and Statistics, 2021, 61(1):

1–22.

[29] Lewis, J. T. and Pfister, C.-E., Thermodynamic probability theory: Some aspects of large deviations, Russian Mathematical Surveys, 1995, 50(2): 279–317.

[30] Macci, M. and Trapani, S., Large deviations for posterior distributions on the parameter of a multivariate AR(p) process, Annals of the Institute of Statistical Mathematics, 2013, 65: 703–719.

[31] Mann, H. B. and Wald, A., On the statistical treatment of linear stochastic difference equations, Econometrica, 1943, 11(3): 173–200.

[32] Mas, A. and Menneteau, L., Large and moderate deviations for infinite-dimensional autoregressive processes, Journal of Multivariate Analysis, 2003, 87(2): 241–260.

[33] McLeod, A. I. and Jiménez, C., Nonnegative definiteness of the sample autocovariance function, The American Statistician, 1984, 38(4): 297–298.

[34] Miao, Y., Large deviation principles for moving average processes of real stationary sequences, Acta Applicandae Mathematicae, 2009, 106: 177–184.

[35] Nikolski, N., Toeplitz Matrices and Operators, Cambridge University Press, Cambridge, 2020.

[36] Robertson, S. and Almost, C., Large Deviation Principles, Available at https://www.andrew.cmu.edu/user/calmost/pdfs/21-882-ldp_lec.pdf, 2010.

[37] Rockafellar, R. T., Convex Analysis, Princeton University Press, New Jersey, 2016.

[38] Rozovskii, L. V., Probabilities of large deviations of sums of independent random variables with common distribution function in the domain of attraction of the normal law, Theory of Probability & Its Applications, 1989, 34(4): 625–644.

[39] Rozovskii, L. V., Large deviations of sums of independent random variables from the domain of attraction of a stable law, Journal of Mathematical Sciences, 1999, 93(3): 421–435.

[40] Shumway, R. H. and Stoffer, D. S., Time Series Analysis and its Applications: With R Examples, 4th ed., Springer, New York, 2016.

[41] Tyrtyshnikov, E. E., Influence of matrix operations on the distribution of eigenvalues and singular values of Toeplitz matrices, Linear Algebra and its Applications, 1994, 207: 225–249.

[42] Wu, L., On large deviations for moving average processes, In: Lai, T. L., Yang, H. and Yung, S.-P.(eds.), Probability, Finance and Insurance: Proceedings of a Workshop, the University of Hong Kong, World Scientific, 2004, 15–49.

[43] Zaigraev, A., Multivariate large deviations with stable limit laws, Probability and Mathematical Statistics, 1999, 19(2): 323–335.

[44] Zani, M., Sample path large deviations for squares of stationary Gaussian processes, Theory of Probability & its Applications, 2013, 57(2): 347–357.