stochastic processes and their applications

ELSEVIER Stochastic Processes and their Applications 71 (1997) 75–90

Large deviations for quadratic forms of stationary Gaussian processes

B. Bercu[a,*], F. Gamboa[b], A. Rouault[c]

[a] Lab. de Statistiques, Université Paris Sud, F-91405 Orsay, France

[b] Lab. de Statistiques, Université Paris Sud, F-91405 Orsay and Université Paris-Nord, Institut Galilée, F-93430 Villetaneuse, France

[c] Département de Mathématiques, Bâtiment Fermat, Université de Versailles, F-78035 Versailles, France

Received 10 June 1996; received in revised form 9 June 1997

Abstract

A large deviation principle is proved for Toeplitz quadratic forms of centred stationary Gaussian processes. The rate function is obtained by a sharp study of the behaviour of eigenvalues of a product of two Toeplitz matrices. Some statistical applications such as the likelihood ratio test and the estimation of the parameter of an autoregressive Gaussian process are also provided. © 1997 Elsevier Science B.V.

AMS classification: 60F10; 11E25; 60G15; 47B35

Keywords: Large deviations; Quadratic forms; Gaussian processes; Toeplitz matrices

1. Introduction

One of the main tools in the study of a centred stationary real Gaussian process (X_n) is the sequence of the empirical spectral measures $(\mathscr{W}_n)$ acting on bounded measurable real functions f on the torus $\mathbb{T} = [-\pi, \pi[$ by

$$\mathscr{W}_n(f) = \frac{1}{2\pi n} \int_{\mathbb{T}} f(t) \left| \sum_{j=1}^{n} X_j \exp(\mathrm{i}jt) \right|^2 \mathrm{d}t. \tag{1.1}$$

Some of its well-known properties are given in Azencott and Dacunha-Castelle (1986). If the spectral measure μ defined by

$$E(X_j X_k) = \frac{1}{2\pi} \int_{\mathbb{T}} \exp(\mathrm{i}(j-k)t)\,\mathrm{d}\mu(t) \tag{1.2}$$

has a density $g \in L^{\infty}(\mathbb{T})$, then we have (Avram, 1988; Azencott and Dacunha-Castelle, 1986)

$$\mathscr{W}_n \Rightarrow \mu \tag{1.3}$$

* Corresponding author. E-mail: bernard.bercu@math.u-psud.fr.

PII S0304-4149(97)00071-9

in probability where $\mu(f)=\frac{1}{2\pi}\int_{\mathbb{T}} f(x)g(x)\,\mathrm{d}x$ and $\Rightarrow$ denotes the weak convergence. We investigate the large deviations properties of the process $(\mathscr{W}_n)$. In a first stage, we are interested in the large deviations properties for its marginals. We now recall the standard large deviations definition.

Definition. We say that a sequence of probability measures (P_n) on a regular Hausdorff space $(\mathscr{X},\mathscr{B}(\mathscr{X}))$ satisfies a Large Deviation Principle (LDP) with rate function I, if I is a lower semicontinuous function $I:\mathscr{X}\to[0,+\infty]$ such that

(a) *Upper bound*: For any closed set F of $\mathscr{X}$

$$\limsup_{n\to\infty}\frac{1}{n}\log P_n(F)\leqslant-\inf_{x\in F} I(x),$$

(b) *Lower bound*: For any open set G of $\mathscr{X}$

$$-\inf_{x\in G} I(x)\leqslant\liminf_{n\to\infty}\frac{1}{n}\log P_n(G).$$

I is a good rate function if its level sets are compact subsets of $\mathscr{X}$. For any real function $f\in L^\infty(\mathbb{T})$, let $T_n(f)$ be the Toeplitz matrix associated with f

$$T_n(f)=\left(\frac{1}{2\pi}\int_{\mathbb{T}}\exp(\mathrm{i}(j-k)t)f(t)\,\mathrm{d}t\right)_{1\leqslant j,k\leqslant n}. \tag{1.4}$$

One can rewrite (1.1) as

$$\mathscr{W}_n(f)=\frac{1}{n}X^{(n)*}T_n(f)X^{(n)}, \tag{1.5}$$

where $X^{(n)*}=(X_1,\ldots,X_n)$. We assume in all the sequel that the spectral density g is not the zero function. By an orthogonal change of basis, it is easy to see that

$$\mathscr{W}_n(f)=\frac{1}{n}\sum_{i=1}^{n}\lambda_i^n Z_i^n, \tag{1.6}$$

where $\lambda_1^n,\ldots,\lambda_n^n$ are the eigenvalues of $T_n(g)^{1/2}T_n(f)T_n(g)^{1/2}$ (which are also the eigenvalues of $T_n(f)T_n(g)$) and $Z_1^n,\ldots,Z_n^n$ are i.i.d. with $\chi^2(1)$ distribution. An application of the Gärtner–Ellis theorem (see e.g. Dembo and Zeitouni, 1993, Theorem 2.3.6) needs the convergence of the normalized cumulant generating function

$$L_n(\lambda)=\frac{1}{n}\log E(\mathrm{e}^{n\lambda\mathscr{W}_n(f)}), \tag{1.7}$$

$$L_n(\lambda)=-\frac{1}{2n}\log\det(I_n-2\lambda T_n(g)^{1/2}T_n(f)T_n(g)^{1/2})=-\frac{1}{2n}\sum_{i=1}^{n}\log(1-2\lambda\lambda_i^n), \tag{1.8}$$

where for convenience and in all the sequel $\log z=-\infty$ if $z\leqslant 0$. Let us consider two simple examples, the white noise case corresponding to (X_n) i.i.d. with g constant, say $g=1$, and the sum of squares case where we take $f=1$ so that $\mathscr{W}_n(1)=\frac{1}{n}\sum X_i^2$.

In those cases, $T_n(g)^{1/2}T_n(f)T_n(g)^{1/2}$ is a Toeplitz matrix $T_n(h)$ with $h=f$ or g. The asymptotic distribution of the eigenvalues of $T_n(h)$ is known (Grenander and Szegö, 1958; Avram, 1988). If we set $m_h = \operatorname{essinf} h$ and $M_h = \operatorname{esssup} h$, we have

$$m_h \leqslant \min_i \lambda_i^n \leqslant \max_i \lambda_i^n \leqslant M_h, \tag{1.9}$$

$$\lim_{n\to\infty} \min_i \lambda_i^n = m_h, \quad \lim_{n\to\infty} \max_i \lambda_i^n = M_h, \tag{1.10}$$

$$\frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i^n} \Rightarrow P_h, \tag{1.11}$$

where P_h is the image probability of the normalized Lebesgue measure on the torus $\mathbb{T}$ by the mapping h. In those cases, as $h=f$ or g, the limit of the sequence $(L_n(\lambda))$ is given by

$$L(h,\lambda) = -\frac{1}{4\pi}\int_{\mathbb{T}} \log(1-2\lambda h(t))\,\mathrm{d}t \tag{1.12}$$

for $\lambda \in \mathscr{D}_h$ defined by

$$\mathscr{D}_h = \begin{cases} \left]\dfrac{1}{2m_h}, \dfrac{1}{2M_h}\right[& \text{if } m_h < 0 \text{ and } M_h > 0, \\ \left]-\infty, \dfrac{1}{2M_h}\right[& \text{if } m_h > 0, \\ \left]\dfrac{1}{2m_h}, +\infty\right[& \text{if } M_h < 0 \end{cases} \tag{1.13}$$

and $L(h,\lambda) = +\infty$ outside $\overline{\mathscr{D}_h}$. As noticed by Bryc and Dembo (1997), a direct application of the Gärtner–Ellis theorem is not possible in general. The problem arises at the boundary of $\mathscr{D}_h$. Steepness is not guaranteed. Indeed, it depends on the regularity of h in the neighbourhood of the set $\{h=m_h\}\cup\{h=M_h\}$. The asymptotic cumulant generating function does not contain apparently the whole information on the large deviation property of the process. There is a loss of information passing to the limit. Bryc and Dembo (1997) go around this key point via a time-varying change of probability. In the present paper, we develop a simple modification of the Gärtner–Ellis theorem which extensively uses relation (1.6) together with the independence structure of $Z_1^n, \ldots, Z_n^n$. We prove a LDP whose good rate function K_h is the Fenchel–Legendre dual of $L(h,\cdot)$.

$$K_h(x) = \sup_{y\in\mathbb{R}} \left(xy + \frac{1}{4\pi}\int_{\mathbb{T}} \log(1-2yh(t))\,\mathrm{d}t \right) \quad (x \in \mathbb{R}). \tag{1.14}$$

In the general case, $T_n(g)^{1/2}T_n(f)T_n(g)^{1/2}$ or $T_n(f)T_n(g)$ is not a Toeplitz matrix. In view of the Toeplitz asymptotic homomorphism, we prove the analogous of (1.11) with $h=fg$. The limits of the minimum and the maximum eigenvalues of $T_n(f)T_n(g)$ are not in general m_{fg} and M_{fg}. If $I_{fg} = [m_{fg}, M_{fg}]$, the eigenvalues $\lambda_1^n, \ldots, \lambda_n^n$ may not all lie in I_{fg} but are bounded by $\|f\|_\infty \|g\|_\infty$. We have to take into account the asymptotic behaviour of the *bad* eigenvalues i.e. the $\lambda_i^n \notin I_{fg}$. To our knowledge, this has been omitted in the literature leading to wrong large deviations functional (e.g. Avram, 1988; Bucklew, 1990). Finally, we prove a LDP for subsequences of $(\mathscr{W}_n(f))$. The rate functions are infimal convolutions (Rockafeller, 1970, p. 34) of K_{fg} with linear functions. Their slopes are connected with limit points of *bad* extremal eigenvalues. These results are extended to locally stationary Gaussian processes in Zani (1997).

The paper is organized as follows. In Section 2, we establish a LDP for subsequences of $(\mathscr{W}_n(f))$. This result was announced in Bercu et al. (1996) for continuous functions. Then, we extend these results to Toeplitz-like matrices. Section 3 is devoted to the proofs of the results of Section 2. Statistical applications are developed in Section 4. First, we prove a LDP for the likelihood ratio statistic of $g=g_0$ against $g=g_1$ improving the partial results of Dacunha-Castelle (1979), Bouaziz (1993) and Barone et al. (1995). Next, we also obtain LDP for the least squares and Yule–Walker estimators of the parameter of the stationary autoregressive Gaussian process exhibiting a nonconvex rate function. Proofs of statistical results are collected in Section 5.

2. Main results

Fixing $f,g\in L^\infty(\mathbb{T})$, let $\underline{a}_n(f,g)$ and $\overline{a}_n(f,g)$ be the minimum and the maximum eigenvalues of $T_n(f)T_n(g)$, respectively. The sequences $(\underline{a}_n(f,g))$ and $(\overline{a}_n(f,g))$ are bounded sequences. Therefore, one can define the set $\mathscr{A}(f,g)\subset\mathbb{R}^2$ of all limit points of sequences $(\underline{a}_n(f,g),\overline{a}_n(f,g))$. For $(\underline{a},\overline{a})\in\mathscr{A}(f,g)$, we clearly have $\max(|\underline{a}|,|\overline{a}|)\leqslant \|f\|_\infty\|g\|_\infty$. Denote by $\mathscr{N}(\underline{a},\overline{a})$ the set of all integers increasing subsequences (n_j) such that $(\underline{a}_{n_j}(f,g),\overline{a}_{n_j}(f,g))$ converges to $(\underline{a},\overline{a})\in\mathscr{A}(f,g)$ as j goes to infinity. We show that some linear parts can appear in the large deviation rate function, depending on the location of $\underline{a}$ and $\overline{a}$. Set

$$J(x)=\begin{cases} K_{fg}(x) & \text{if } x\in\,]x_1,x_2[,\\[2mm] K_{fg}(x_1)+\dfrac{1}{2\underline{a}}(x-x_1) & \text{if } x\in\,]-\infty,x_1],\\[2mm] K_{fg}(x_2)+\dfrac{1}{2\overline{a}}(x-x_2) & \text{if } x\in[x_2,+\infty[,\end{cases}\tag{2.1}$$

where x_1 and x_2 are given by

$$x_1=\begin{cases} L'\left(fg,\dfrac{1}{2\underline{a}}\right) & \text{if } \underline{a}<0 \text{ and } \underline{a}<m_{fg},\\[2mm] -\infty & \text{otherwise.}\end{cases}\tag{2.2}$$

$$x_2=\begin{cases} L'\left(fg,\dfrac{1}{2\overline{a}}\right) & \text{if } \overline{a}>0 \text{ and } \overline{a}>M_{fg},\\[2mm] +\infty & \text{otherwise.}\end{cases}\tag{2.3}$$

Theorem 1. *For $(n_j)\in\mathscr{N}(\underline{a},\overline{a})$ with $(\underline{a},\overline{a})\in\mathscr{A}(f,g)$ the subsequence $(\mathscr{W}_{n_j}(f))$ satisfies a LDP with good rate function J.*

Corollary 2. (a) *If* $fg \geqslant 0$ *a.e.* (*resp.* $fg \leqslant 0$ *a.e.*) *then the sequence* $(\mathscr{W}_n(f))$ *satisfies a LDP with good rate function* J *if and only if the sequence* $(\bar{a}_n(f,g))$ (*resp.* $(\underline{a}_n(f,g))$), *has only one limit point.*

(b) *Otherwise, the sequence* $(\mathscr{W}_n(f))$ *satisfies a LDP with good rate function* J *if and only if* $\mathscr{A}(f,g)$ *is a singleton.*

Remark. In the particular case where f is positive and $\|fg\|_\infty = \|f\|_\infty \|g\|_\infty$ then $(\mathscr{W}_n(f))$ satisfies a LDP with good rate function K_{fg}. For example, $\mathscr{W}_n(1) = \frac{1}{n}\sum_{i=1}^n X_i^2$ satisfies a LDP with good rate function K_g. This last result was recently proved by Bryc and Dembo (1997). See also Bryc and Smolenski (1993) for the autoregressive process, Bucklew (1990, p. 103) for a heuristic approach.

Hermitian forms. Theorem 1 and Corollary 2 may be generalized by the following proposition. For $n \geqslant 1$, let M_n be an order n Hermitian matrix. Denote by $\lambda_1^n, \ldots, \lambda_n^n$ the eigenvalues of $T_n(g)^{1/2} M_n T_n(g)^{1/2}$ and let $\underline{a}_n$ and $\bar{a}_n$ be the minimum and the maximum ones.

Proposition 3. *Assume that* $(\underline{a}_n)$ *and* $(\bar{a}_n)$ *are bounded sequences and that, as* $n \to \infty$

$$\sigma_n = \frac{1}{n}\sum_{i=1}^n \delta_{\lambda_i^n} \Rightarrow P_{fg} \tag{2.4}$$

for some measurable function f *on* $\mathbb{T}$ *such that* $fg \in L^\infty(\mathbb{T})$. *Then, Theorem* 1 *and Corollary* 2 *hold replacing* $\mathscr{W}_n(f)$ *by*

$$Z_n = \frac{1}{n} X^{(n)*} M_n X^{(n)}. \tag{2.5}$$

Proof. We prove Proposition 3 exactly as Theorem 1 replacing $\mathscr{W}_n(f)$ by Z_n. □

Analytic extension. Theorem 1 and Corollary 2 may be extended to more general quadratic forms, namely built on additive analytic function of Toeplitz matrices. For any $d \geqslant 1$, let $\varphi = (\varphi_1, \ldots, \varphi_d) \in L^\infty(\mathbb{R}^d; \mathbb{T})$. For $j = 1, \ldots, d$, let $\Omega_j = [\operatorname{essinf} \varphi_j, \operatorname{esssup} \varphi_j]$ and denote by F_j an analytic function defined on an open set containing Ω_j. For $x = (x_1, \ldots, x_d) \in \Omega = \prod_{j=1}^d \Omega_j$, we set

$$F(x) = \sum_{j=1}^d F_j(x_j).$$

Since for every j and every n the spectrum of $T_n(\varphi_j)$ is included in Ω_j, it is possible to define the Hermitian matrix $F_j(T_n(\varphi_j))$. From now, we use the notation of Proposition 3. Set

$$M_n = F(T_n(\varphi_1), \ldots, T_n(\varphi_d)), \tag{2.6}$$

$$Z_n = \frac{1}{n} X^{(n)*} M_n X^{(n)}. \tag{2.7}$$

Corollary 4. *Theorem* 1 *and Corollary* 2 *hold with* $f = F(\varphi)$.

3. Proofs

3.1. Proof of Theorem 1

We now give two lemmas which are useful to prove Theorem 1. The first give the behaviour of good and bad eigenvalues.

Lemma 5. *Set*

$$\sigma_n = \frac{1}{n}\sum_{i=1}^{n}\delta_{\lambda_i^n}. \tag{3.1}$$

Then $\sigma_n \Rightarrow P_{fg}$ and there exists a nonincreasing sequence of positive real numbers (ε_n) such that

(a) $\lim_{n\to\infty}\varepsilon_n = 0$.
(b) $\widehat{\sigma}_n = \mathbb{1}_{[m_{fg}-\varepsilon_n, M_{fg}+\varepsilon_n]}\sigma_n \Rightarrow P_{fg}$.
(c) $\overline{\sigma}_n = \sigma_n - \widehat{\sigma}_n \Rightarrow 0$.

Proof. We first show that (σ_n) converges weakly to P_{fg}. This result is classical for the distribution of the eigenvalues of a single Toeplitz matrix (Grenander and Szegö, 1958). Actually, we shall use the same method. As we saw at the beginning of Section 2, all the eigenvalues of $T_n(f)T_n(g)$ are uniformly bounded. Hence, it is enough to show the convergence of the moments

$$\forall l \geqslant 1, \quad I_l^n = \int_{\mathbb{R}} x^l \,\mathrm{d}\sigma_n(x) \underset{n\to\infty}{\longrightarrow} \int_{\mathbb{R}} x^l \,\mathrm{d}P_{fg}(x). \tag{3.2}$$

We have

$$I_l^n = \frac{1}{n}\sum_{i=1}^{n}(\lambda_i^n)^l = \frac{1}{n}\mathrm{Tr}([T_n(f)T_n(g)]^l). \tag{3.3}$$

Now, from Theorem 1 of Avram (1988)

$$\lim_{n\to\infty}\frac{1}{n}\mathrm{Tr}([T_n(f)T_n(g)]^l) = \frac{1}{2\pi}\int_{\mathbb{T}}(f(t)g(t))^l\,\mathrm{d}t, \tag{3.4}$$

so that (3.2) holds. As the support of P_{fg} is a subset of $[m_{fg}, M_{fg}]$ and $\sigma_n \Rightarrow P_{fg}$ we may find a nonincreasing sequence of positive real number (ε_n) with $\lim_{n\to\infty}\varepsilon_n = 0$ and $\hat{\sigma}_n = \mathbb{1}_{[m_{fg}-\varepsilon_n, M_{fg}+\varepsilon_n]}\sigma_n \Rightarrow P_{fg}$. □

Lemma 6. *Let $(m(n))$ be a sequence of integers increasing to infinity. For $n \geqslant 1$, assume that the random variables $Z_1^n, \ldots, Z_{m(n)}^n$ are i.i.d. with $\chi^2(1)$ distribution. Consider a triangular array $a_1^n, \ldots, a_{m(n)}^n$ of nonnegative real numbers. Set*

$$U_n = \frac{1}{n}\sum_{i=1}^{m(n)} a_i^n Z_i^n. \tag{3.5}$$

We make the following two assumptions:
(a) $m(n) = o(n)$.
(b) $\max_i a_i^n \to a > 0$.
Define for all $x \in \mathbb{R}$

$$G_a(x) = \begin{cases} \dfrac{x}{2a} & \text{if } x \geqslant 0, \\ +\infty & \text{otherwise.} \end{cases} \tag{3.6}$$

Then, (U_n) *satisfies a LDP with good rate function* G_a.

Proof. For the upper bound, we may assume, without loss of generality, that $a_1^n = \max_i a_i^n$ and that F is a closed set of $\mathbb{R}^+$ with $x = \min F > 0$. From Markov inequality, we get

$$\forall y \in \left]0, \frac{1}{2a_1^n}\right[, \quad \frac{1}{n}\log P(U_n \in F) \leqslant -\frac{1}{2n}\sum_{i=1}^{m(n)} \log(1 - 2a_i^n y) - xy.$$

Consequently,

$$\forall y \in \left]0, \frac{1}{2a}\right[, \quad \limsup_{n\to\infty} \frac{1}{n}\log P(U_n \in F) \leqslant -xy,$$

and the upper bound is proved.

It is enough to prove the lower bound for any open half line. For $x > 0$, we have

$$P(U_n > x) \geqslant P(a_1^n Z_1^n > nx) = 2P\left(Y > \sqrt{\frac{nx}{a_1^n}}\right),$$

where Y has a standard normal distribution. Thus, it implies that

$$\liminf_{n\to\infty} \frac{1}{n}\log P(U_n > x) \geqslant \frac{-x}{2a}. \qquad \square$$

Proof of Theorem 1. Fix $(\underline{a}, \overline{a}) \in \mathscr{A}(f,g)$ and assume for the sake of simplicity that $(\underline{a}_n(f,g), \overline{a}_n(f,g))$ converges to $(\underline{a}, \overline{a})$. We will prove Theorem 1 in the case where $\underline{a} < 0$, $\overline{a} > 0$ and both do not lie in I_{fg}. The other cases can be tackled identically. Let (ε_n) be the sequence defined in Lemma 5. We may rewrite (1.6) as $\mathscr{W}_n(f) = W_n^1 + W_n^2 + W_n^3$ where $W_n^1 = W_n^+ + W_n^-$ and

$$W_n^+ = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{(\lambda_i^n \in [0, M_{fg}+\varepsilon_n])} \lambda_i^n Z_i^n, \qquad W_n^- = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{(\lambda_i^n \in [m_{fg}-\varepsilon_n, 0[)} \lambda_i^n Z_i^n,$$

$$W_n^2 = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{(\lambda_i^n \in]-\infty, m_{fg}-\varepsilon_n[)} \lambda_i^n Z_i^n, \qquad W_n^3 = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{(\lambda_i^n \in]M_{fg}+\varepsilon_n, +\infty[)} \lambda_i^n Z_i^n.$$

We shall prove that each of this four parts satisfies a LDP. As, for any λ, the function $\log(1 - 2\lambda x)$ vanishes at $x = 0$, we may (from Lemma 5(b)) calculate the limit of the normalized cumulant generating function of W_n^+

$$\lim_{n\to\infty} \frac{1}{n} \log E(e^{n\lambda W_n^+}) = -\frac{1}{4\pi} \int_{fg\geqslant 0} \log(1 - 2\lambda fg(t))\,dt = L((fg)^+, \lambda) \quad \text{if } \lambda < \frac{1}{2M_{fg}}$$

$$= +\infty = L((fg)^+, \lambda) \quad \text{if } \lambda > \frac{1}{2M_{fg}}.$$

We do not know the behaviour of the above limit for $\lambda = 1/2M_{fg}$. This particular point will not alterate the LDP we are looking for. We set

$$L((fg)^+, 1/2M_{fg}) = \lim_{\lambda \nearrow (1/2M_{fg})} L((fg)^+, \lambda).$$

With this choice the function $L((fg)^+, \cdot)$ is lower semicontinuous. Now, we have two cases to consider.

Case 1: The function $L((fg)^+, \cdot)$ is steep. Then, we may apply Gärtner–Ellis Theorem (see Dembo and Zeitouni 1993, Theorem 2.3.6). Consequently, (W_n^+) satisfies a LDP with good rate function $K_{(fg)^+}$.

Case 2: The function $L((fg)^+, \cdot)$ is not steep. Let

$$x^* = \lim_{\lambda \nearrow (1/2M_{fg})} L'((fg)^+, \lambda).$$

Then, it is simple to check that for $x \geqslant x^*$,

$$K_{(fg)^+}(x) = K_{(fg)^+}(x^*) + \frac{1}{2M_{fg}}(x - x^*).$$

We shall prove that (W_n^+) still satisfies a LDP with good rate function $K_{(fg)^-}$. The upper bound is given by the Gärtner–Ellis Theorem. Now, it remains to check that the lower bound holds for any half lines. This may be deduced again from the Gärtner–Ellis Theorem if $x < x^*$. Thus, we shall show that, for any $x \geqslant x^*$,

$$\liminf_{n\to\infty} \frac{1}{n} \log P(W_n^+ > x) \geqslant -K_{(fg)^-}(x). \tag{3.7}$$

Let $i_n \in \{1, \ldots, n\}$ with $\lambda_{i_n}^n = \max\{\lambda_i^n; \lambda_i^n \in [0, M_{fg} + \varepsilon_n]\}$. By the independence structure of $Z_1^n, \ldots, Z_n^n$, we have for any $0 < \varepsilon < x^*$

$$P(W_n^+ > x) \geqslant P\left(W_n^+ - \frac{\lambda_{i_n}^n}{n} Z_{i_n}^n > x^* - \varepsilon\right) P\left(\frac{\lambda_{i_n}^n}{n} Z_{i_n}^n > x - x^* + \varepsilon\right). \tag{3.8}$$

On the one hand, it is easy to see that the limit of the normalized cumulant generating function of $W_n^+ - (\lambda_{i_n}^n/n)Z_{i_n}^n$ is the same as the one of W_n^+. This implies, using again the Gärtner–Ellis Theorem, that

$$\liminf_{n\to\infty} \frac{1}{n} \log P\left(W_n^+ - \frac{\lambda_{i_n}^n}{n} Z_{i_n}^n > x^* - \varepsilon\right) \geqslant -K_{(fg)^+}(x^* - \varepsilon). \tag{3.9}$$

On the other hand, we have obviously $\lim_{n\to\infty} \lambda_{i_n}^n = M_{fg}$ so that Lemma 6 gives

$$\liminf_{n\to\infty} \frac{1}{n} \log P\left(\frac{\lambda_{i_n}^n}{n} Z_{i_n}^n > x - x^* + \varepsilon\right) \geqslant -\frac{x - x^* + \varepsilon}{2M_{fg}}. \tag{3.10}$$

Now, as ε is arbitrary and $K_{(fg)^+}$ is continuous, adding (3.9) to (3.10) leads to (3.7). Using the same arguments, we see that (W_n^-) satisfies a LDP with good rate function $K_{(fg)^-}$. Hence, as these two sequences are independent, we obtain that (W_n^1) satisfies a LDP with rate function given by the infimal convolution $K_{(fg)^-}\Box K_{(fg)^+}=K_{fg}$ (see Rockafellar, 1970, p. 34 and Theorem 16.4. p. 145 for the last equality).

Now, taking into account Lemma 5(c), Lemma 6 implies that $(-W_n^2)$ and (W_n^3) also satisfy a LDP with good rate function $G_{-\underline{a}}$ and $G_{\bar{a}}$ respectively. Hence, as W_n^1, W_n^2, W_n^3 are independent, we obtain via the contraction principle that $(\mathcal{W}_n(f))$ satisfies a LDP with rate function given by $J=K_{fg}\Box G_{-\underline{a}}\Box G_{\bar{a}}$

$$J(x)=\inf_{y_1+y_2+y_3=x}(K_{fg}(y_1)+G_{-\underline{a}}(-y_2)+G_{\bar{a}}(y_3)) \quad (x\in\mathbb{R}). \tag{3.11}$$

It can be easily checked that J is given by (2.1)–(2.3). □

3.2. Proof of Corollary 4

We want to apply Proposition 3. First, it is easy to see that $(\underline{a}_n)$ and $(\bar{a}_n)$ are bounded sequences and that $fg\in L^\infty(\mathbb{T})$. We now prove (2.4). For every $l\in\mathbb{N}$, the function Φ_l defined by

$$\Phi_l(x_0,x_1,\dots,x_d)=(x_0F(x_1,\dots,x_d))^l \tag{3.12}$$

is analytic on an open set containing $I_g\times\Omega$. For every $\varepsilon>0$, one can find a polynomial $p(x_0,\dots,x_d)$ such that

$$\sup_{(x_0,\dots,x_d)}|\Phi_l(x_0,x_1,\dots,x_d)-p(x_0,\dots,x_d)|\leqslant\varepsilon. \tag{3.13}$$

We proceed as in the proof of Lemma 5 for p so that

$$\lim_{n\to\infty}\frac{1}{n}\mathrm{Tr}(p(T_n(g),T_n(\varphi_1),\dots,T_n(\varphi_d)))=\frac{1}{2\pi}\int_{\mathbb{T}}p(g(t),\varphi_1(t),\dots,\varphi_d(t))\,\mathrm{d}t \tag{3.14}$$

and we apply the known inequality

$$|\mathrm{Tr}(B)|\leqslant\|B\|_1\leqslant n\|B\|_\infty \tag{3.15}$$

for every square matrix B of order n. Finally, Corollary 4 immediately follows from Proposition 3. □

4. Statistical applications

We now provide two statistical applications. The first deals with the likelihood ratio test on spectral densities and the second is about the least squares and the Yule–Walker estimators of the parameter of the autoregressive stationary Gaussian process.

4.1. Likelihood ratio test

Let g_0 and g_1 be two spectral densities. In this subsection, we study the asymptotic level and power for the Neyman–Pearson test of g_0 against g_1. To be more precise, observing the Gaussian random variables $X_1, \dots, X_n$, we wish to test

$$\mathrm{H}_0 : g = g_0 \quad \text{against} \quad H_1 : g = g_1.$$

For this simple hypothesis, the most powerful test is based on the likelihood ratio statistic (see e.g. Lehman, 1959, Theorem 1, p. 65)

$$\mathscr{L}_n = \frac{1}{2n}\left(\log\frac{\det T_n(g_0)}{\det T_n(g_1)} + X^{(n)*}[T_n(g_0)^{-1} - T_n(g_1)^{-1}]X^{(n)}\right). \tag{4.1}$$

The study of the large deviation properties of $(\mathscr{L}_n)$ under hypothesis H_0 or H_1 is useful to fit asymptotically the threshold or the power of the test, respectively. To our knowledge, these properties have been first investigated by Dacunha–Castelle (1979) (see also Coursol et al., 1979; Bouaziz, 1993 for the same kind of results under weaker assumptions). Nevertheless, in these papers, the large deviations are only given in a neighbourhood of the limit of $(\mathscr{L}_n)$ and no complete LDP is provided. More recently, in the special case of an ARMA process where g_0 and g_1 are rational functions, a LDP is established in Barone et al. (1995). Using our methods, we prove in the general case that $(\mathscr{L}_n)$ satisfies a LDP under very weak assumptions. Some applications of this result are given in Bercu et al. (1997). The keystone is that there is no *bad* eigenvalue for the Toeplitz-like matrix. We now make use of the two following assumptions:
(A_1) The spectral density g_0 is in the Szegö class, i.e. $\log g_0 \in L^1(\mathbb{T})$.
(A_2) The ratio $g_0/g_1 \in L^\infty(\mathbb{T})$.

Proposition 7. *Assume that* (A_1) *and* (A_2) *are satisfied. Then, under the null hypothesis* H_0, *the sequence* $(\mathscr{L}_n)$ *satisfies a LDP with good rate function* $K_h(\cdot - b)$ *where*

$$h = \frac{1}{2}\left(1 - \frac{g_0}{g_1}\right) \quad \text{and} \quad b = \frac{1}{4\pi}\int_{\mathbb{T}} \log\left(\frac{g_0}{g_1}(t)\right)\,\mathrm{d}t.$$

The proof of Proposition 7 is given in Section 5.

4.2. Autoregressive Gaussian process

Consider the autoregressive process

$$X_{n+1} = \theta X_n + \varepsilon_{n+1}, \quad |\theta| < 1, \tag{4.2}$$

where (ε_n) is i.i.d. with $\mathcal{N}(0,\sigma^2)$ distribution. Assume that X_0 is independent of (ε_n) with $\mathcal{N}(0, \sigma^2/(1-\theta^2))$ distribution. (X_n) is a centred stationary Gaussian process with spectral density defined for all $x \in \mathbb{T}$ by $g(x) = \sigma^2(1 + \theta^2 - 2\theta\cos x)^{-1}$. Let $\hat{\theta}_n$ be the least-squares estimator of the parameter θ

$$\hat{\theta}_n = \frac{\sum_{i=1}^n X_i X_{i-1}}{\sum_{i=1}^n X_{i-1}^2}. \tag{4.3}$$

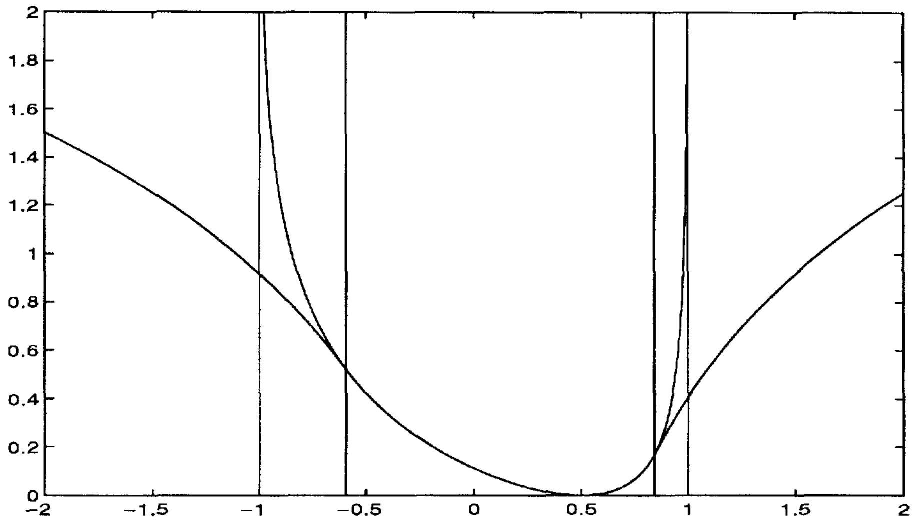

Fig. 1. The rate functions R and S with $\theta = \frac{1}{2}$.

It is well-known that $\hat{\theta}_n \to \theta$ a.s. and $\sqrt{n}(\hat{\theta}_n - \theta) \Rightarrow \mathcal{N}(0, 1-\theta^2)$. One can also estimate θ by the Yule–Walker estimator

$$\tilde{\theta}_n = \frac{\sum_{i=1}^n X_i X_{i-1}}{\sum_{i=0}^n X_i^2}. \tag{4.4}$$

The least squares and the Yule–Walker estimators share the same almost sure property and the same central limit theorem. As far as the authors know, the large deviation properties of $\hat{\theta}_n$ and $\tilde{\theta}_n$ are not available. We show in the next proposition that the large deviation behaviour of the Yule–Walker estimator is better than the one of the least squares for the estimation of the parameter θ. Set

$$a = \frac{\theta - \sqrt{\theta^2 + 8}}{4} \quad \text{and} \quad b = \frac{\theta + \sqrt{\theta^2 + 8}}{4}. \tag{4.5}$$

Proposition 8. *$(\hat{\theta}_n)$ satisfies a LDP with good rate function*

$$R(x) = \begin{cases} \dfrac{1}{2}\log\left(\dfrac{1+\theta^2-2\theta x}{1-x^2}\right) & \textit{if } x \in [a,b], \\ \log|\theta - 2x| & \textit{otherwise.} \end{cases} \tag{4.6}$$

In addition, $(\tilde{\theta}_n)$ satisfies a LDP with good rate function (Fig. 1)

$$S(x) = \begin{cases} \dfrac{1}{2}\log\left(\dfrac{1+\theta^2-2\theta x}{1-x^2}\right) & \textit{if } x \in]-1,1[, \\ +\infty & \textit{otherwise.} \end{cases} \tag{4.7}$$

Remark. The proof of Proposition 8 is given in Section 5. On the one hand, we cannot directly obtain a LDP for $(\hat{\theta}_n)$ via one Toeplitz quadratic form as the random variable X_n^2 does not appear in the denominator of (4.3). On the other hand, a LDP at the process level is known (Donsker and Varadhan, 1985). Unfortunately, we cannot directly use the contraction principle by lack of continuity.

5. Proofs of statistical results

Let (M_n) be a sequence of order n Hermitian matrices and $Z_n = (1/n)X^{(n)*}M_nX^{(n)}$. Denote by $\lambda_1^n, \ldots, \lambda_n^n$ the eigenvalues of $T_n(g)^{1/2}M_nT_n(g)^{1/2}$ and set

$$\sigma_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i^n}. \tag{5.1}$$

By a classical Gaussian argument, we have for all $\lambda \in \mathbb{R}$

$$\log E(\exp(n\lambda Z_n)) = -\tfrac{1}{2}\log\det(I_n - 2\lambda M_nT_n(g)) \tag{5.2}$$

if $I_n - 2\lambda M_nT_n(g)$ is positive definite and $+\infty$ otherwise. Since we have

$$L_n(\lambda) = \frac{1}{n}\log E(\exp(n\lambda Z_n)) = -\frac{1}{2}\int_{\mathbb{R}} \log(1-2\lambda x)\,\sigma_n(\mathrm{d}x), \tag{5.3}$$

we prove the statistical results via Proposition 3 where the weak convergence (2.4) is obtained by use of the following lemma.

Lemma 9. *Let (σ_n) be a sequence of probability measures on $\mathbb{R}$ whose supports are included in a fixed compact set. For a probability measure σ with compact support, assume that there exists $r>0$ such that*

$$\lim_{n\to\infty}\int_{\mathbb{R}} \log(1-2\lambda x)\,\sigma_n(\mathrm{d}x) = \int_{\mathbb{R}} \log(1-2\lambda x)\,\sigma(\mathrm{d}x) \tag{5.4}$$

for $|\lambda| \leqslant r$. Then, (σ_n) converges weakly to σ.

Proof. The proof is similar to the one of Grenander-Szegö (1958, p. 63). □

5.1. Proof of Proposition 8

Consider the likelihood ratio statistic

$$\mathscr{L}_n = \frac{1}{2n}\left(\log\frac{\det T_n(g_0)}{\det T_n(g_1)} + X^{(n)*}[T_n(g_0)^{-1} - T_n(g_1)^{-1}]X^{(n)}\right). \tag{5.5}$$

From assumptions (A_1) and (A_2), g_0 and g_1 are in the Szegö class and we have for the deterministic part

$$\lim_{n\to\infty}\frac{1}{2n}\log\frac{\det T_n(g_0)}{\det T_n(g_1)} = \frac{1}{4\pi}\int_{\mathbb{T}}\log g_0(t)\,\mathrm{d}t - \frac{1}{4\pi}\int_{\mathbb{T}}\log g_1(t)\,\mathrm{d}t. \tag{5.6}$$

The stochastic part of $\mathscr{L}_n$ is a quadratic form of the process given by

$$Z_n = \frac{1}{n}X^{(n)*}M_nX^{(n)}, \tag{5.7}$$

where

$$M_n = \tfrac{1}{2}(T_n(g_0)^{-1} - T_n(g_1)^{-1}). \tag{5.8}$$

We want to apply Proposition 3 with $g = g_0$ and $f = g_0^{-1} - g_1^{-1}$. Assumption (A_2) implies $fg \in L^\infty(\mathbb{T})$. In addition, there is no *bad* eigenvalues by the following lemma.

Lemma 10. *Assume that the ratio* $g_0/g_1 \in L^\infty(\mathbb{T})$. *If* $m = \operatorname{essinf}(g_0/g_1)$ *and* $M = \operatorname{esssup}(g_0/g_1)$, *then any eigenvalue of* $T_n(g_1)^{-1/2}T_n(g_0)T_n(g_1)^{-1/2}$ *lies in* $[m, M]$.

Proof. We have a.e.

$$g_0 - mg_1 \geqslant 0 \quad \text{and} \quad Mg_1 - g_0 \geqslant 0 \tag{5.9}$$

so that the corresponding Toeplitz matrices are both nonnegative. This implies that both $T_n(g_1)^{-1/2}T_n(g_0)T_n(g_1)^{-1/2} - mI_n$ and $MI_n - T_n(g_1)^{-1/2}T_n(g_0)T_n(g_1)^{-1/2}$ are also non-negative. The result of Lemma 10 immediately follows. □

Proof of Proposition 7. To check condition (2.4), we apply Lemma 9 with $\sigma = P_{fg}$. Set $r^{-1} = 2(M+1)$. For all $|\lambda| \leqslant r$, we have $(1-\lambda)g_1 + \lambda g_0 \geqslant g_1/2$ so that $(1-\lambda)g_1 + \lambda g_0$ lies in the Szegö class. Therefore, from the equality

$$\int_{\mathbb{R}} \log(1 - 2\lambda x)\,\sigma_n(\mathrm{d}x) = \frac{1}{n}\log\det[T_n((1-\lambda)g_1 + \lambda g_0)] - \frac{1}{n}\log\det[T_n(g_1)] \tag{5.10}$$

we find that

$$\begin{aligned}
&\lim_{n\to\infty} \int_{\mathbb{R}} \log(1 - 2\lambda x)\,\sigma_n(\mathrm{d}x) \\
&= \frac{1}{2\pi}\left(\int_{\mathbb{T}} \log((1-\lambda)g_1 + \lambda g_0)(t)\,\mathrm{d}t - \int_{\mathbb{T}} \log g_1(t)\,\mathrm{d}t\right) \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \log(1 - 2\lambda x)\,\sigma(\mathrm{d}x).
\end{aligned} \tag{5.11}$$

Conditions of Proposition 3 are fulfilled so Corollary 2 gives the LDP for the stochastic part of $\mathscr{L}_n$ with rate function K_{fg}. From (5.5), the LDP for $\mathscr{L}_n$ follows. □

5.2. Proof of Proposition 9

Consider the autoregressive process

$$X_{n+1} = \theta X_n + \varepsilon_{n+1}, \quad |\theta| < 1, \tag{5.12}$$

where we can take, without loss of generality, the noise variance $\sigma^2 = 1$. Its spectral density is $g(x) = (1 + \theta^2 - 2\theta\cos x)^{-1}$ with $x \in \mathbb{T}$. The least-squares estimator of θ is given by

$$\hat{\theta}_n = \frac{\sum_{i=1}^n X_i X_{i-1}}{\sum_{i=1}^n X_{i-1}^2}. \tag{5.13}$$

The large deviation properties of $(\hat{\theta}_n)$ are related with the ones of

$$Z_n^c = \frac{1}{n}\left(\sum_{i=1}^{n} X_i X_{i-1} - c\sum_{i=1}^{n} X_{i-1}^2\right) \tag{5.14}$$

with $c \in \mathbb{R}$, since

$$P(\hat{\theta}_n \geqslant c) = P(Z_n^c \geqslant 0). \tag{5.15}$$

If for all $c \in \mathbb{R}$, the LDP for the sequence (Z_n^c) can be performed, the LDP for $(\hat{\theta}_n)$ will immediately follow. We may use a similar argument for $(\tilde{\theta}_n)$. We only give details for $(\hat{\theta}_n)$. We want to apply Proposition 3 with the tridiagonal matrix of order n

$$M_n = \frac{1}{2}\begin{pmatrix} -2c & 1 & 0 & \cdots \\ 1 & -2c & 1 & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & 1 & -2c & 1 \\ \cdots & 0 & 1 & 0 \end{pmatrix} \tag{5.16}$$

and $f(x) = \cos x - c$ with $x \in \mathbb{T}$. One can remark here that M_n is quite similar to the Toeplitz matrix $T_n(f)$. As we have seen at the beginning of Section 5, $L_n(\lambda)$ is associated with the matrix $I_n - 2\lambda M_n T_n(g)$. Since M_n and $T_n^{-1}(g)$ are both tridiagonal matrices, it is more easy to work with the matrix $D_n = T_n^{-1}(g) - 2\lambda M_n$

$$D_n = \begin{pmatrix} r & q & 0 & \cdots \\ q & p & q & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & q & p & q \\ \cdots & 0 & q & 1 \end{pmatrix} \tag{5.17}$$

with $p = 1 + \theta^2 + 2c\lambda$, $q = -\theta - \lambda$ and $r = p - \theta^2$. We now make precise the domain where D_n and so $I_n - 2\lambda M_n T_n(g)$ are positive definite.

Lemma 11. *For n large enough, the tridiagonal matrix D_n is positive definite only on the domain $\mathscr{D} = \mathscr{D}_1 \cup \mathscr{D}_2 \cup \mathscr{D}_3$ with $\mathscr{D}_1 = \{\theta^2 < p \leqslant 2\theta^2$ and $q^2 \leqslant \theta^2(p - \theta^2)\}$, $\mathscr{D}_2 = \{2\theta^2 < p < 2$ and $p > 2|q|\}$, $\mathscr{D}_3 = \{p \geqslant 2$ and $q^2 \leqslant p - 1\}$.*

By a sharp study of the domain $\mathscr{D}$, we find four boundaries in the variable λ

$$\lambda_1 = -2\theta(1 - \theta c) \quad \text{and} \quad \lambda_2 = 2(c - \theta), \tag{5.18}$$

$$\lambda_3 = -\frac{(1+\theta)^2}{2(1+c)} \quad \text{and} \quad \lambda_4 = \frac{(1-\theta)^2}{2(1-c)}. \tag{5.19}$$

By use of the notations at the beginning of Section 2, we obtain that $\mathscr{A}(f,g)$ is a singleton $(\underline{a}, \overline{a})$, where

$$\underline{a}, \overline{a} \in \left\{\frac{1}{2\lambda_1}, \frac{1}{2\lambda_2}, \frac{1}{2\lambda_3}, \frac{1}{2\lambda_4}\right\}. \tag{5.20}$$

These values are to be compared with

$$m_{fg}, M_{fg} \in \left\{ \frac{1}{2\lambda_3}, \frac{1}{2\lambda_4} \right\}. \tag{5.21}$$

The above classification depends on the location of c. The product fg is constant if and only if $c = c_0$ with $c_0 = (1+\theta^2)/2\theta$. We have for instance

$$m_{fg} = \frac{1}{2\lambda_3} \quad \text{and} \quad M_{fg} = \frac{1}{2\lambda_4} \tag{5.22}$$

if $c > c_0$ and vice versa if $c < c_0$. One can observe here that there are *bad* eigenvalues. In order to apply Lemma 9, it remains to prove that convergence (5.4) holds.

Lemma 12. *For $p > 2|q|$, set*

$$\alpha = \frac{p + \sqrt{p^2 - 4q^2}}{2} \quad \textit{and} \quad \beta = \frac{p - \sqrt{p^2 - 4q^2}}{2}. \tag{5.23}$$

Then, for n large enough, we have

$$L_n(\lambda) = \frac{1}{2n}\log((1-\theta^2)(\alpha-\beta)) - \frac{1}{2n}\log((r-\beta)(1-\beta)\alpha^n - (r-\alpha)(1-\alpha)\beta^n)$$

if $\lambda \in \mathscr{D}$ and $L_n(\lambda) = +\infty$ otherwise.

We immediately obtain from Lemma 12 that the limit of the sequence (L_n) is given by

$$L(c,\lambda) = \begin{cases} -\frac{1}{2}\log\alpha = -\frac{1}{2}\log\left(\frac{p + \sqrt{p^2-4q^2}}{2}\right) & \text{if } \lambda \in \mathscr{D}, \\ +\infty & \text{otherwise.} \end{cases} \tag{5.24}$$

It implies that (5.4) holds and we immediately obtain the LDP for (Z_n^c) by Proposition 3.

Proof of Proposition 8. From (5.15), after some calculations (see (2.1)–(2.3)), we can deduce the LDP for $(\hat{\theta}_n)$. The proof of the LDP for $(\tilde{\theta}_n)$ follows the same lines. □

Proofs of Lemmas 11 and 12. One can rewrite (5.17) as

$$D_n = \begin{pmatrix} r & u^* & 0 \\ u & T_n & v \\ 0 & v^* & 1 \end{pmatrix}, \tag{5.25}$$

where $u^* = (q\ 0 \cdots 0)$, $v^* = (0 \cdots 0\ q)$ and $T_n = T_n(g^{-1} - 2\lambda f)$. Hence, we obtain by a block matrix product that

$$\det(D_n) = (\alpha-\beta)^{-1}((r-\beta)(1-\beta)\alpha^{n-1} - (r-\alpha)(1-\alpha)\beta^{n-1}) \tag{5.26}$$

with α and β given by (5.23). Similar result can be found in Jensen (1995, p. 274). The proofs of Lemmas 11 and 12 follow. □

Acknowledgements

The authors thank the referee for comments and well-pointed remarks which helped to improve the manuscript substantially. Many thanks are also due to M. Bouaziz, J. Coursol, D. Dacunha-Castelle and A. Touati for helpful discussions.

References

Avram, F., 1988. On bilinear forms in Gaussian random variables and Toeplitz matrices. Probab. Theory Related Fields 79, 37–45.

Azencott, R., Dacunha-Castelle, D., 1986. Series of Irregular Observations Forecasting and Model Building. Springer, Berlin.

Barone, P., Gigli, A., Piccioni, M., 1995. Optimal importance sampling for some quadratic forms of A.R.M.A. processes. IEEE Trans. Inform. Theory 41, 1834–1844.

Bercu, B., Gamboa, F., Lavielle, M., 1997. Sharp large deviations for Gaussian quadratic forms with statistical applications. Preprint.

Bercu, B., Gamboa, F., Rouault, A., 1996. Grandes déviations pour des formes quadratiques de processus gaussiens stationnaires. Compte-rendus de l'Académie des Sciences, Série I t. 322, 695–698.

Bouaziz, M., 1993. Testing Gaussian sequences and asymptotic inversion of Toeplitz operators. Probab. Math. Statist. 14, 207–222.

Bryc, W., Dembo, A., 1997. Large deviations for quadratic functionals of gaussian processes. Journal of Theoretical Probability 10, 307–332.

Bryc, W., Smolenski, W., 1993. On large deviation principle for a quadratic functional of the autoregressive process. Statist. Probab. Lett. 17, 281–285.

Bucklew, J.A., 1990. Large Deviations Techniques in Decision, Simulation, and Estimation. Wiley, New York.

Coursol, J., Dacunha-Castelle, D., 1979. Sur la formule de Chernoff pour deux processus gaussiens stationnaires. Compte-rendus de l'académie des sciences, Série I t. 288, 769–770.

Dacunha-Castelle, D., 1979. Remarque sur l'étude asymptotique du rapport de vraisem-blance de deux processus gaussiens. Compte-rendus de l'académie des sciences, Série I t. 288, 225–228.

Dembo, A., Zeitouni, O., 1993. Large Deviations Techniques and Applications. Jones and Barblett Pub. Boston.

Donsker, M.S., Varadhan, S.R.S., 1985. Large deviations for stationary Gaussian processes. Comm. Math. Phys. 97, 187–210.

Grenander, V., Szegö, G., 1958. Toeplitz Forms and Their Applications. University of California Press. Berkeley, CA.

Jensen, J.L., 1995. Saddlepoint Approximations. Oxford Statistical Science Series, 16. Oxford.

Lehmann, E.L., 1959. Testing Statistical Hypotheses. Wiley, New York.

Rockafellar, R.T., 1970. Convex Analysis. Princeton University Press, Princeton, NJ.

Zani, M., 1997. Large deviations for quadratic forms built with locally stationary Gaussian processes. Preprint.