Hindawi Publishing Corporation
Advances in Decision Sciences
Volume 2012, Article ID 572919, 15 pages
doi:10.1155/2012/572919

Research Article

Large-Deviation Results for Discriminant Statistics of Gaussian Locally Stationary Processes

Junichi Hirukawa

Faculty of scince, Niigata University, 8050 Ikarashi 2-no-cho, Nishi-ku, Niigata 950-2181, Japan

Correspondence should be addressed to Junichi Hirukawa, hirukawa@math.sc.niigata-u.ac.jp

Received 16 February 2012; Accepted 9 April 2012

Academic Editor: Kenichiro Tamaki

Copyright © 2012 Junichi Hirukawa. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper discusses the large-deviation principle of discriminant statistics for Gaussian locally stationary processes. First, large-deviation theorems for quadratic forms and the log-likelihood ratio for a Gaussian locally stationary process with a mean function are proved. Their asymptotics are described by the large deviation rate functions. Second, we consider the situations where processes are misspecified to be stationary. In these misspecified cases, we formally make the log-likelihood ratio discriminant statistics and derive the large deviation theorems of them. Since they are complicated, they are evaluated and illustrated by numerical examples. We realize the misspecification of the process to be stationary seriously affecting our discrimination.

1. Introduction

Consider a sequence of random variables $S_1, S_2, \ldots$ converging (in probability) to a real constant c. By this we mean that $\Pr\{|S_T - c| > \varepsilon\} \to 0$ as $T \to \infty$ for all $\varepsilon > 0$. The simplest setting in which to obtain large-deviation results is that considering sums of independent identically distributed (iid) random variables on the real line. For example, we would like to consider the large excursion probabilities of sums as the sample average:

$$S_T = T^{-1} \sum_{i=1}^{T} X_i, \tag{1.1}$$

where the X_i, $i = 1, 2, \ldots$, are i.i.d., and T approaches infinity. Suppose that $E(X_i) = m$ exists and is finite. By the law of large numbers, we know that S_T should be converging to m. Hence,

c is merely the expected value of the random process. It is often the case that not only does $\Pr\{|S_T - c| > \varepsilon\}$ go to zero, but it does so exponentially fast. That is,

$$\Pr\{|S_T - c| > \varepsilon\} \approx K(\varepsilon, c, T) \exp\{-TI(\varepsilon, c)\}, \tag{1.2}$$

where $K(\varepsilon, c, T)$ is a slowly varying function of T (relative to the exponential), and $I(\varepsilon, c)$ is a positive quantity. Loosely, if such a relationship is satisfied, we will say that the sequence $\{S_n\}$ satisfies a large-deviation principle. Large-deviation theory is concerned primarily with determining the quantities $I(\varepsilon, c)$ and (to a lesser extent) $K(\varepsilon, c, T)$. The reason for the nomenclature is that for a fixed $\varepsilon > 0$ and a large index T, a large-deviation from the nominal value occurs if $|S_T - c| > \varepsilon$. Large-deviation theory can rightly be considered as a generalization or extension of the law of large numbers. The law of large numbers says that certain probabilities converge to zero. Large-deviation theory is concerned with the rate of convergence. Bucklew [1] describes the historical statements of large-deviation in detail.

There have been a few works on the large-deviation theory for time series data. Sato et al. [2] discussed the large-deviation theory of several statistics for short- and long-memory stationary processes. However, it is still hard to find the large-deviation results for nonstationary processes. Recently, Dahlhaus [3, 4] has formulated an important class of nonstationary processes with a rigorous asymptotic theory, which he calls locally stationary. A locally stationary process has a time-varying spectral density whose spectral structure changes smoothly with time. There are several papers which discuss discriminant analysis for locally stationary processes (e.g., Chandler and Polonik [5], Sakiyama and Taniguchi [6], and Hirukawa [7]). In this paper, we discuss the large-deviation theory of discriminant statistics of Gaussian locally stationary processes. In Section 2 we present the Gärtner-Ellis theorem which establishes a large-deviation principle of random variables based only upon convergence properties of the associated sequence of cumulant generating functions. Since no assumptions are made about the dependency structure of random variables, we can apply this theorem to non-stationary time series data. In Section 3, we deal with a Gaussian locally stationary process with a mean function. First, we prove the large-deviation principle for a general quadratic form of the observed stretch. We also give the large-deviation principle for the log-likelihood ratio and the misspecified log-likelihood ratio between two hypotheses. These fundamental statistics are important not only in statistical estimation and testing theory but in discriminant problems. The above asymptotics are described by the large-deviation rate functions. In our stochastic models, the rate functions are very complicated. Thus, in Section 4, we evaluate them numerically. They demonstrate that the misspecifications of non-stationary has serious effects. All the proofs of the theorems presented in Section 3 are given in the Appendix.

2. Gärtner-Ellis Theorem

Cramér's theorem (e.g., Bucklew [1]) is usually credited with being the first large-deviation result. It gives the large-deviation principle for sums of independent identically distributed random variables. One of the most useful and surprising generalizations of this theorem is the one due to Gärtner [8] and, more recently, Ellis [9]. These authors established a large-deviation principle of random variables based only upon convergence properties of the associated sequence of moment generating functions $\Phi(\omega)$. Their methods thus allow large-deviation results to be derived for dependent random processes such as Markov chains and

functionals of Gaussian random processes. Gärtner [8] assumed throughout that $\Phi(\omega) < \infty$ for all ω. By extensive use of convexity theory, Ellis [9] relaxed this fairly stringent condition.

Suppose that we are given an infinite sequence of random variables $\{Y_T, \ T \in \mathbb{N}\}$. No assumptions are made about the dependency structure of this sequence. Define

$$\psi_T(\omega) \equiv T^{-1} \log E\{\exp(\omega Y_T)\}. \tag{2.1}$$

Now let us list two assumptions.

Assumption 2.1. $\psi(\omega) \equiv \lim_{T \to \infty} \psi_T(\omega)$ exists for all $\omega \in \mathbb{R}$, where we allow ∞ both as a limit value and as an element of the sequence $\{\psi_T(\omega)\}$.

Assumption 2.2. $\psi(\omega)$ is differentiable on $D_\psi \equiv \{\omega : \psi(\omega) < \infty\}$.

Define the large-deviation rate function by

$$I(x) \equiv \sup_\omega \{\omega x - \psi(\omega)\}; \tag{2.2}$$

this function plays a crucial role in the development of the theory. Furthermore, define

$$\psi'(D_\psi) \equiv \{\psi'(\omega) : \omega \in D_\psi\}, \tag{2.3}$$

where ψ' indicates the derivative of ψ. Before proceeding to the main theorem, we first state some properties of this rate function.

Property 1. $I(x)$ is convex.

We remark that a convex function $I(\cdot)$ on the real line is continuous everywhere on $D_I \equiv \{x : I(x) < \infty\}$, the domain of $I(\cdot)$.

Property 2. $I(x)$ has its minimum value at $m = \lim_{T \to \infty} T^{-1} E(Y_T)$, and $I(m) = 0$.

We now state a simple form of a general large-deviation theorem which is known as the Gärtner and Ellis theorem (e.g., Bucklew [1]).

Lemma 2.3 (Gärtner-Ellis). *Let (a, b) be an interval with $[a, b] \cap D_I \neq \emptyset$. If Assumption 2.1 holds and $a < b$, then*

$$\limsup_{T \to \infty} T^{-1} \log \Pr\left\{T^{-1} Y_T \in [a, b]\right\} \le - \inf_{x \in [a,b]} I(x). \tag{2.4}$$

If Assumptions 2.1 and 2.2 hold and $(a, b) \subset \psi'(D_\psi)$, then

$$\liminf_{T \to \infty} T^{-1} \log \Pr\left\{T^{-1} Y_T \in (a, b)\right\} \ge - \inf_{x \in (a,b)} I(x). \tag{2.5}$$

Large-deviation theorems are usually expressed as two separate limit theorem: an upper bound for closed sets and a lower bound for open sets. In the case of interval subsets

of $\mathbb{R}$, it can be guaranteed that the upper bound equals the lower bound by the continuity of $I(\cdot)$. For the applications that we have in mind, the interval subsets will be sufficient.

3. Large-Deviation Results for Locally Stationary Processes

In this section, using the Gärtner-Ellis theorem, we develop the large-deviation principle for some non-stationary time series statistics. When we deal with non-stationary processes, one of the difficult problems to solve is how to set up an adequate asymptotic theory. To overcome this problem, an important class of non-stationary process has been formulated in rigorous asymptotic framework by Dahlhaus [3, 4], called locally stationary processes. Locally stationary processes have time-varying densities, whose spectral structures smoothly change in time. We give the precise definition of locally stationary processes which is due to Dahlhaus [3, 4].

Definition 3.1. A sequence of stochastic processes $X_{t,T}$ $(t = 1, \ldots, T; T \geq 1)$ is called locally stationary with transfer function A° and trend μ if there exists a representation:

$$X_{t,T} = \mu\left(\frac{t}{T}\right) + \int_{-\pi}^{\pi} \exp(i\lambda t) A_{t,T}^\circ(\lambda) d\xi(\lambda), \tag{3.1}$$

where

(i) $\xi(\lambda)$ is a stochastic process on $[-\pi, \pi]$ with $\overline{\xi(\lambda)} = \xi(-\lambda)$ and

$$\mathrm{cum}\{d\xi(\lambda_1), \ldots, d\xi(\lambda_k)\} = \eta\left(\sum_{j=1}^{k} \lambda_j\right) \nu_k(\lambda_1, \ldots, \lambda_{k-1}) d\lambda_1 \cdots d\lambda_k, \tag{3.2}$$

where $\mathrm{cum}\{\ldots\}$ denotes the cumulant of k-th order, $\nu_1 = 0$, $\nu_2(\lambda) = 1$, $|\nu_k(\lambda_1, \ldots, \lambda_{k-1})| \leq \mathrm{const}_k$ for all k and $\eta(\lambda) = \sum_{j=-\infty}^{\infty} \delta(\lambda + 2\pi j)$ is the period 2π extension of the Dirac delta function. To simplify the problem, we assume in this paper that the process $X_{t,T}$ is Gaussian, namely, we assume that $\nu_k(\lambda) = 0$ for all $k \geq 3$;

(ii) there exists constant K and a 2π-periodic function $A : [0,1] \times \mathbb{R} \to \mathbb{C}$ with $A(u, \lambda) = \overline{A(u, -\lambda)}$ and

$$\sup_{t,\lambda} \left| A_{t,T}^\circ(\lambda) - A\left(\frac{t}{T}, \lambda\right) \right| \leq KT^{-1}, \tag{3.3}$$

for all T. $A(u, \lambda)$ and $\mu(u)$ are assumed to be continuous in u.

The function $f(u, \lambda) := |A(u, \lambda)|^2$ is called the time-varying spectral density of the process. In the following, we will always denote by s and t time points in the interval $[1, T]$, while u and v will denote time points in the rescaled interval $[0, 1]$, that is $u = t/T$.

We discuss the asymptotics away from the expectation of some statistics used for the problem of discriminating between two Gaussian locally stationary processes with specified

mean functions. Suppose that $\{X_{t,T}, t = 1, \ldots, T; T \geq 1\}$ is a Gaussian locally stationary process which under the hypothesis Π_j has mean function $\mu^{(j)}(u)$ and time-varying spectral density $f^{(j)}(u, \lambda)$ for $j = 1, 2$. Let $\mathbf{X}_T = (X_{1,T}, \ldots, X_{T,T})'$ be a stretch of the series $\{X_{t,T}\}$, and let $p^{(j)}(\cdot)$ be the probability density function of $\mathbf{X}_T$ under $\Pi_j (j = 1, 2)$. The problem is to classify $\mathbf{X}_T$ into one of two categories Π_1 and Π_2 in the case that we do not have any information on the prior probabilities of Π_1 and Π_2.

Set $\boldsymbol{\mu}_T^{(j)} = \{\mu^{(j)}(1/T), \ldots, \mu^{(j)}(T/T)\}'$ and $\boldsymbol{\Sigma}_T^{(j)} = \boldsymbol{\Sigma}_T(A^{(j)}, A^{(j)})$, where

$$\boldsymbol{\Sigma}_T(A, B) = \left\{ \int_{-\pi}^{\pi} A_{s,T}^{\circ}(\lambda) B_{t,T}^{\circ}(-\lambda) \exp(i\lambda(s-t)) d\lambda \right\}_{s,t=1,\ldots,T}. \tag{3.4}$$

Initially, we make the following assumption.

Assumption 3.2. (i) We observe a realisation $X_{1,T}, \ldots, X_{T,T}$ of a Gaussian locally stationary process with mean function $\mu^{(j)}$ and transfer function $A^{(j)\circ}$, under Π_j, $j = 1, 2$;

(ii) the $A^{(j)}(u, \lambda)$ are uniformly bounded from above and below, and are differentiable in u and λ with uniformly continuous derivatives $(\partial/\partial u)(\partial/\partial \lambda)A^{(j)}$;

(iii) the $\mu^{(j)}(u)$ are differentiable in u with uniformly continuous derivatives.

In time series analysis, the class of statistics which are quadratic forms of $\mathbf{X}_T$ is fundamental and important. This class of statistics includes the first-order terms (in the expansion with respect to T) of quasi-Gaussian maximum likelihood estimator (QMLE), tests and discriminant statistics, and so forth Assume that G° is the transfer function of a locally stationary process, where the corresponding G satisfies Assumption 3.2 (ii) and $g(u)$ is a continuous function of u which satisfies Assumption 3.2 (iii), if we replace $A^{(j)}$ by G and $\mu^{(j)}(u)$ by $g(u)$, respectively. And set $\mathbf{G}_T \equiv \boldsymbol{\Sigma}_T(G, G)$, $f_G(u, \lambda) \equiv |G(u, \lambda)|^2$, $\mathbf{g}_T \equiv \{g(1/T), \ldots, g(T/T)\}'$ and $Q_T \equiv \mathbf{X}_T' \mathbf{G}_T^{-1} \mathbf{X}_T + \mathbf{g}_T' \mathbf{X}_T$. Henceforth, $E^{(j)}(\cdot)$ stands for the expectation with respect to $p^{(j)}(\cdot)$. Set $S_T^{(j)}(Q) \equiv Q_T - E^{(j)}(Q_T)$ for $j = 1, 2$. We first prove the large-deviation theorem for this quadratic form Q_T of $\mathbf{X}_T$. All the proofs of the theorems are in the Appendix.

Theorem 3.3. *Let Assumption 3.2 hold. Then under Π_1,*

$$\lim_{T \to \infty} T^{-1} \log \Pr_1 \left\{ T^{-1} S_T^{(1)}(Q) > x \right\} = \inf_{\omega} \left\{ \psi_Q \left(\omega; f^{(1)} \right) - \omega \max(x, 0) \right\}, \tag{3.5}$$

and under Π_2,

$$\lim_{T \to \infty} T^{-1} \log \Pr_2 \left\{ T^{-1} S_T^{(2)}(Q) < x \right\} = \inf_{\omega} \left\{ \psi_Q \left(\omega; f^{(2)} \right) - \omega \min(x, 0) \right\}, \tag{3.6}$$

where for $j = 1, 2$, $\psi_Q(\omega; f^{(j)})$ equals

$$\frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \left[\log \frac{f_G(u, \lambda)}{f_G(u, \lambda) - 2\omega f^{(j)}(u, \lambda)} - \frac{2\omega f^{(j)}(u, \lambda)}{f_G(u, \lambda)} \right.$$
$$\left. + \frac{\omega^2 \{f_G(u, 0)g(u) + 2\mu^{(j)}(u)\}^2 f^{(j)}(u, 0)}{2\pi \{f_G(u, 0) - 2\omega f^{(j)}(u, 0)\} f_G(u, 0)} \right] d\lambda \, du. \tag{3.7}$$

Next, one considers the log-likelihood ratio statistics. It is well known that the log-likelihood ratio criterion:

$$\Lambda_T \equiv \log \frac{p^{(2)}(\mathbf{X}_T)}{p^{(1)}(\mathbf{X}_T)} \tag{3.8}$$

gives the optimal discrimination rule in the sense that it minimizes the probability of misdiscrimination (Anderson [10]). Set $S_T^{(j)}(\Lambda) \equiv \Lambda_T - E^{(j)}(\Lambda_T)$ for $j = 1, 2$. For discrimination problem one gives the large-deviation principle for Λ_T.

Theorem 3.4. *Let Assumption 3.2 hold. Then under Π_1,*

$$\lim_{T \to \infty} T^{-1} \log \mathrm{Pr}_1\left\{T^{-1} S_T^{(1)}(\Lambda) > x\right\} = \inf_{\omega}\left\{\psi_L\left(\omega; f^{(1)}, f^{(2)}\right) - \omega \max(x, 0)\right\}, \tag{3.9}$$

where $\psi_L(\omega; f^{(1)}, f^{(2)})$ equals

$$\frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \left[\log \frac{f^{(2)}(u, \lambda)}{(1 - \omega) f^{(2)}(u, \lambda) + \omega f^{(1)}(u, \lambda)} + \omega \left\{ \frac{f^{(1)}(u, \lambda)}{f^{(2)}(u, \lambda)} - 1 \right\} \right. $$
$$\left. + \frac{\omega^2 \left\{ \mu^{(1)}(u) - \mu^{(2)}(u) \right\}^2 f^{(1)}(u, 0)}{2\pi \left\{ (1 - \omega) f^{(2)}(u, 0) + \omega f^{(1)}(u, 0) \right\} f^{(2)}(u, 0)} \right] d\lambda du. \tag{3.10}$$

Similarly, under Π_2,

$$\lim_{T \to \infty} T^{-1} \log \mathrm{Pr}_2\left\{T^{-1} S_T^{(2)}(\Lambda) < x\right\} = \inf_{\omega}\left\{\psi_L\left(-\omega; f^{(2)}, f^{(1)}\right) - \omega \min(x, 0)\right\}. \tag{3.11}$$

In practice, misspecification occurs in many statistical problems. We consider the following three situations. Although actually $\{X_{t,T}\}$ has the time-varying mean functions $\mu^{(j)}(u)$ and the time-varying spectral densities $f^{(j)}(u, \lambda)$, under Π_j, $j = 1, 2$, respectively,

    (i) the mean functions are misspecified to $\mu^{(j)}(u) \equiv 0$, $j = 1, 2$;

    (ii) the spectral densities are misspecified to $f^{(j)}(u, \lambda) \equiv f^{(j)}(0, \lambda)$, $j = 1, 2$;

    (iii) the mean functions and the spectral densities are misspecified to $\mu^{(j)}(u) \equiv 0$ and $f^{(j)}(u, \lambda) \equiv f^{(j)}(0, \lambda)$, $j = 1, 2$. Namely, $\mathbf{X}_T$ is misspecified to stationary.

In each misspecified case, one can formally make the log-likelihood ratio in the form:

$$
M_{1,T} = \frac{1}{2}\left[\log\left\{\frac{\left|\mathbf{\Sigma}^{(1)}\right|}{\left|\mathbf{\Sigma}^{(2)}\right|}\right\} + \mathbf{X}_T'\left\{\mathbf{\Sigma}_T^{(1)^{-1}} - \mathbf{\Sigma}_T^{(2)^{-1}}\right\}\mathbf{X}_T\right],
$$

$$
M_{2,T} = \frac{1}{2}\left[\log\left\{\frac{\left|\widetilde{\mathbf{\Sigma}}^{(1)}\right|}{\left|\widetilde{\mathbf{\Sigma}}^{(2)}\right|}\right\} + \left(\mathbf{X}_T - \boldsymbol{\mu}_T^{(1)}\right)'\widetilde{\mathbf{\Sigma}}_T^{(1)-1}\left(\mathbf{X}_T - \boldsymbol{\mu}_T^{(1)}\right)\right.
$$

$$
\left. -\left(\mathbf{X}_T - \boldsymbol{\mu}_T^{(2)}\right)'\widetilde{\mathbf{\Sigma}}_T^{(2)-1}\left(\mathbf{X}_T - \boldsymbol{\mu}_T^{(2)}\right)\right],
$$
(3.12)

$$
M_{3,T} = \frac{1}{2}\left[\log\left\{\frac{\left|\widetilde{\mathbf{\Sigma}}^{(1)}\right|}{\left|\widetilde{\mathbf{\Sigma}}^{(2)}\right|}\right\} + \mathbf{X}_T'\left\{\widetilde{\mathbf{\Sigma}}_T^{(1)-1} - \widetilde{\mathbf{\Sigma}}_T^{(2)-1}\right\}\mathbf{X}_T\right],
$$

where

$$
\widetilde{\mathbf{\Sigma}}_T^{(j)} = \left\{\int_{-\pi}^{\pi}\exp(i\lambda(t-s))f^{(j)}(0,\lambda)d\lambda\right\}_{s,t=1,\dots,T}.
$$
(3.13)

Set $S_T^{(j)}(M_k) \equiv M_{k,T} - E^{(j)}(M_{k,T})$ for $j = 1,2$ and $k = 1,2,3$. The next result is a large-deviation theorem for the misspecified log-likelihood ratios $M_{k,T}$. It is useful in investigating the effect of misspecification.

Theorem 3.5. *Let Assumption 3.2 hold. Then under Π_1,*

$$
\lim_{T\to\infty} T^{-1}\log \mathrm{Pr}_1\left\{T^{-1}S_T^{(1)}(M_k) > x\right\} = \inf_{\omega}\left\{\psi_{M_k}\left(\omega; f^{(1)}, f^{(2)}, \mu^{(1)}, \mu^{(2)}\right) - \omega\max(x,0)\right\},
$$
(3.14)

where $\psi_{M_1}(\omega; f^{(1)}, f^{(2)}, \mu^{(1)}, \mu^{(2)})$ equals

$$
\frac{1}{4\pi}\int_0^1\int_{-\pi}^{\pi}\left[\log\frac{f^{(2)}(u,\lambda)}{(1-\omega)f^{(2)}(u,\lambda) + \omega f^{(1)}(u,\lambda)} + \omega\left\{\frac{f^{(1)}(u,\lambda)}{f^{(2)}(u,\lambda)} - 1\right\}\right.
$$

$$
\left. +\frac{\left[\omega\mu^{(1)}(u)\{f^{(1)}(u,0) - f^{(2)}(u,0)\}\right]^2}{2\pi\{(1-\omega)f^{(2)}(u,0) + \omega f^{(1)}(u,0)\}f^{(1)}(u,0)f^{(2)}(u,0)}\right]d\lambda du,
$$
(3.15)

$\psi_{M_2}(\omega; f^{(1)}, f^{(2)}, \mu^{(1)}, \mu^{(2)})$ *equals*

$$
\begin{aligned}
\frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \Bigg[&\log \frac{f^{(1)}(0,\lambda) f^{(2)}(0,\lambda)}{f^{(1)}(0,\lambda) f^{(2)}(0,\lambda) - \omega f^{(1)}(u,\lambda)\{f^{(2)}(0,\lambda) - f^{(1)}(0,\lambda)\}} \\
&+ \omega \left\{ \frac{f^{(1)}(u,\lambda)}{f^{(2)}(0,\lambda)} - \frac{f^{(1)}(u,\lambda)}{f^{(1)}(0,\lambda)} \right\} \\
&+ \frac{\omega^2 \{\mu^{(2)}(u) - \mu^{(1)}(u)\}^2 f^{(1)}(u,0) f^{(1)}(0,0)/f^{(2)}(0,0)}{2\pi \left[f^{(1)}(0,0) f^{(2)}(0,0) - \omega f^{(1)}(u,0)\{f^{(2)}(0,0) - f^{(1)}(0,0)\} \right]} \Bigg] d\lambda\, du
\end{aligned}
\tag{3.16}
$$

and $\psi_{M_3}(\omega; f^{(1)}, f^{(2)}, \mu^{(1)}, \mu^{(2)})$ *equals*

$$
\begin{aligned}
\frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \Bigg[&\log \frac{f^{(1)}(0,\lambda) f^{(2)}(0,\lambda)}{f^{(1)}(0,\lambda) f^{(2)}(0,\lambda) - \omega f^{(1)}(u,\lambda)\{f^{(2)}(0,\lambda) - f^{(1)}(0,\lambda)\}} \\
&+ \omega \left\{ \frac{f^{(1)}(u,\lambda)}{f^{(2)}(0,\lambda)} - \frac{f^{(1)}(u,\lambda)}{f^{(1)}(0,\lambda)} \right\} \\
&+ \frac{\omega^2 \mu^{(1)}(u)^2 f^{(1)}(u,0)\{f^{(1)}(0,0) - f^{(2)}(0,0)\}^2/\{f^{(1)}(0,0) f^{(2)}(0,0)\}}{2\pi \left[f^{(1)}(0,0) f^{(2)}(0,0) - \omega f^{(1)}(u,0)\{f^{(2)}(0,0) - f^{(1)}(0,0)\} \right]} \Bigg] d\lambda\, du.
\end{aligned}
\tag{3.17}
$$

Similarly, under Π_2,

$$
\lim_{T \to \infty} T^{-1} \log \, \Pr{}_2 \left\{ T^{-1} S_T^{(2)}(M_k) < x \right\} = \inf_{\omega} \left\{ \psi_{M_k}\left(-\omega; f^{(2)}, f^{(1)}, \mu^{(2)}, \mu^{(1)} \right) - \omega \min(x,0) \right\}.
\tag{3.18}
$$

Now, we turn to the discussion of our discriminant problem of classifying $\mathbf{X}_T$ into one of two categories described by two hypotheses a follows:

$$
\Pi_1 : \mu^{(1)}(u), f^{(1)}(u,\lambda), \qquad \Pi_2 : \mu^{(2)}(u), f^{(2)}(u,\lambda).
\tag{3.19}
$$

We use Λ_T as the discriminant statistic for the problem (3.19), namely, if $\Lambda_T > 0$ we assign $\mathbf{X}_T$ into Π_2, and otherwise into Π_1. Taking $x = -\lim_{T\to\infty} T^{-1} E^{(1)}(\Lambda_T)$ in (3.9), we can evaluate the probability of misdiscrimination of $\mathbf{X}_T$ from Π_1 into Π_2 as follows:

$$P(2 \mid 1) \equiv \mathrm{Pr}_1\{\Lambda_T > 0\}$$

$$\approx \exp\left[T \inf_{\omega}\left\{\frac{1}{4\pi}\int_0^1\int_{-\pi}^{\pi}\left[\log\frac{f^{(1)}(u,\lambda)^{\omega}f^{(2)}(u,\lambda)^{1-\omega}}{(1-\omega)f^{(2)}(u,\lambda)+\omega f^{(1)}(u,\lambda)}\right.\right.\right.$$

$$\left.\left.\left.+\frac{\omega(\omega-1)\{\mu^{(1)}(u)-\mu^{(2)}(u)\}^2}{2\pi\{(1-\omega)f^{(2)}(u,0)+\omega f^{(1)}(u,0)\}}\right]d\lambda du\right\}\right]. \tag{3.20}$$

Thus, we see that the rate functions play an important role in the discriminant problem.

4. Numerical Illustration for Nonstationary Processes

We illustrate the implications of Theorems 3.4 and 3.5 by numerically evaluating the large-deviation probabilities of the statistics Λ_T and $M_{k,T}$, $k = 1, 2, 3$ for the following hypotheses:

$$(\text{Stationary white noise})\quad \Pi_1 : \mu^{(1)}(u) \equiv 0,\quad f^{(1)}(u,\lambda) \equiv 1,$$

$$(\text{Time-varying AR(1)})\quad \Pi_2 : \mu^{(2)}(u) = \mu(u),\quad f^{(2)}(u,\lambda) = \frac{\sigma(u)^2}{\left|1-a(u)e^{i\lambda}\right|^2}, \tag{4.1}$$

where $\mu(u) = (1/2)\exp(-u^2)$, $\sigma(u) = (1/2)\exp\{-(u-1)^2\}$ and $a(u) = (1/2)\exp\{-4(u-1/2)^2\}$, $u \in [0,1]$, respectively. Figure 1 plots the mean function $\mu(u)$ (the solid line), the coefficient functions $\sigma(u)$ (the dashed line), and $a(u)$ (the dotted line). The time-varying spectral density $f^{(2)}(u,\lambda)$ is plotted in Figure 2.

From these figures, we see that the magnitude of the mean function is large at u close to 0, while the magnitude of the time-varying spectral density is large at u close to 1.

Specifically, we use the formulae in those theorems concerning Π_2 to evaluate the limits of the large-deviation probabilities:

$$\mathrm{LDP}(\Lambda) = \lim_{T\to\infty} T^{-1}\log \mathrm{Pr}_2\left\{T^{-1}S_T^{(2)}(\Lambda) < x\right\},$$

$$\mathrm{LDP}(M_k) = \lim_{T\to\infty} T^{-1}\log \mathrm{Pr}_2\left\{T^{-1}S_T^{(2)}(M_k) < x\right\},\quad k = 1, 2, 3. \tag{4.2}$$

Though the result is an asymptotic theory, we perform the simulation with a limited sample size. Therefore, we use some levels of x to expect fairness, that is, we take $x = -0.1, -1, -10$. The results are listed in Table 1.

For each value x, the large-deviation rate of Λ_T is the largest and that of $M_{3,T}$ is the smallest. Namely, we see that the correctly specified case is the best, and on the other hand the misspecified to stationary case is the worst. Furthermore, the large-deviation rates $-\mathrm{LDP}(M_2)$
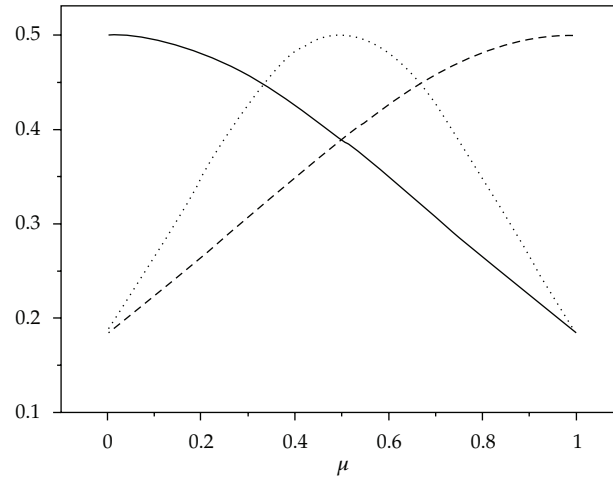
Figure 1: The mean function $\mu(u)$ (the solid line), the coefficient functions $\sigma(u)$ (the dashed line), and $a(u)$ (the dotted line).
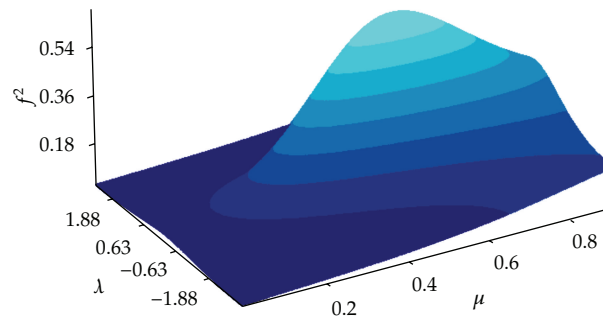

Figure 2: The time-varying spectral density $f^{(2)}(u, \lambda)$.

Table 1: The limits of the large-deviation probabilities of Λ_T and $M_{k,T}$, $k = 1, 2, 3$.

	$x = -0.1$	$x = -1$	$x = -10$
LDP(Λ)	−0.012078	−0.562867	−9.460066
LDP(M_1)	−0.009895	−0.486088	−8.857859
LDP(M_2)	−0.000348	−0.026540	−0.703449
LDP(M_3)	−0.000290	−0.022313	−0.629251

and $-\text{LDP}(M_2)$ are significantly small, comparing with $-\text{LDP}(M_1)$. This fact implies that the misspecification of the spectral density to be constant in the time seriously affects the large-deviation rate.

Figures 3, 4, 5, and 6 show the large-deviation probabilities of Λ_T and $M_{k,T}$, $k = 1, 2, 3$, for $x = -1$, at each time u and frequency λ.

We see that the large-deviation rate of Λ_T keeps the almost constant value at all the time u and frequency λ. On the other hand, that of $M_{1,T}$ is small at u close to 0 and those of $M_{2,T}$ and $M_{3,T}$ are small at u close to 1 and λ close to 0. That is, the large-deviation probability of $M_{1,T}$ is violated by the large magnitude of the mean function, while those of $M_{2,T}$ and
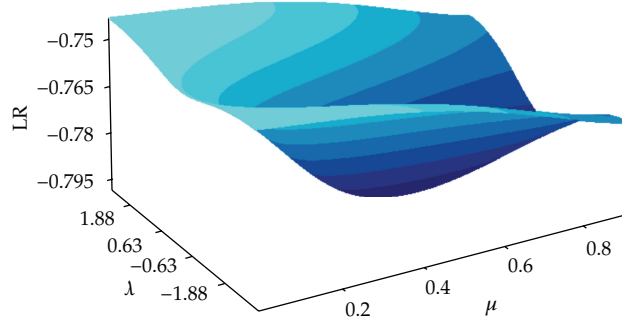
Figure 3: The time-frequency plot of the large-deviation probabilities of Λ_T.
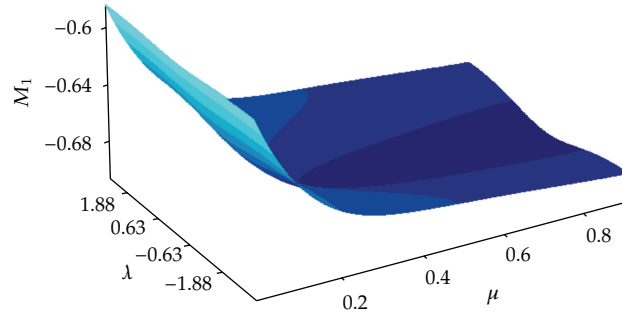

Figure 4: The time-frequency plot of the large-deviation probabilities of $M_{1,T}$.

$M_{3,T}$ are violated by that of the time-varying spectral density. Hence, we can conclude the misspecifications seriously affect our discrimination.

Appendix

We sketch the proofs of Theorems 3.3–3.5. First, we summarize the assumptions used in this paper.

Assumption A.1. (i) Suppose that $A : [0,1] \times \mathbb{R} \to \mathbb{C}$ is a 2π-periodic function with $A(u,\lambda) = \overline{A(u,-\lambda)}$ which is differentiable in u and λ with uniformly bounded derivative $(\partial/\partial u)(\partial/\partial \lambda)A$. $f_A(u,\lambda) \equiv |A(u,\lambda)|^2$ denotes the time-varying spectral density. $A_{t,T}^{\circ} : \mathbb{R} \to \mathbb{C}$ are 2π-periodic functions with

$$\sup_{t,\lambda}\left| A_{t,T}^{\circ}(\lambda) - A\left(\frac{t}{T},\lambda\right)\right| \leq KT^{-1}, \tag{A.1}$$

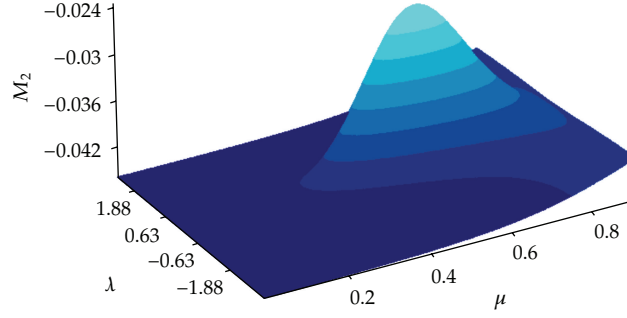(ii) suppose that $\mu : [0,1] \to \mathbb{R}$ is differentiable with uniformly bounded derivative.

Figure 5: The time-frequency plot of the large-deviation probabilities of $M_{2,T}$.
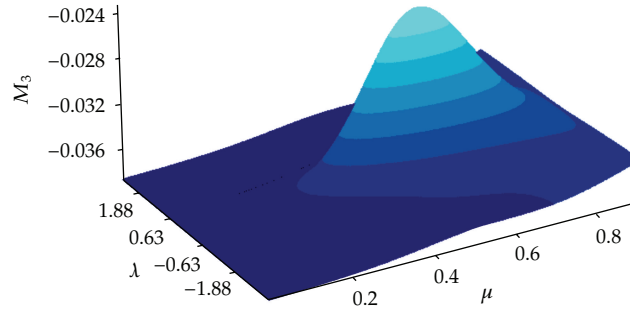

Figure 6: The time-frequency plot of the large-deviation probabilities of $M_{3,T}$.

We introduce the following matrices (see Dahlhaus [4] p.154 for the detailed definition):

$$\mathbf{W}_T(\phi) = \frac{S}{N} \sum_{j=1}^{M} \mathbf{K}_T^{(j)'} \mathbf{W}_T^{(j)}(\phi) \mathbf{K}_T^{(j)}, \tag{A.2}$$

where

$$\mathbf{W}_T^{(j)}(\phi) = \left\{ \int_{-\pi}^{\pi} \phi(u_j, \lambda) \exp(i\lambda(k-l)) d\lambda \right\}_{k,l=1,\dots,L_j}, \tag{A.3}$$

and $\mathbf{K}_T^{(j)} = (\mathbf{0}_{j1}, \mathbf{I}_{L_j}, \mathbf{0}_{j2})$. According to Lemmata 4.4 and 4.7 of Dahlhaus [4], we can see that

$$\|\mathbf{\Sigma}_T(A, A)\| \leq C + o(1), \qquad \left\|\mathbf{\Sigma}_T(A, A)^{-1}\right\| \leq C + o(1), \tag{A.4}$$

and $\mathbf{W}_T(f_A)$ and $\mathbf{W}_T(\{4\pi^2 f_A\}^{-1})$ are the approximations of $\mathbf{\Sigma}_T(A, A)$ and $\mathbf{\Sigma}_T(A, A)^{-1}$, respectively. We need the following lemmata which are due to Dahlhaus [3, 4]. Lemma A.2 is Lemma A.5 of Dahlhaus [3] and Lemma A.3 is Theorem 3.2 (ii) of Dahlhaus [4].

Lemma A.2. *Let* $k \in \mathbb{N}$, A_l, B_l *fulfill Assumption A.1 (i) and* μ_1, μ_2 *fulfill Assumption A.1 (ii).* *Let* $\mathbf{\Sigma}_l = \mathbf{\Sigma}_T(A_l, A_l)$ *or* $\mathbf{W}_T(f_{A_l})$. *Furthermore, let* $\mathbf{\Gamma}_l = \mathbf{\Sigma}_T(B_l, B_l)$, $\mathbf{W}_T(\{4\pi^2\}^{-1}f_{B_l})$ *or* $\mathbf{\Gamma}_l^{-1} = \mathbf{W}_T(\{4\pi^2 f_{B_l}\}^{-1})$. *Then we have*

$$
T^{-1}\operatorname{tr}\left\{\prod_{l=1}^{k}\mathbf{\Gamma}_l^{-1}\mathbf{\Sigma}_l\right\} = \frac{1}{2\pi}\int_0^1\int_{-\pi}^{\pi}\left\{\prod_{l=1}^{k}\frac{f_{A_l}(u,\lambda)}{f_{B_l}(u,\lambda)}\right\}d\lambda\,du + O\left(T^{-1/2}\log^{2k+2}T\right),
$$

$$
T^{-1}\boldsymbol{\mu}_{1,T}'\left\{\prod_{l=1}^{k-1}\mathbf{\Gamma}_l^{-1}\mathbf{\Sigma}_l\right\}\mathbf{\Gamma}_k^{-1}\boldsymbol{\mu}_{2,T} = \frac{1}{2\pi}\int_0^1\left\{\prod_{l=1}^{k-1}\frac{f_{A_l}(u,0)}{f_{B_l}(u,0)}\right\}f_{B_k}(u,0)^{-1}\mu_1(u)\mu_2(u)du
$$
$$
\qquad\qquad + O\left(T^{-1/2}\log^{2k+2}T\right).
$$

(A.5)

Lemma A.3. *Let* $D°$ *be the transfer function of a locally stationary process* $\{Z_{t,T}\}$, *where the corresponding* D *is bounded from below and has uniformly bounded derivative* $(\partial/\partial u)(\partial/\partial\lambda)D$. $f_D(u,\lambda) \equiv |D(u,\lambda)|^2$ *denotes the time-varying spectral density of* $Z_{t,T}$. *Then, for* $\mathbf{\Sigma}_T(d) \equiv \mathbf{\Sigma}_T(D,D)$, *we have*

$$
\lim_{T\to\infty}T^{-1}\log|\mathbf{\Sigma}_T(d)| = \frac{1}{2\pi}\int_0^1\int_{-\pi}^{\pi}\log 2\pi f_D(u,\lambda)d\lambda\,du.
$$

(A.6)

We also remark that if $\mathbf{U}_T$ and $\mathbf{V}_T$ are real nonnegative symmetric matrices, then

$$
\operatorname{tr}\{\mathbf{U}_T\mathbf{V}_T\} \leq \operatorname{tr}\{\mathbf{U}_T\}\|\mathbf{V}_T\|.
$$

(A.7)

Proof of Theorems 3.3–3.5. We need the cumulant generating function of the quadratic form in normal variables $\mathbf{X}_T \sim \mathcal{N}(\boldsymbol{\nu}_T^{(j)}, \mathbf{\Sigma}_T^{(j)})$. It is known that the quadratic form $S_T^{(j)} \equiv \mathbf{X}_T'\mathbf{H}_T\mathbf{X}_T + \mathbf{h}_T'\mathbf{X}_T - E^{(j)}(\mathbf{X}_T'\mathbf{H}_T\mathbf{X}_T + \mathbf{h}_T'\mathbf{X}_T)$ has cumulant generating function $\log E^{(j)}(e^{\omega S_T^{(j)}})$ equals to

$$
-\frac{1}{2}\log\left|\mathbf{\Sigma}_T^{(j)}\right| - \frac{1}{2}\log\left|\mathbf{\Sigma}_T^{(j)^{-1}} - 2\omega\mathbf{H}_T\right| - \omega\operatorname{tr}\left\{\mathbf{H}_T\mathbf{\Sigma}_T^{(j)}\right\}
$$
$$
+\frac{1}{2}\omega^2\left\{\mathbf{h}_T + 2\mathbf{H}_T\boldsymbol{\mu}_T^{(j)}\right\}'\left\{\mathbf{\Sigma}_T^{(j)^{-1}} - 2\omega\mathbf{H}_T\right\}^{-1}\left\{\mathbf{h}_T + 2\mathbf{H}_T\boldsymbol{\mu}_T^{(j)}\right\}
$$

(A.8)

(see Mathai and Provost [11] Theorem 3.2a.3). Theorems 3.3, 3.4, and 3.5 correspond to the respective cases:

$$\mathbf{H}_T(Q) = \mathbf{G}_T^{-1}, \qquad \mathbf{h}_T(Q) = \mathbf{g}_T;$$

$$\mathbf{H}_T(\Lambda) = \frac{1}{2}\left(\boldsymbol{\Sigma}_T^{(1)^{-1}} - \boldsymbol{\Sigma}_T^{(2)^{-1}}\right), \qquad \mathbf{h}_T(\Lambda) = -\boldsymbol{\Sigma}_T^{(1)^{-1}}\boldsymbol{\mu}_T^{(1)} + \boldsymbol{\Sigma}_T^{(2)^{-1}}\boldsymbol{\mu}_T^{(2)};$$

$$\mathbf{H}_T(M_1) = \frac{1}{2}\left(\boldsymbol{\Sigma}_T^{(1)^{-1}} - \boldsymbol{\Sigma}_T^{(2)^{-1}}\right), \qquad \mathbf{h}_T(M_1) = \mathbf{0}; \tag{A.9}$$

$$\mathbf{H}_T(M_2) = \frac{1}{2}\left(\tilde{\boldsymbol{\Sigma}}_T^{(1)-1} - \tilde{\boldsymbol{\Sigma}}_T^{(2)-1}\right), \qquad \mathbf{h}_T(M_2) = -\tilde{\boldsymbol{\Sigma}}_T^{(1)-1}\boldsymbol{\mu}_T^{(1)} + \tilde{\boldsymbol{\Sigma}}_T^{(2)-1}\boldsymbol{\mu}_T^{(2)};$$

$$\mathbf{H}_T(M_3) = \frac{1}{2}\left(\tilde{\boldsymbol{\Sigma}}_T^{(1)-1} - \tilde{\boldsymbol{\Sigma}}_T^{(2)-1}\right), \qquad \mathbf{h}_T(M_3) = \mathbf{0}.$$

We prove Theorem 3.5 for $M_{3,T}$ (under Π_1) only. Theorems 3.3 and 3.4 are similarly obtained. In order to use the Gärtner-Ellis theorem, consider

$$\psi_T^{(j)}(\omega) \equiv T^{-1}\log\left[E^{(j)}\left\{\exp\left(\omega S_T^{(j)}(M_3)\right)\right\}\right]. \tag{A.10}$$

Setting $\mathbf{H}_T = \mathbf{H}_T(M_3)$ and $\mathbf{h}_T = \mathbf{h}_T(M_3)$ in (A.8), we have under Π_1 the following:

$$\begin{aligned}
\psi_T^{(1)}(\omega) = &-(2T)^{-1}\left[\log\left|\boldsymbol{\Sigma}_T^{(1)}\right| - \log\left|\boldsymbol{\Sigma}_T^{(1)^{-1}} - \omega\left(\tilde{\boldsymbol{\Sigma}}_T^{(1)-1} - \tilde{\boldsymbol{\Sigma}}_T^{(2)-1}\right)\right|\right. \\
&\left. - \omega\operatorname{tr}\left\{\left(\tilde{\boldsymbol{\Sigma}}_T^{(1)-1} - \tilde{\boldsymbol{\Sigma}}_T^{(2)-1}\right)\boldsymbol{\Sigma}_T^{(1)}\right\} + \omega^2\boldsymbol{\mu}_T^{(1)'}\left\{\tilde{\boldsymbol{\Sigma}}_T^{(1)-1} - \tilde{\boldsymbol{\Sigma}}_T^{(2)-1}\right\}\right. \\
&\left. \times\left\{\boldsymbol{\Sigma}_T^{(1)^{-1}} - \omega\left(\tilde{\boldsymbol{\Sigma}}_T^{(1)-1} - \tilde{\boldsymbol{\Sigma}}_T^{(2)-1}\right)\right\}^{-1}\left\{\tilde{\boldsymbol{\Sigma}}_T^{(1)-1} - \tilde{\boldsymbol{\Sigma}}_T^{(2)-1}\right\}\boldsymbol{\mu}_T^{(1)}\right].
\end{aligned} \tag{A.11}$$

Using the inequality (A.7), we then replace $\{\boldsymbol{\Sigma}_T^{(1)^{-1}} - \omega(\tilde{\boldsymbol{\Sigma}}_T^{(1)-1} - \tilde{\boldsymbol{\Sigma}}_T^{(2)-1})\}$ by $\mathbf{W}_T(f_0^{(1)}f_0^{(2)} - \omega f^{(1)}\{f_0^{(2)} - f_0^{(1)}\}/4\pi^2 f^{(1)}f_0^{(1)}f_0^{(2)})$, where $f_0^{(j)}$ denote $f^{(j)}(0,\lambda)$, $j = 1,2$, that is, we obtain the approximation

$$\begin{aligned}
-(2T)^{-1}&\left[-\log\left|\boldsymbol{\Sigma}_T^{(1)}\right| - \log\left|\mathbf{W}_T\left(\frac{f_0^{(1)}f_0^{(2)} - \omega f^{(1)}\left\{f_0^{(2)} - f_0^{(1)}\right\}}{4\pi^2 f^{(1)}f_0^{(1)}f_0^{(2)}}\right)\right| - \omega\operatorname{tr}\left\{\left(\tilde{\boldsymbol{\Sigma}}_T^{(1)-1} - \tilde{\boldsymbol{\Sigma}}_T^{(2)-1}\right)\boldsymbol{\Sigma}_T^{(1)}\right\}\right. \\
&\left. + \omega^2\boldsymbol{\mu}_T^{(1)'}\left\{\tilde{\boldsymbol{\Sigma}}_T^{(1)-1} - \tilde{\boldsymbol{\Sigma}}_T^{(2)-1}\right\}\left\{\mathbf{W}_T\left(\frac{f_0^{(1)}f_0^{(2)} - \omega f^{(1)}\left\{f_0^{(2)} - f_0^{(1)}\right\}}{4\pi^2 f^{(1)}f_0^{(1)}f_0^{(2)}}\right)\right\}^{-1}\right. \\
&\left. \times\left\{\tilde{\boldsymbol{\Sigma}}_T^{(1)-1} - \tilde{\boldsymbol{\Sigma}}_T^{(2)-1}\right\}\boldsymbol{\mu}_T^{(1)}\right].
\end{aligned}$$

$$\tag{A.12}$$

In view of Lemmas A.2 and A.3, the above $\psi_T^{(1)}(\omega)$ converges to ψ_{M_3}, given in Theorem 3.5. Clearly, ψ_{M_3} exists for $\omega \in D_{\psi_{M_3}} = \{\omega : 1 - \omega\{(f^{(1)}(u,\lambda)/f^{(1)}(0,\lambda)) - (f^{(1)}(u,\lambda)/f^{(2)}(0,\lambda))\} > 0\}$ and is convex and continuously differentiable with respect to ω. For a sequence $\{\omega_m\} \to \omega_0 \in \partial D_{\psi_{M_3}}$ as $m \to \infty$, we can show that

$$\frac{\partial \psi_{M_3}\left(\omega_m; f^{(1)}, f^{(2)}, \mu^{(1)}, \mu^{(2)}\right)}{\partial \omega} \longrightarrow \infty, \qquad \frac{\partial \psi_{M_3}\left(0; f^{(1)}, f^{(2)}, \mu^{(1)}, \mu^{(2)}\right)}{\partial \omega} = 0. \qquad \text{(A.13)}$$

Hence, $\psi'_{M_3}(D_{\psi_{M_3}}) \supset (x, \infty)$ for every $x > 0$. Application of the Gärtner-Ellis theorem completes the proof. □

Acknowledgments

References

[1] J. A. Bucklew, *Large Deviation Techniques in Decision, Simulation, and Estimation*, Wiley, New York, NY, USA, 1990.

[2] T. Sato, Y. Kakizawa, and M. Taniguchi, "Large deviation results for statistics of short- and long-memory Gaussian processes," *Australian & New Zealand Journal of Statistics*, vol. 40, no. 1, pp. 17–29, 1998.

[3] R. Dahlhaus, "Maximum likelihood estimation and model selection for locally stationary processes," *Journal of Nonparametric Statistics*, vol. 6, no. 2-3, pp. 171–191, 1996.

[4] R. Dahlhaus, "On the Kullback-Leibler information divergence of locally stationary processes," *Stochastic Processes and their Applications*, vol. 62, no. 1, pp. 139–168, 1996.

[5] G. Chandler and W. Polonik, "Discrimination of locally stationary time series based on the excess mass functional," *Journal of the American Statistical Association*, vol. 101, no. 473, pp. 240–253, 2006.

[6] K. Sakiyama and M. Taniguchi, "Discriminant analysis for locally stationary processes," *Journal of Multivariate Analysis*, vol. 90, no. 2, pp. 282–300, 2004.

[7] J. Hirukawa, "Discriminant analysis for multivariate non-Gaussian locally stationary processes," *Scientiae Mathematicae Japonicae*, vol. 60, no. 2, pp. 357–380, 2004.

[8] J. Gärtner, "On large deviations from an invariant measure," *Theory of Probability and its Applications*, vol. 22, no. 1, pp. 24–39, 1977.

[9] R. S. Ellis, "Large deviations for a general class of random vectors," *The Annals of Probability*, vol. 12, no. 1, pp. 1–12, 1984.

[10] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, Wiley, New York, NY, USA, 2nd edition, 1984.

[11] A. M. Mathai and S. B. Provost, *Quadratic Forms in Random Variables: Theory and Applications*, Marcel Dekker, New York, NY, USA, 1992.

Submit your manuscripts at
http://www.hindawi.com

Advances in
Operations Research

Advances in
Decision Sciences

Journal of
Applied Mathematics

Algebra

Journal of
Probability and Statistics

The Scientific
World Journal

International Journal of
Differential Equations

International Journal of
Combinatorics

Advances in
Mathematical Physics

Journal of
Complex Analysis

Journal of
Mathematics

Mathematical Problems
in Engineering

Abstract and
Applied Analysis

Discrete Dynamics in
Nature and Society

International
Journal of
Mathematics and
Mathematical
Sciences

Journal of
Discrete Mathematics

Journal of
Function Spaces

International Journal of
Stochastic Analysis

Journal of
Optimization