Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Confidence estimation of the covariance function of stationary and locally stationary processes

Giurcanu, Mihai and Spokoiny, Vladimir

submitted: 17. March 2002

Technical University,
Strasse des 17. Juni 136,
D-10623 Berlin
E-Mail: giurcanu@math.tu-berlin.de

Weierstrass Institute
and Humboldt University Berlin,
Mohrenstr. 39, 10117 Berlin, Germany
E-Mail: spokoiny@wias-berlin.de
URL: http://www.wias-berlin.de/~spokoiny

No. 726

Berlin 2002

1991 *Mathematics Subject Classification.* 62M10; Secondary 62G15.

Key words and phrases. covariance function estimation, confidence intervals, local stationarity .

The authors thank R. Dahlhaus and R. von Sachs for very helpful remarks and discussion.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
D — 10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

In this note we consider the problem of confidence estimation of the covariance function of a stationary or locally stationary zero mean Gaussian process. The constructed confidence intervals are based on the usual empirical covariance estimate and a special estimate of its variance. The results about coverage probability are stated in nonasymptotic way and apply for small and moderate sample size under mild conditions on the model. The presented numerical results are in agreement with the theoretical issues and demonstrate applicability of the method.

1 Introduction

Let $X_0, X_1, \ldots$ be a zero mean Gaussian process. Then its distribution is completely determined by the covariance function $r_{s,t} = \boldsymbol{E} X_s X_t$. If, in addition, the process $\{X_t\}$ is stationary, then the correlation function $r_{s,t}$ depends only upon the difference $|s-t|$, that is, $r_k = \boldsymbol{E} X_s X_{s+k}$ does not depend on s. There exists a vast literature about estimation of the covariance function in the stationary case. A natural unbiased estimator of $r_k = \boldsymbol{E} X_s X_{s+k}$ from the sample of $X_0, \ldots, X_{n-1}$ is given by

$$\widehat{r}_k = \frac{1}{n-k} \sum_{s=0}^{n-k-1} X_s X_{s+k} \tag{1.1}$$

An alternative estimator which is often used in time series analysis is given by

$$\widehat{\widehat{r}}_k = \frac{1}{n} \sum_{s=0}^{n-k-1} X_s X_{s+k}.$$

The estimator $\widehat{r}_k$ is unbiased (in the sense that $\boldsymbol{E}\widehat{r}_k = r_k$) while $\widehat{\widehat{r}}_k$ is slightly biased but has smaller variance. Both estimators are root-n normal under weak regularity conditions, that is, $\sqrt{n}\,(\widehat{r}_k - r_k)$ is asymptotically as $n \to \infty$ normal with parameters $(0, s_k^2)$ with some constant s_k^2 and similarly for $\widehat{\widehat{r}}_k$, see e.g. Brockwell and Davis (1991). This result allows to construct asymptotic confidence intervals for r_k. Indeed, if $\widehat{\sigma}_k^2$ is an estimator of $\mathrm{Var}(\widehat{r}_k)$ such that $n\widehat{\sigma}_k^2$ is a consistent estimator of s_k^2, then such a confidence interval can be taken of the form $(\widehat{r}_k - \lambda\widehat{\sigma}_k, \widehat{r}_k + \lambda\widehat{\sigma}_k)$ where λ is a proper quantile of the standard normal law. The goal of the present note is to extend this nice result in two directions. First, we aim to construct a nonasymptotic finite sample confidence interval for r_k. Second, we show that the construction continues to apply even if the assumption of stationarity is slightly violated.

The assumption of stationarity can be too restrictive for many practical applications, since it does not allow to model external perturbations (like shocks of the financial market) or slowly varying structural changes. Apart financial data, the other examples of nonstationary time series are frequently met in weather analysis, economic data, sound and speech recognition, where the records show important changes over time. So, an extension of the classical stationary modelling that allows to model non-stationary time series is required.

Different approaches for describing and modelling of locally stationary time series have been developed in the last years. Adak (1998), Ombao et al. (2002) discussed an approximation of the underlying process by piecewise stationary processes. The method is based on a data-driven segmentation of the time interval into time subintervals such that the assumption of the stationarity is not significantly violated within each subinterval. Another approach assumes a smooth change of the model structure in time. The basic idea originated from Priestley (1965) and developed by Dahlhaus (1997a, 1997b) is to assume the existence of a spectral density "model" that smoothly varies in time. The assumption allows to apply the well developed nonparametric estimation theory to the estimation of time varying parameters, time varying spectral densities (Neumann and von Sachs, 1997; Dahlhaus, 1996b) or time varying covariances (Dahlhaus, 1996c). An alternative is the use of Maximum Likelihood methods, Dahlhaus (1996b, 2000). Picard (1985), Giraitis and Leipus (1992), Rozenholc (1995), Sakiyama und Taniguchi (2000), and von Sachs und Neumann (2000) among others proposed some tests of stationarity for time series in different setups.

A disadvantage of all the above described results is their asymptotic nature. That means, that the theory applies only asymptotically, when the sample size grows to infinity. Therefore, the applicability of the mentioned results for small or moderate sample sizes is questionable. In this paper, we present nonasymptotic results concerning the estimator $\widehat{r}_k$ from (1.1) for two setups. First, we consider the classical stationary situation and show that the confidence intervals of the form $(\widehat{r}_k - \lambda\widehat{\sigma}_k, \widehat{r}_k + \lambda\widehat{\sigma}_k)$ continues to apply for moderate sample size for a proper variance estimator $\widehat{\sigma}_k^2$. Second, we extend this result to the non stationary case, when the correlation function $r_{s,t}$ of the underlying process can be approximated by a correlation function of a stationary process at the rate $n^{-1/2}$ or a smaller rate, cf. Malat, Papanicolau and Zhang (2000). All the results continue to apply almost in the same form for $\widehat{\widehat{r}}_k$.

The paper is organized as follows. Section 2 presents some properties of the estimator $\widehat{r}_k$ in the stationary case. An extension to the non stationary situation is given in Section 3. In Section 4 we introduce and study an estimator $\widetilde{\sigma}_k^2$ of $\mathrm{Var}(\widehat{r}_k)$. The confidence intervals and bands for r_k are presented in Section 5. Some numerical results demonstrating the performance of the proposed estimator are given in Section 6. The proofs are collected in Section 7.

2 Some properties of $\widehat{r}_k$ in the stationary case

Denote

$$\sigma_k^2 = \mathbf{Var}(\widehat{r}_k) = \boldsymbol{E}\,(\widehat{r}_k - r_k)^2$$

To get an explicit expression for the variance σ_k^2, some matrix notations are useful. Denote by V the covariance matrix of the vector $X = (X_j)_{j=0,\ldots,n-1}$, that is, $V = (r_{st})_{s,t=0,\ldots,n-1}$. In the stationary case, the matrix V has a Toeplitz structure: $V = (r_{s-t})_{s,t=0,\ldots,n-1}$. Here we assume $r_{-s} = r_s$ for s negative. Let also A_k be the $n\times n$-matrix with the entries $a_{st} = \frac{1}{2(n-k)}\left(\mathbf{1}_{\{s-t=k\}} + \mathbf{1}_{\{t-s=k\}}\right)$.

Proposition 2.1. *Let $(X_t)_t$ be a zero mean stationary Gaussian process. Then*

$$\widehat{r}_k = X^\top A_k X, \qquad r_k = \mathrm{tr}(A_k V), \qquad \sigma_k^2 = 2\mathrm{tr}(A_k V)^2. \tag{2.1}$$

Moreover, suppose that the correlation function r_k satisfies the following regularity condition:

$$\sum_{k\in\mathbf{Z}} |r_k| = K_1 < \infty, \qquad 2\sum_{k\in\mathbf{Z}} |k| r_k^2 = K_2 < \infty. \tag{2.2}$$

Define

$$v_k = \frac{1}{2}\sum_{u\in\mathbf{Z}} (r_{u+k} + r_{u-k})^2.$$

Then for every $k \le [n/2]$

$$\sigma_k^2 \le v_0/(n-k).$$

Moreover, under the condition (2.2) σ_k^2 fulfills

$$\frac{n-2k}{(n-k)^2}v_k + \frac{kv_0 - 2K_2}{2(n-k)^2} \le \sigma_k^2 \le \frac{n-2k}{(n-k)^2}v_k + \frac{kv_0 + 2K_2}{2(n-k)^2}. \tag{2.3}$$

It is easy to see that v_k is close to $v_0/2$ for large k. By Proposition 2.1 this implies e.g. for n big enough and $k \le [n/3]$, that $\sigma_k^2 \ge \frac{v_0}{4(n-k)}$. It is therefore non restrictive to suppose the following lower bound for the variance σ_k^2:

$$\sigma_k^2 \ge s_0^2/(n-k), \qquad \forall k \le [n/3] \tag{2.4}$$

for some positive s_0.

The next result states an important probability bound for the deviation of $\widehat{r}_k$ from the true value r_k.

Proposition 2.2. *Let conditions (2.2) and (2.4) hold true. Then, for every* $k \le [n/3]$*,*

$$\boldsymbol{P}\left(\pm\sigma_k^{-1}\left(\widehat{r}_k - \boldsymbol{E}\widehat{r}_k\right) \ge \lambda\right) \le e^{-\lambda^2/4}$$

for all positive λ, n *satisfying the condition*

$$\lambda \le 2\sigma_k/(3\|A_k V\|_\infty) \tag{2.5}$$

or a stronger condition $\lambda \le C_1\sqrt{n-k}$ *with* $C_1 = 2s_0/(3K_1)$.

This result is useful for constructing a non-asymptotic confidence interval for the estimator $\widehat{r}_k$. Unfortunately, we cannot apply it directly since the variance σ_k^2 is typically unknown. In the next section we propose an estimator for the variance σ_k^2 and construct non-asymptotic confidence intervals for the covariance r_k.

The result of Proposition 2.2 allows to bound the standardized errors $\sigma_k^{-1}(\widehat{r}_k - r_k)$ uniformly over $k \le m_n$ for a given m_n.

Proposition 2.3. *Let conditions (2.2) and (2.4) hold true and, for a given* $m_n \le [n/3]$*, let* $\lambda \le C_1\sqrt{n-m_n}$ *with* $C_1 = 2s_0/(3K_1)$*. Then*

$$\boldsymbol{P}\left(\sup_{k=0,\dots,m_n-1} \sigma_k^{-1}\left|\widehat{r}_k - \boldsymbol{E}\widehat{r}_k\right| \ge \lambda\right) \le m_n e^{-\lambda^2/4}.$$

Propositions 2.3 and 2.1 imply for λ_n satisfying $\lambda_n \le C_1\sqrt{n-m_n}$

$$\boldsymbol{P}\left(\sup_{k=0,\dots,m_n-1} \left|\widehat{r}_k - r_k\right| \ge \frac{\lambda_n\sqrt{v_0 n}}{n-m_n}\right) \le m_n e^{-\lambda_n^2/4}. \tag{2.6}$$

We see that for the choice, say, $\lambda_n = 2\sqrt{\log n}$ the probability for $|\widehat{r}_k - r_k|$ being larger than $\lambda_n\sqrt{v_0 n}/(n-m_n)$ is at most m_n/n which is small provided that n is big enough.

3 Some properties of $\widehat{r}_k$ in the non-stationary case

In this section we consider the situation corresponding to the analysis of a locally stationary process $\{X_t\}$. The latter can be understood in the sense that the assumption of stationarity is approximately fulfilled within a local interval of every time instant. More precisely, we consider the situation that the correlation function $r_{s,t}$ of the process X_t within the considered interval is close to a correlation function of a stationary process. This suggests to proceed similarly to the stationary case and consider $\widehat{r}_k$ from (1.1) as the estimator of the correlation function of the approximating stationary process. In the rest of this section we discuss some properties of $\widehat{r}_k$ in the locally stationary situation.

The next assertion can be proved similarly to Proposition 2.1. Let V denote the $n \times n$-matrix whose elements are $r_{s,t}$ for $s,t = 0,\ldots,n-1$.

Proposition 3.1. *Let $(X_t)_t$ be a zero mean Gaussian time series. Then the estimator $\widehat{r}_k$ from (1.1) fulfills $\widehat{r}_k = X^\top A_k X$, $\boldsymbol{E}\widehat{r}_k = \operatorname{tr}(A_k V)$ and $\mathbf{Var}(\widehat{r}_k) = 2\operatorname{tr}(A_k V)^2$. Moreover, if there exists a sequence of positive numbers $\{r_k^*, k \in \mathbb{Z}\}$ such that $|r_{s,s+k}| \le r_k^*$ for all s,k and*

$$\sum_{s\in\mathbb{Z}} r_s^* = K_1 < \infty, \qquad 2\sum_{s\in\mathbb{Z}} |s| r_s^{*2} = K_2 < \infty. \tag{3.1}$$

then

$$\sigma_k^2 = \mathbf{Var}(\widehat{r}_k) \le v_0/(n-k)$$

where $v_0 = 2\sum_{s\in\mathbb{Z}}(r_s^*)^2$.

Under stationarity $\widehat{r}_k$ is an unbiased estimator of the true covariance r_k and the accuracy of estimation is of order σ_k, where $\sigma_k^2 = \mathbf{Var}(\widehat{r}_k)$. The assumption of local stationarity would mean that the departure from stationarity within the considered time interval is statistically insignificant, that is, smaller in order than the accuracy of estimation. Therefore, it is natural to measure the deviation from stationarity for the process (X_t) by the value

$$\delta_n = \inf_{\overline{r}} \max_{s,k\ge 0, s+k<n} \sigma_k^{-1} \left| r_{s,s+k} - \overline{r}_k \right| \tag{3.2}$$

where the infimum is taken over the class of the covariance functions $\overline{r}_k$ of stationary processes. In what follows we assume that the infimum is attained and there exists a covariance function $\overline{r}$ such that

$$\delta_n = \max_{s,k\ge 0,\ s+k<n} \sigma_k^{-1} \left| r_{s,s+k} - \overline{r}_k \right|. \tag{3.3}$$

Without loss of generality we also suppose that $\bar{r}_k \le r_k^*$ where r_k^* is from (3.1).

A straightforward corollary of the definition (3.2) is the following bound for the 'bias' $\left|\boldsymbol{E}\widehat{r}_k - \bar{r}_k\right|$ and for the quadratic risk of $\widehat{r}_k$.

Proposition 3.2. *Let $\{X_t\}$ be a Gaussian time series and let δ_n be defined in (3.2). Then*

$$\left|\boldsymbol{E}\widehat{r}_k - \bar{r}_k\right| \le \sigma_k \delta_n, \qquad \boldsymbol{E}\left(\widehat{r}_k - \bar{r}_k\right)^2 \le \sigma_k^2(1+\delta_n^2). \tag{3.4}$$

Similarly to the stationary case, we suppose that

$$\sigma_k^2 = \mathbf{Var}(\widehat{r}_k) \ge s_0^2/(n-k), \qquad k \le [n/3]. \tag{3.5}$$

Proposition 2.2 continues to hold in the non stationary situation without any change.

Proposition 3.3. *Let conditions (3.1) and (3.5) hold true. Then for every $k \le [n/3]$*

$$\begin{aligned} \boldsymbol{P}\left(\pm\sigma_k^{-1}\left(\widehat{r}_k - \boldsymbol{E}\widehat{r}_k\right) \ge \lambda\right) &\le e^{-\lambda^2/4}, \\ \boldsymbol{P}\left(\pm\sigma_k^{-1}\left(\widehat{r}_k - \bar{r}_k\right) \ge \lambda + \delta_n\right) &\le e^{-\lambda^2/4}, \end{aligned}$$

for all positive λ, n satisfying the condition

$$\lambda \le \frac{2\sigma_k}{3\|A_k V\|_\infty}$$

or a stronger condition $\lambda \le C_1\sqrt{n-k}$ with $C_1 = 2s_0/(3K_1)$.

Similarly to Proposition 2.3 we can derive a uniform bound for the deviation $|\widehat{r}_k - \bar{r}_k|$.

Proposition 3.4. *Let $m_n \le [n/3]$ and let the conditions (3.1) and (3.5) be fulfilled. If $\lambda \le C_1\sqrt{n-m_n}$, then*

$$\begin{aligned} \boldsymbol{P}\left(\sup_{k=0,\ldots,m_n-1} \sigma_k^{-1}\left|\widehat{r}_k - \boldsymbol{E}\widehat{r}_k\right| \ge \lambda\right) &\le m_n e^{-\lambda^2/4}, \\ \boldsymbol{P}\left(\sup_{k=0,\ldots,m_n-1} \sigma_k^{-1}\left|\widehat{r}_k - \bar{r}_k\right| \ge \lambda + \delta_n\right) &\le m_n e^{-\lambda^2/4}. \end{aligned}$$

The latter result and Proposition 3.1 imply

Proposition 3.5. *Under the conditions (3.1) and (3.5) and $\lambda_n \le C_1\sqrt{n-m_n}$*

$$\boldsymbol{P}\left(\sup_{k=0,\ldots,m_n-1} \left|\widehat{r}_k - \bar{r}_k\right| \ge \frac{(\lambda_n + \delta_n)\sqrt{nv_0}}{n-m_n}\right) \le m_n e^{-\lambda_n^2/4}.$$

4 Estimation of the variance σ_k^2

To construct the confidence interval for the estimator $\widehat{r}_k$ we need to estimate the variance σ_k^2. We again discuss first the *stationary* case. Due to (2.1) it holds $\sigma_k^2 = 2\mathrm{tr}(A_k V)^2$ where $V = (r_{s-t}, s, t = 0, \dots, n-1)$. A natural estimator of σ_k^2 is $\widehat{\sigma}_k^2 = 2\mathrm{tr}(A_k \widehat{V})^2$ where $\widehat{V}$ is the matrix with the entries $\widehat{r}_{|s-t|}$. Note however that $\widehat{r}_k$ estimates the true value r_k with the rate $n^{-1/2}$. At the same time, because of our hypothesis (2.2) there exists a subsequence $(r(k_n))_{n\in N}$ such that $r(k_n) \le 1/k_n$. Therefore, for large k it is more reasonable to simply estimate r_k by zero. We therefore fix some $m_n' = [n^\alpha]$ for $\alpha \in (0, 1/2)$ and apply the following estimator

$$\widetilde{\sigma}_k^2 = 2\mathrm{tr}(A_k \widetilde{V})^2 \tag{4.1}$$

where $\widetilde{V}$ is the matrix with the entries $\widetilde{r}_{|s-t|}$:

$$\widetilde{r}_s = \begin{cases} \widehat{r}_s, & s = 0, \dots, m_n' - 1, \\ 0, & s = m_n', \dots, n-1. \end{cases} \tag{4.2}$$

By Proposition 2.1 σ_k^2 is of order n^{-1}. Below in this section we show that $\widetilde{\sigma}_k^2$ estimates the true variance σ_k^2 with the accuracy of smaller order than n^{-1}, that is, $n\left(\widetilde{\sigma}_k^2 - \sigma_k^2\right) = o_n(1)$. Moreover, this result continues to apply in the non stationary situation when δ_n from (3.2) is small.

Due to (2.6) there exists a random set A with $\boldsymbol{P}(A) \ge 1 - m_n' e^{-\lambda_n^2/4}$ such that

$$|\widehat{r}_s - r_s| \le \frac{\lambda_n \sqrt{v_0 n}}{n - m_n'} \tag{4.3}$$

for all $s \le m_n' - 1$. Define

$$\varepsilon = \frac{\lambda_n \sqrt{v_0 n}}{n - m_n'}.$$

Then $|\widehat{r}_s - r_s| \le \varepsilon$ on A for all $s \le m_n' - 1$. We now show that a good quality in estimating the covariance function r_s implies automatically a good quality of the estimator $\widetilde{\sigma}_k^2$.

Proposition 4.1. *Let $\{X_t\}$ be a zero mean stationary Gaussian time series such that (2.2) holds true. If $|\widehat{r}_s - r_s| \le \varepsilon$ for $s \le m_n' - 1$, then for every number $m_n \le [n/3]$*

$$n\widetilde{\sigma}_k^2 \ge n\sigma_k^2 - \psi_n \qquad \text{and} \qquad n\left|\widetilde{\sigma}_k^2 - \sigma_k^2\right| \le \psi_n' \qquad \text{for all } k = 0, 1, \dots, m_n - 1$$

where

$$\psi_n = \frac{4n^2 K_1 \tau_n}{(n - m_n)^2}, \qquad \psi_n' = \frac{4n^2 K_1 \tau_n}{(n - m_n)^2} + \frac{2n\tau_n^2}{n - m_n} = \psi_n + \frac{2n\tau_n^2}{n - m_n} \tag{4.4}$$

and

$$\tau_n = m'_n\varepsilon + \sum_{s=m'_n}^{n-1} |r_s|.$$

Remark 4.1. It is easy to see that, if $m'_n = [n^\alpha]$ for $\alpha < 1/2$, then $\psi_n, \psi'_n = o_n(1)$.

Now we consider the *non stationary* situation. We suppose that there exists a correlation function $\overline{r}_k$ corresponding to some stationary process such that $\left|r_{s,s+k} - \overline{r}_k\right| \le \delta_n \sigma_k$ for all possible pairs (s,k). The value δ_n measures the departure from stationarity and local stationarity within the considered time interval can be roughly understood in the sense that δ_n is small.

Let $\overline{V}$ be the matrix whose elements are $\overline{r}_{|s-t|}$ for $s,t = 0,\ldots,n-1$. Then $\overline{\sigma}_k^2 = 2\mathrm{tr}(A_k\overline{V})^2$ is an approximation of the variance $\sigma_k^2 = 2\mathrm{tr}(A_kV)^2$. The result of Proposition 4.1 can be easily extended to the locally stationary situation in the sense that $\widetilde{\sigma}_k^2$ is a reasonable estimator of $\overline{\sigma}_k^2$. Namely, define

$$\overline{\varepsilon} = \frac{(\lambda_n + \delta_n)\sqrt{v_0 n}}{n - m'_n}. \tag{4.5}$$

Due to Proposition 3.5 there exists a random set A with $\boldsymbol{P}(A) \ge 1 - m'_n e^{-\lambda_n^2/4}$ such that $|\widehat{r}_k - \overline{r}_k| \le \overline{\varepsilon}$ for all $k \le m'_n - 1$.

Proposition 4.2. *Let $\{X_t\}$ be a zero mean Gaussian time series such that (3.1) holds true. If $|\widehat{r}_s - \overline{r}_s| \le \overline{\varepsilon}$ for $s < m'_n$, then for every $m_n \le [n/3]$*

$$n\widetilde{\sigma}_k^2 \ge n\overline{\sigma}_k^2 - \psi_n \qquad \textit{and} \qquad n\left|\widetilde{\sigma}_k^2 - \overline{\sigma}_k^2\right| \le \psi'_n \qquad k = 0,\ldots,m_n - 1$$

where ψ_n and ψ'_n are defined in (4.4) with

$$\tau_n = m'_n\overline{\varepsilon} + \sum_{s=m'_n}^{n-1} r_s^*. \tag{4.6}$$

For constructing the confidence intervals, we need to bound from below the difference $\widetilde{\sigma}_k^2 - \sigma_k^2$.

Proposition 4.3. *Under the conditions of Proposition 4.2,*

$$n\widetilde{\sigma}_k^2 \ge n\sigma_k^2 - \psi''_n$$

where, for ψ_n defined by (4.4) and (4.6),

$$\psi''_n = \psi_n + \frac{2n^2K_1\delta_n\sqrt{v_0}}{(n-m_n)^{5/2}}.$$

Remark 4.2. If δ_n from (3.2) is bounded by an absolute constant or, moreover, $\delta_n = o_n(1)$, then similarly to the stationary situation it follows that $\psi_n, \psi'_n, \psi''_n = o_n(1)$.

5 Confidence interval for the covariance function

The next two theorems can be used for constructing non-asymptotic confidence intervals and bands for the sample autocovariance function in the stationary case. Afterwards, these results are extended to the non stationary situation.

In what follows we suppose that $\widetilde{\sigma}_k^2$ is the estimator of the variance σ_k^2 from (4.1) for a given $m_n' \le [n^\alpha]$ for $\alpha \in (0, 1/2)$, and k varies from zero to m_n with a prescribed $m_n \le [n/3]$.

Theorem 5.1. *Let $\{X_t\}$ be a zero mean stationary Gaussian time series satisfying the conditions (2.2) and (2.4). Moreover, let n, k, λ_n and λ be such that $\lambda_n \le C_1\sqrt{n - m_n'}$, $(n-k)\psi_n < ns_0^2$ and $\lambda \le C_1\sqrt{n-k}$ with $C_1 = 2s_0/(3K_1)$ and ψ_n from (4.4). Then, for every $k \le [n/3]$,*

$$\boldsymbol{P}\left(|\widehat{r}_k - r_k| \ge \lambda' \widetilde{\sigma}_k\right) \le e^{-\lambda^2/4} + m_n' e^{-\lambda_n^2/4} \tag{5.1}$$

where $\lambda' = \lambda\left(1 - \frac{(n-k)\psi_n}{ns_0^2}\right)^{-1/2}$.

Similarly one can describe the confidence bands for the covariance function r_k.

Theorem 5.2. *Let $\{X_t\}$ be a zero mean stationary Gaussian time series satisfying the conditions (2.2) and (2.4). Moreover, let n, λ_n and λ be such that $\lambda_n \le C_1\sqrt{n - m_n'}$, $\psi_n < s_0^2$ and $\lambda \le C_1\sqrt{n - m_n}$. Then*

$$\boldsymbol{P}\left(\sup_{k=0,\dots,m_n-1} \frac{|\widehat{r}_k - r_k|}{\widetilde{\sigma}_k} \ge \lambda'\right) \le m_n e^{-\lambda^2/4} + m_n' e^{-\lambda_n^2/4} \tag{5.2}$$

where $\lambda' = \lambda\left(1 - \psi_n/s_0^2\right)^{-1/2}$.

Remark 5.1. Since $\psi_n = o_n(1)$, the condition $\psi_n/s_0^2 < 1$ is fulfilled for n big enough. Moreover, $|\lambda'/\lambda - 1| = o_n(1)$.

Now we briefly discuss the *non stationary* case.

Theorem 5.3. *Let $\{X_t\}$ be a zero mean Gaussian time series satisfying the conditions (3.1) and (3.5). Moreover, let n, k, λ_n and λ fulfill $\lambda_n \le C_1\sqrt{n - m_n'}$, $(n-k)\psi_n < ns_0^2$ and $\lambda \le C_1\sqrt{n-k}$ with $C_1 = 2s_0/(3K_1)$. Then, for every $k \le [n/3]$,*

$$\boldsymbol{P}\left(\left|\widehat{r}_k - \overline{r}_k\right| \ge \lambda' \widetilde{\sigma}_k\right) \le e^{-(\lambda - \delta_n)^2/4} + m_n' e^{-\lambda_n^2/4}. \tag{5.3}$$

where λ' is defined by $\lambda'\sqrt{1 - \frac{(n-k)\psi_n''}{ns_0^2}} = \lambda$ with ψ_n'' from Proposition 4.3.

Finally, we state the confidence band result for the non stationarity case.

Theorem 5.4. *Let $\{X_t\}$ be a zero mean Gaussian time series satisfying the conditions (3.1) and (3.5). Moreover, let n, λ_n and λ be such that $\lambda_n \leq C_1\sqrt{n - m'_n}$, $\psi_n < s_0^2$ and $\lambda \leq C_1\sqrt{n - m_n}$ with $C_1 = 2s_0/(3K_1)$. Then*

$$\boldsymbol{P}\left(\sup_{k=0,\ldots,m_n-1} \frac{\left|\widehat{r}_k - \overline{r}_k\right|}{\widetilde{\sigma}_k} \geq \lambda'\right) \leq m_n e^{-(\lambda-\delta_n)^2/4} + m'_n e^{-\lambda_n^2/4} \tag{5.4}$$

where λ' is defined by $\lambda'\sqrt{1 - \psi''_n/s_0^2} = \lambda$ with ψ''_n from Proposition 4.3.

6 Simulation results

In this section we present some numerical results demonstrating the finite size performance of the proposed procedure. We present the results for two artificial models. The first one is stationary and it is described by a classical autoregressive model. The other one is nonstationary and obtained from the first one by allowing the coefficients to vary with time.

The stationary process is given by the following linear difference equation:

$$X_t = \frac{1}{2}X_{t-1} + \frac{1}{3}X_{t-2} - \frac{1}{6}X_{t-3} + Z_t, \text{ where } Z_t \sim N(0,1), \qquad t = 1, \ldots, n. \tag{6.1}$$

The initial values X_{-2}, X_{-1}, X_0 are generated randomly from the stationary distribution. The non stationary process is obtained from (6.1) with the coefficients "perturbed":

$$\begin{aligned} X_t &= \{1/2 + 0.1\cos(\pi t/25)\}\, X_{t-1} + \{1/3 + 0.08\sin(\pi t/15)\}\, X_{t-2} \\ &\quad - \left\{1/6 + 0.05\cos^2(\pi t/10)\right\} X_{t-3} + Z_t, \text{ where } Z_t \sim N(0,1). \end{aligned} \tag{6.2}$$

To compute the true covariances $r(k)$ for the process defined by (6.1), we use the Yule Walker equations:

$$\begin{cases} R\Phi = r \\ \sigma^2 = r_0 - \Phi' r \end{cases}$$

where $R = (r(s-t)_{s,t=1,2,3})$, $r = (r_1, r_2, r_3)$, $\Phi = (1/2, 1/3, -1/6)$, $\sigma^2 = 1$. To compute the true variances r_k for $k \geq 4$, one can use the following iterative equation:

$$r_k = \phi_1 r_{k-1} + \phi_2 r_{k-2} + \phi_3 r_{k-3}$$

The covariances r_k for the process (6.2) are computed by Monte Carlo techniques, using 10000 realizations. Afterwards, the variances of the estimators $\widehat{r}_k$ are computed by

Proposition 2.1 or Proposition 3.1. For the non stationary process, we consider the approximating autocovariance value $\overline{r}_k$ defined by:

$$\overline{r}_k = \frac{1}{n-k} \sum_{i=0}^{n-k-1} r_{i,i+k}$$

where $n = 100$ is the sample size, $k = 0, \ldots, 99$.

Figure 1 shows the boxplot for the standardized autocovarince values $\widetilde{\sigma}_k^{-1}(\widehat{r}_k - r_k)$ for $k = 0, \ldots, 7$ for the process given by (6.1) and the same for the process given by (6.2). The results are based on 200 realizations with the sample size $n = 100$.

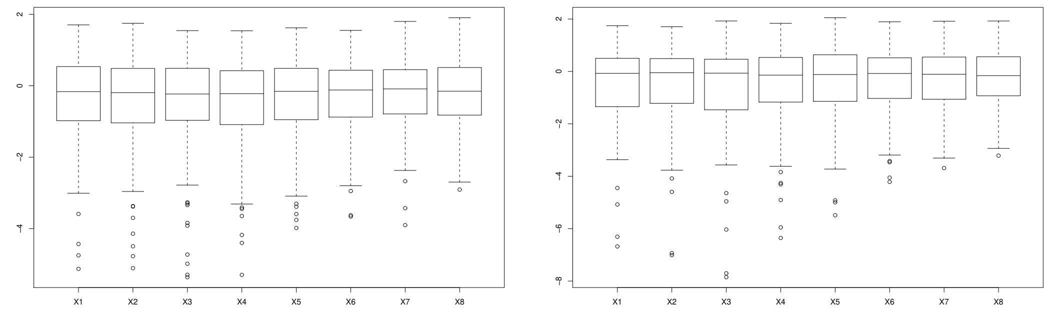

Figure 1: Boxplots of $\widetilde{\sigma}_k^{-1}(\widehat{r}_k - r_k)$ for $k = 0, \ldots, 7$ for the process (6.1), on the left, and for the process given by (6.2), on the right, from 200 realizations, with the sample size $n = 100$

Figure 2 shows the empirical probability confidence bands as function of λ, for the standardized sample variances $\widetilde{\sigma}_k^{-1}(\widehat{r}_k - r_k)$ for $k = 0, \ldots, 7$, for the processes (6.1) and (6.2). Here 500 realizations of sample size 100 were run.

To judge about the quality of the estimator $\widetilde{\sigma}_k$ of the variance of $\widehat{r}_k$, we present some results using the mean absolute error measure: $MAE_k = \widehat{\boldsymbol{E}} \left| \frac{\widetilde{\sigma}_k}{\sigma_k} - 1 \right|$, see Table 1. Here we consider different sample sizes n varying from 50 to 300 to see the dependence of the estimation quality on n.

To get some impression about how large is the departure of the process like (6.2) from stationarity we computed the value δ_n, see (3.3), for the following family of non-stationary processes indexed by the numerical parameter a:

$$\begin{aligned} X_t &= \{1/2 + 0.1a\cos(\pi t/25)\} X_{t-1} + \{1/3 + 0.08a\sin(\pi t/15)\} X_{t-2} \\ &\quad - \left\{1/6 + 0.05a\cos^2(\pi t/10)\right\} X_{t-3} + Z_t, \text{ where } Z_t \sim N(0,1) \end{aligned} \tag{6.3}$$

The covariances $\widehat{r}_k$ have been computed by Monte Carlo techiques, using 1000 realizations

n	MAE_0	MAE_1	MAE_2	MAE_3	MAE_4
50	0.374	0.369	0.401	0.377	0.380
100	0.295	0.270	0.304	0.330	0.326
200	0.229	0.251	0.246	0.287	0.276
300	0.202	0.213	0.231	0.244	0.235

Table 1: The empirical mean absolute error for the standardized variances at different lags, for the non stationary process given by (6.2), from 500 realizations

for every process and the sample size $n = 100$. The results for different values of a from 0.025 to 1 are shown in Figure 3.

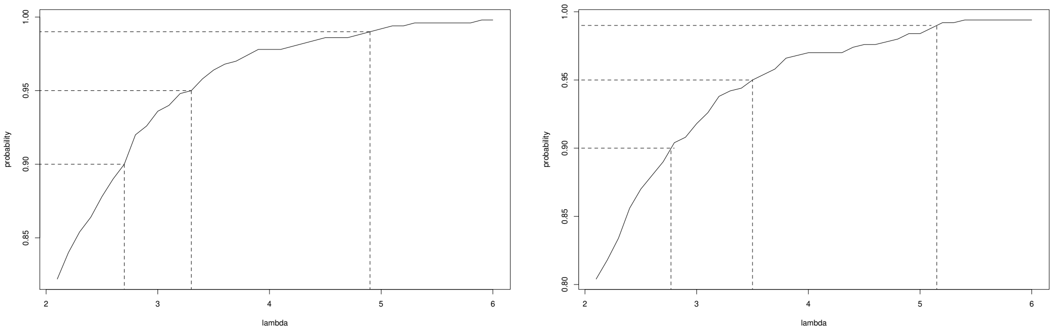

Figure 2: The empirical confidence bands probability of $\widetilde{\sigma}_k^{-1}(\widehat{r}_k - r_k),\ k = 0, \ldots, 7$ for the process given by (6.1) on the left and for the non stationary process given by (6.2) on the right, from 500 realizations with $n = 100$ for different values of λ.

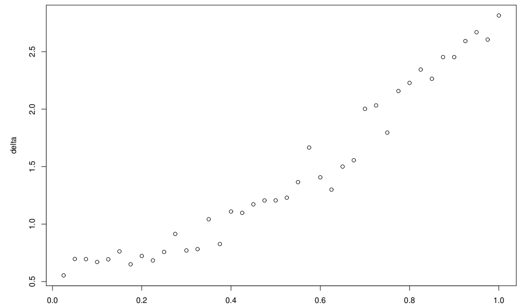

Figure 3: δ_n as a function of a, on a grid of length 0.025, for the process (6.3), the sample size $n = 100$

7 Proofs

In this section we collect the proofs of the main assertions.

Proof of Proposition 2.1

For the proof of (2.1) see e.g. Spokoiny (2002). Next

$$
\begin{aligned}
\sigma_k^2 &= 2\mathrm{tr}(A_k V)^2 = 2\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}(A_k V)_{ij}(A_k V)_{ji} \\
&= \frac{1}{2(n-k)^2}\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}\left(r_{i-k-j}\mathbf{1}_{\{i-k\geq 0\}} + r_{i+k-j}\mathbf{1}_{\{i+k\leq n-1\}}\right) \\
&\qquad \left(r_{j-k-i}\mathbf{1}_{\{j-k\geq 0\}} + r_{j+k-i}\mathbf{1}_{\{j+k\leq n-1\}}\right) \\
&= \frac{1}{2(n-k)^2}\left[\sum_{i=k}^{n-1}\sum_{j=k}^{n-1} r_{i-k-j}\, r_{j-k-i} + \sum_{i=k}^{n-1}\sum_{j=0}^{n-k-1} r_{i-k-j}^2 \right.\\
&\qquad \left. + \sum_{i=0}^{n-k-1}\sum_{j=k}^{n-1} r_{i+k-j}^2 + \sum_{i=0}^{n-k-1}\sum_{j=0}^{n-k-1} r_{i+k-j}\, r_{j+k-i}\right] \\
&= \frac{1}{(n-k)^2}\left[\sum_{i=0}^{n-k-1}\sum_{j=0}^{n-k-1} r_{i-k-j}\, r_{j-k-i} + \sum_{i=0}^{n-k-1}\sum_{j=0}^{n-k-1} r_{i-j}^2\right]
\end{aligned}
$$

This easily implies

$$
\sigma_k^2 = \frac{1}{(n-k)^2}\sum_{|u|\leq n-k-1}(n-k-|u|)\left(r_{u+k}\, r_{u-k} + r_u^2\right) \leq \frac{v_0}{n-k}.
$$

Now we bound the variance σ_k^2 from below and above to get (2.3). The following inequalities are true:

$$
\begin{aligned}
&\left|(n-k)^2\sigma_k^2 - (n-2k)\sum_{u\in\mathbf{Z}} r_{u+k}\,r_{u-k} - (n-k)\sum_{u\in\mathbf{Z}} r_u^2\right| \\
= &\left|\sum_{|u|\le n-k-1} (k-|u|) r_{u+k} r_{u-k} - \sum_{|u|\le n-k-1} |u| r_u^2\right. \\
&\left. - \sum_{|u|\ge n-k} (n-2k) r_{u+k} r_{u-k} + \sum_{|u|\ge n-k} (n-k) r_u^2\right| \\
\le &\sum_{|u|\le n-k-1} \frac{1}{2}\left|k-|u|\right| \left(r_{u+k}^2 + r_{u-k}^2\right) + \sum_{|u|\le n-k-1} |u| r_u^2 \\
&\quad + \sum_{|u|\ge n-k} \frac{1}{2}(n-2k)\left(r_{u+k}^2 + r_{u-k}^2\right) + \sum_{|u|\ge n-k} (n-k) r_u^2 \\
\le &\ 2\sum_{u\in\mathbf{Z}} |u| r_u^2 = K_2.
\end{aligned}
$$

Finally we obtain:

$$
\frac{n-2k}{(n-k)^2} v_k + \frac{k v_0 - 2K_2}{2(n-k)^2} \le \sigma_k^2 \le \frac{n-2k}{(n-k)^2} v_k + \frac{k v_0 + 2K_2}{2(n-k)^2}
$$

Proof of Proposition 2.2

It is easy to check that

$$
\|VA_k\|_\infty \le \|V\|_\infty \|A_k\|_\infty \le \frac{1}{n-k}\sum_{k\in\mathbf{Z}} |r_k| = \frac{K_1}{n-k}. \tag{7.1}
$$

An application of the general result for Gaussian quadratic forms from Spokoiny (2002) yields

$$
\begin{aligned}
&\boldsymbol{P}\left(\pm\sigma_k^{-1}(\widehat{r}_k - \boldsymbol{E}\widehat{r}_k) \ge \lambda\right) = \boldsymbol{P}\left(\pm(\widehat{r}_k - \boldsymbol{E}\widehat{r}_k) \ge \sigma_k\lambda\right) \\
&\quad = \boldsymbol{P}\left(\pm\left(X^\top A_k X - \boldsymbol{E}(X^\top A_k X)\right) \ge \sigma_k\lambda\right) \le \max\left\{e^{-\lambda^2/4}, e^{-\frac{\lambda\sigma_k}{6\|VA_k\|_\infty}}\right\}.
\end{aligned}
$$

This, under the condition $\lambda \le 2\sigma_k/(3\|VA_k\|_\infty)$, implies

$$
\boldsymbol{P}\left(\pm\sigma_k^{-1}(\widehat{r}_k - \boldsymbol{E}\widehat{r}_k) \ge \lambda\right) \le e^{-\lambda^2/4}
$$

It remains to note that in view of (2.4) and (7.1), the condition $\lambda < 2s_0/(3K_1)\sqrt{n-k}$ is sufficient for $\lambda \le 2\sigma_k/(3\|VA_k\|_\infty)$.

Proof of Proposition 2.3

Since

$$\boldsymbol{P}\left(\sup_{k=0,\ldots,m_n-1}\sigma_k^{-1}\left|\widehat{r}_k-\boldsymbol{E}\widehat{r}_k\right|\geq\lambda\right)\leq\sum_{k=0}^{m_n-1}\boldsymbol{P}\left(\left|\widehat{r}_k-\boldsymbol{E}\widehat{r}_k\right|\geq\lambda\sigma_k\right)$$

follows immediately from Proposition 2.2.

Proof of Proposition 3.2

It follows directly from (3.3) that

$$\left|\boldsymbol{E}\widehat{r}_k-\overline{r}_k\right|=\left|\frac{1}{n-k}\sum_{j=0}^{n-k-1}r_{j,j+k}-\overline{r}_k\right|\leq\frac{1}{n-k}\sum_{j=0}^{n-k-1}\left|r_{j,j+k}-\overline{r}_k\right|\leq\sigma_k\delta_n.$$

Now

$$\boldsymbol{E}\left(\widehat{r}_k-\overline{r}_k\right)^2=\mathbf{Var}(\widehat{r}_k)+\left(\boldsymbol{E}\widehat{r}_k-\overline{r}_k\right)^2\leq\sigma_k^2+\sigma_k^2\delta_n^2$$

as required.

Proof of Proposition 4.1

Let us fix an arbitrary fixed $k<m_n$ and consider the function $f(x_0,\ldots,x_{n-1}):\mathbb{R}^n\to\mathbb{R}$ given by:

$$f(x_0,\ldots,x_{n-1})=2\mathrm{tr}(A_kV)^2,$$

where $V=\left(x_{|s-t|}\,,\,s,t=0,\ldots,n-1\right)$ and $A_k=(a_{st}\,,\,s,t=0,\ldots,n-1)$ with $a_{st}=\frac{1}{2(n-k)}\left(\mathbf{1}_{\{s-t=k\}}+\mathbf{1}_{\{t-s=k\}}\right)$. It is clear that f is a quadratic form of the vector $\mathbf{x}=(x_0,\ldots,x_{n-1})^\top$ and therefore it holds for every pair $\mathbf{x},\mathbf{x}_0\in\mathbb{R}^n$

$$f(\mathbf{x})=f(\mathbf{x}_0)+(\mathbf{x}-\mathbf{x}_0)^T\nabla f(\mathbf{x}_0)+0.5(\mathbf{x}-\mathbf{x}_0)^T\nabla^2 f(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0)$$

where $\nabla^2 f$ denotes the Hessian of f. We now apply this formula for $\mathbf{x}=\widetilde{\mathbf{r}}=(\widetilde{r}_s\,,\,s=0,\ldots,n-1)$ and $\mathbf{x}_0=\mathbf{r}=(r_s\,,\,s=0,\ldots,n-1)$. This yields

$$\widetilde{\sigma}_k^2=\sigma_k^2+(\widetilde{\mathbf{r}}-\mathbf{r})^T\nabla f(\mathbf{r})+0.5(\widetilde{\mathbf{r}}-\mathbf{r})^T\nabla^2 f(\mathbf{r})(\widetilde{\mathbf{r}}-\mathbf{r}).\tag{7.2}$$

First we compute the gradient and the Hessian of the function f. Denote by E_k the matrix with the entries $e_{st}=\frac{1}{2}\left(\mathbf{1}_{\{s-t=k\}}+\mathbf{1}_{\{t-s=k\}}\right)$ for $s,t=0,\ldots,n-1$, so that

$A_k = (n-k)^{-1}E_k$. In an obvious way it holds for the matrix $V = V(\mathbf{x})$ with the entries $x_{|s-t|}$ for $s,t = 0,\dots,n-1$

$$\frac{\partial V(\mathbf{x})}{\partial x_s} = 2E_s, \qquad s = 0,\dots,n-1,$$

and thus

$$\frac{\partial f(\mathbf{x})}{\partial x_s} = \frac{\partial 2\mathrm{tr}\,(A_k V)^2}{\partial x_s} = \frac{4}{(n-k)^2}\mathrm{tr}\left(E_k V E_k \frac{\partial V(\mathbf{x})}{\partial x_s}\right) = \frac{8}{(n-k)^2}\mathrm{tr}\,(V E_k E_s E_k)\,.$$

It is easy to check that all the entries of the matrix $E_k E_s E_k$ are nonnegative and bounded by $1/2$. Hence

$$\mathrm{tr}\,(V E_k E_s E_k) \le \frac{1}{2}\sum_{s,t=0}^{n-1} |r_{s-t}| \le nK_1/2$$

and

$$\left|\frac{\partial f(\mathbf{x})}{\partial x_s}\right| \le \frac{4nK_1}{(n-k)^2}.$$

Similarly, for every pair $s,t \in 0,\dots,n-1$

$$\frac{\partial^2 f}{\partial x_s \partial x_t}(\mathbf{x}) = \frac{8}{(n-k)^2}\mathrm{tr}\,(E_s E_k E_t E_k)$$

and $\mathrm{tr}\,(E_s E_k E_t E_k) \le (n-k)/2$. Therefore

$$\left|\frac{\partial^2 f}{\partial x_s \partial x_t}(\mathbf{x})\right| \le \frac{4}{(n-k)}. \tag{7.3}$$

Now the decomposition (7.2) implies

$$\widetilde{\sigma}_k^2 \ge \sigma_k^2 - \left|(\widetilde{\mathbf{r}} - \mathbf{r})^T \nabla f(\mathbf{r})\right| \ge \sigma_k^2 - \|\widetilde{\mathbf{r}} - \mathbf{r}\|_1 \max_{s<n}\left|\frac{\partial f(\mathbf{x})}{\partial x_s}\right| \ge \sigma_k^2 - \frac{4nK_1}{(n-k)^2}\|\widetilde{\mathbf{r}} - \mathbf{r}\|_1$$

where

$$\|\widetilde{\mathbf{r}} - \mathbf{r}\|_1 = \sum_{s=0}^{n-1} |\widetilde{r}_k - r_k| \le m_n' \sup_{k \in \overline{0, m_n' - 1}} |\widehat{r}_k - r_k| + \sum_{s=m_n'}^{n-1} |r_k|.$$

This and (4.3) imply

$$n\widetilde{\sigma}_k^2 \ge n\sigma_k^2 - \psi_n$$

with $\psi_n = 4n^2 K_1 \tau_n/(n-m_n)^2$ and the first assertion follows.

Next, by (7.3)

$$\left|(\widetilde{\mathbf{r}} - \mathbf{r})^T \nabla^2 f(\mathbf{r})(\widetilde{\mathbf{r}} - \mathbf{r})\right| \le \frac{4}{n-k}\|\widetilde{\mathbf{r}} - \mathbf{r}\|_1^2 \le \frac{4\tau_n^2}{n-k}$$

and

$$n\left|\widetilde{\sigma}_k^2 - \sigma_k^2\right| \le \psi_n' = \psi_n + \frac{2n\tau_n^2}{n-m_n}$$

as required.

Proof of Proposition 4.3

For the proof, it suffices to bound the value $n|\overline{\sigma}_k^2 - \sigma_k^2|$. The definition of δ_n implies for every indices s, t that $|(A_k\overline{V} - A_k V)_{st}| \le \delta_n \sigma_k/(n-k)$. Therefore,

$$
\begin{aligned}
n|\overline{\sigma}_k^2 - \sigma_k^2| &= 2n\left|\mathrm{tr}[A_k\overline{V}]^2 - \mathrm{tr}[A_k V]^2\right| \\
&= 2n\left|\sum_{s,t=0}^{n-1}\left\{(A_k\overline{V} - A_k V)_{st}\right\}\left\{(A_k\overline{V} + A_k V)_{st}\right\}\right| \\
&\le \frac{2n\delta_n\sigma_k}{n-k}\sum_{s,t=0}^{n-1}\left|(A_k\overline{V} + A_k V)_{st}\right| \\
&\le \frac{2n\delta_n\sigma_k}{(n-k)^2}\sum_{s=0}^{n-1}\sum_{t\in\mathbb{Z}}\left(|\overline{r}_t| + r_t^*\right) \le \frac{2n^2K_1\delta_n\sigma_k}{(n-k)^2}.
\end{aligned}
$$

The desired result now follows in view of the bound $\sigma_k^2 \le v_0/(n-k)$ from Proposition 3.1.

Proof of Theorem 5.1

Let $A = \{|\widehat{r}_k - r_k| \le \varepsilon\}$ with ε from (4.3). By (2.6) $\boldsymbol{P}(A^c) \le m_n'/n$ where A^c is the complement of A. By Proposition 4.1 it holds on A that $n\widetilde{\sigma}_k^2 \ge n\sigma_k^2 - \psi_n$ and therefore, in view of $\sigma_k^2 \ge s_0^2/(n-k)$, see (2.4),

$$
(\lambda'\widetilde{\sigma}_k)^2 \ge (\lambda')^2\left(\sigma_k^2 - \psi_n/n\right) \ge (\lambda'\sigma_k)^2\left(1 - \psi_n(n-k)/(ns_0^2)\right) = (\lambda\sigma_k)^2 .
$$

Now, by Proposition 2.2

$$
\boldsymbol{P}\left(|\widehat{r}_k - r_k| \ge \lambda'\widetilde{\sigma}_k\right) \le \boldsymbol{P}(A^c) + \boldsymbol{P}\left(|\widehat{r}_k - r_k| \ge \lambda\sigma_k\right) \le m_n'/n + e^{-\lambda^2/4}
$$

as required.

Proof of Theorems 5.2

The proof is similar to that of Theorem 5.1 with the use of Proposition 2.3 in place of Proposition 2.2.

Proof of Theorems 5.3 (resp. Theorem 5.4)

Similarly to the proof of Theorem 5.1 one can show that the inequality $n\widetilde{\sigma}_k^2 \ge n\sigma_k^2 - \psi_n''$ and (3.5) imply $(\lambda'\widetilde{\sigma}_k)^2 \ge (\lambda\sigma_k)^2$. Now the result of Theorem 5.3 follows from Propositions 3.3, 3.5 and 4.3.

Theorem 5.4 can be proved in the same way using Proposition 3.4 in place of Proposition 3.3.

References

[1] Adak, S. (1998). Time-dependent spectral analysis of nonstationary time series, *J. American Statist. Ass. 93*, 1488-1501.

[2] Brockwell, P. and Davis, R. (1991). *Time Series: Theory and Methods*, Springer, New York.

[3] Dahlhaus, R. (1996a). On the Kullback-Leibler information divergence of locally stationary processes, *Stoch. Proc. Appl. 62*, 139-168.

[4] Dahlhaus, R. (1996b). Maximum likelihood estimation and model selection for locally stationary processes, *J. Nonparam. Statist. 6*, 171-191.

[5] Dahlhaus, R. (1997a). Fitting time series models to nonstationary processes, *Ann. Statist. 25*, 1-37.

[6] Dahlhaus, R. (1997b). The analysis of non-stationary time series and curve estimation with locally stationary models, Lecture Notes, *Academia Sinica*, Taiwan.

[7] Dahlhaus, R. (2000). A likelihood approximation for locally stationary processes, *Ann. Statist. 28* (6).

[8] Malat, S., Papanicolau, G. and Zhang, Z. (1998). Adaptive covariance estimation of locally stationary processes, *Annals of Statistics, 26*, 1-47.

[9] Neumann, M.H. and von Sachs, R. (1997). Wavelet tresholding in anisotropic function classes and application to adaptive estimation of evolutionary spectra, *Ann. Statist. 25*, 38-76.

[10] Ombao, H., Raz, J., von Sachs, R. and Guo, W. (2002). The SLEX Model of a non-stationary random process, *Ann. Inst. Statist. Math.*, **54**, no. 1, in print.

[11] Picard, D. (1985). Testing and estimating change-points in time series, *Adv. Appl. Probab. 17*, 841-67.

[12] Rozenholc, Y. (1995). Non-parametric tests of change point with tapered data, *Technical report PMA-553*, Université Paris Jussieu

[13] Priestley, M.B. (1965). Evolutionary spectra and nonstationary processes, *J. Roy. Statist. Soc. Ser. B 27*, 204-237.

[14] Sakiyama, K. and Taniguchi, M. (2000). Statistical analysis of locally stationary processes and its applications to testing for stationarity, *Preprint S-49*, Departmeant of Mathematical Science Osaka University.

[15] Spokoiny, V. (2002). Variance estimation for high-dimensional regression models. *J. of Multivariate Analysis*, in print.

[16] von Sachs, R. and Neumann, M.H. (2000). A wavelet-based test for stationarity, *J. Time Series Anal. 21*, 597-613.