

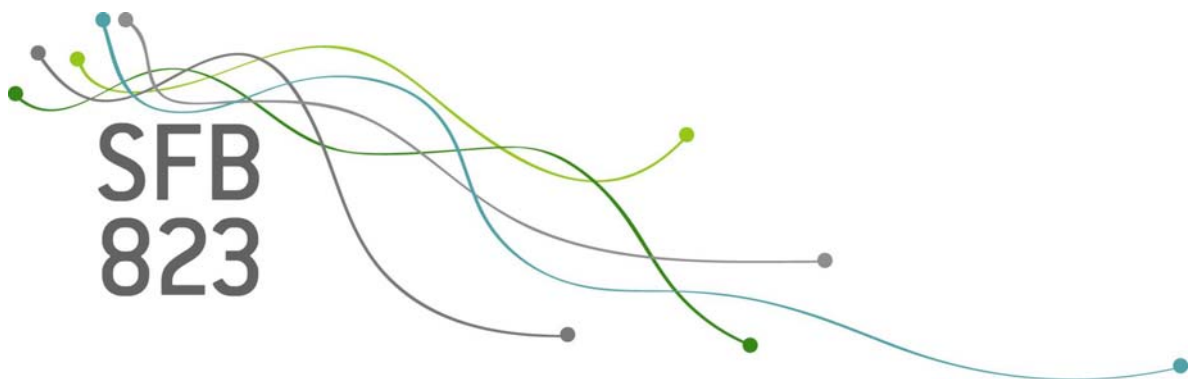
SFB
823

On local power properties of frequency domain based tests for stationarity

Efstathios Paparoditis, Philip Preuß

Nr. 13/2013

Discussion Paper



On local power properties of frequency domain based tests for stationarity

Efstathios Paparoditis¹ and Philip Preuß²

¹*Department of Mathematics and Statistics, University of Cyprus, P.O.Box 20537, CY 1678 Nicosia, CYPRUS*

²*Department of Mathematics, Ruhr Universität Bochum, 44780 Bochum, Germany*

email: stathisp@ucy.ac.cy

email: philip.preuss@ruhr-uni-bochum.de

April 7, 2013

Abstract

A framework for the asymptotic analysis of local power properties of tests of stationarity in time series analysis is developed. Appropriate sequences of locally stationary processes are defined that converge at a controlled rate to a limiting stationary process as the length of the time series increases. Different interesting classes of local alternatives to the null hypothesis of stationarity are then considered and the local power properties of some recently proposed, frequency domain-based tests for stationarity are investigated. Some simulations illustrate our theoretical findings.

1 Introduction

The assumption of second order stationarity has been the dominating paradigm in time series analysis for several decades. This assumption allowed for the development of an asymptotic theory capable of investigating properties of statistical inference procedures like estimation or testing. However, the assumption that the dependence characteristics of a stochastic process remains constant over time is often not justified in practice. Among many sources of non-stationarity, the covariance structure of a process may change as time evolves. In the last decades interest has been directed towards the statistical analysis of stochastic processes the second order characteristics of which vary over time. In an early work, Priestley (1965) introduced processes with evolutionary spectra and a time-varying spectral representation; see also Granger (1964) and Priestley (1988). For processes with time varying spectral characteristics, asymptotic statistical inference has been made possible by introducing the class of locally stationary processes; see Dahlhaus (1997). This class consists of triangular arrays of stochastic processes which together with a time rescaling approach allow for increasing information on the local structure of the process as the sample size increases to infinity. The associated fill-in type asymptotics, enables then the development of a powerful theory for investigating properties of statistical inference procedures for processes with

time varying spectral characteristics; see Dahlhaus and Polonik (2009) and Dahlhaus (2012) for an overview and for recent developments. Approaches similar to that of locally stationary processes have been also proposed in the literature. We refer to Nason et al. (2000) and Ombao et al. (2005) for locally stationary wavelet processes and to Davis et al. (2006) for piecewise stationary processes. The above interest in the statistical analysis of processes with time varying spectral characteristics motivated also the development of procedures for testing stationarity and the appropriateness of the preassigned stationary process class. Priestly and Rao (1969) considered testing the homogeneity of a set of evolutionary spectra evaluated at different instants of time using an analysis of variance framework. For the change-point problem, Picard (1985) developed a test based on spectral distribution functions and Giraitis and Leipus (1992) generalized this approach to linear processes. More recently, von Sachs and Neumann (2000) proposed a test of stationarity based on empirical wavelet coefficients while Dwivedi and Subba Rao (2010) introduced a Portmanteau type test statistic for detecting non-stationarity. Interest has been also directed towards frequency domain-based tests of stationarity. Paparoditis (2009) and Paparoditis (2010) proposed testing stationarity by evaluating the L_2 distance between a non-parametrically estimated local spectral density obtained via smoothing the rescaled local periodogram and a spectral density estimator based on the entire time series. Dette et al. (2011) proposed a test that also exploits the L_2 -distance between local and global spectral density estimates, where smoothing of the local periodogram is replaced by integrating over frequencies. Finally, a test of stationarity based on the Kolmogorov-Smirnov type distance between estimates of the time localized spectral distribution function and its stationary counterpart has been proposed by Dahlhaus (2009) and by Preuß et al. (2012).

Although for some of the aforementioned tests, consistency against a broad class of alternatives has been established, identifying differences in the power behavior of the tests is more involved and requires a sophisticated asymptotic analysis. In this paper, a framework for such an analysis is developed by considering sequences of local alternatives to stationarity, that is by considering appropriately defined sequences of locally stationary processes, that converge at a controlled rate to a stationary process as the length of the time series increases to infinity. Within this framework, different classes of locally stationary alternatives to the null hypothesis of stationarity can be generated two of which are discussed in more detail in this paper. The first class of local alternatives possesses so-called global in time deviations from stationarity, while the second class, generates locally stationary alternatives to stationarity that are more localized in time. For both classes of local alternatives, the power behavior of some recently proposed, frequency domain-based tests of stationarity is investigated and the maximum rate is identified, at which convergence to the null of stationarity is allowed, such that the tests have power bounded away from the level and from unity. Our theoretical findings highlight several interesting features and differences in the power behavior of the tests considered. Notice that although we restrict our comparison to some recently proposed, frequency domain-based tests of stationarity, the theoretical framework developed in this paper can be also used to investigate the local power properties of other tests of stationarity as well.

The paper is organized as follows. Section 2 briefly reviews the frequency domain-based tests of stationarity compared in this paper. The different classes of locally stationary alternatives considered are introduced in Section 3. For these classes of local alternatives, the power properties

of the tests considered are investigated and our main results are presented. Section 4 presents some simulations that illustrate and support our theoretical findings, while Section 5 gives some concluding remarks. All technical proofs are deferred to the Appendix.

2 A brief review of the different tests

2.1 Testing second order stationarity

Consider a triangular array $\{\mathbf{X}_n, n \in \mathbb{N}\}$ of stochastic processes, where $\mathbf{X}_n = \{X_{t,n}, t = 1, 2, \dots, n\}$, and assume that $X_{t,n}$ is generated by

$$X_{t,n} = \sum_{l=-\infty}^{\infty} a_{t,n}(l) \varepsilon_{t-l} \quad (1)$$

where, for simplicity, we assume that $\{\varepsilon_t\}$ is a zero mean, Gaussian white noise process with variance 1. For the coefficients $a_{t,n}(l)$ it is assumed that functions $l(\cdot)$ exist with

$$\sup_{t,n} |a_{t,n}(j)| \leq K/l(j) \quad \text{and} \quad \sum_{j=-\infty}^{\infty} |j|l^{-1}(j) < \infty$$

being satisfied, where the constant K does not depend on n . Furthermore, functions $a(\cdot, j) : [0, 1] \rightarrow \mathbb{R}$ exist such that for some $K > 0$ independent of n ,

$$\sup_{u \in (0,1]} \left| \frac{\partial^2 a(u, j)}{\partial u^2} \right| \leq K/l(j) \quad \text{and} \quad \sup_{1 \leq t \leq n} |a_{t,n}(j) - a(t/n, j)| \leq K/(nl(j)).$$

Let $f(u, \lambda) = (2\pi)^{-1} |A(u, \lambda)|^2$, $A(u, \lambda) = \sum_{j=-\infty}^{\infty} a(u, j) \exp\{-ij\lambda\}$, be the local spectral density of the locally stationary process (1). Denote by $g(\lambda) = \int_0^1 f(u, \lambda) du$ the time averaged local spectral density at frequency λ and notice that $g(\cdot) \geq 0$ and $g(\cdot)$ is a positive-definite function. Thus, $g(\cdot)$ is the spectral density of a stationary process which can be understood as the spectral density of the stationary process that approximates best (in the L_2 -sense) the local spectral density $f(u, \lambda)$ of the underlying locally stationary process $\{\mathbf{X}_n, n \in \mathbb{N}\}$, see Paparoditis (2009) and Dette et al. (2011). The null and alternative hypothesis of interest can then be stated as follows.

$$H_0 : \quad f(u, \cdot) = g(\cdot), \quad \text{a.e.},$$

$$H_1 : f(u, \cdot) \neq g(\cdot) \quad \text{on a set } A \subseteq [0, 1] \text{ with positive Lebesgue measure.}$$

For testing the above null hypothesis, several frequency domain based approaches have been proposed which are briefly reviewed in the sequel.

2.2 L_2 tests based on smoothed normalized local periodograms

Paparoditis (2009) proposed testing the null hypothesis H_0 by comparing local spectral density estimators based on different segments of the observed time series with a (global) spectral density estimator based on the whole stretch of data. To be more specific, let $m_n < n$ be an even time

window length and denote by $I_{m_n}(u, \lambda)$ the local periodogram at frequency $\lambda \in [0, \pi]$ calculated using m_n observations around the time point $[un]$, that is

$$I_{m_n}(u, \lambda) = \frac{1}{2\pi m_n} \left| \sum_{t=1}^{m_n} X_{t+[un]-m_n/2, n} e^{-i\lambda t} \right|^2.$$

Observe that if the null hypothesis is true, then

$$E\left(\frac{I_{m_n}(u, \lambda_j)}{g(\lambda_j)} - 1\right) = O(m_n^{-1}),$$

uniformly in $\lambda_j = 2\pi j/m_n$, $j \in \{-[m_n - 1]/2, \dots, [m_n/2]\}$. Thus, a test of the null hypothesis of stationarity can be constructed by estimating non parametrically the above mean function for a set of equidistant time points u in the interval $(0, 1)$ and evaluating the L_2 distance of the estimator obtained to the null function. To elaborate on, let $\hat{g}(\lambda)$ be a nonparametric, kernel estimator of the stationary spectral density $g(\lambda)$, that is

$$\hat{g}(\lambda) = \frac{1}{n} \sum_j K_h(\lambda - \lambda_j) I_n(\lambda_j),$$

where $I_n(\lambda) = (2\pi n)^{-1} |\sum_{t=1}^n X_{t,n} \exp\{-it\lambda\}|^2$ is the periodogram calculated using the entire set of observations, $K_h(\cdot) = h^{-1}K(\cdot/h)$, K is a smoothing kernel and h a smoothing bandwidth. Let further u_s , $s = 1, 2, \dots, N$, be a number of $N = [n/m_n]$ equidistant time points in the time interval $[\delta_n, 1 - \delta_n]$ with $u_1 = 1 - u_N = \delta_n$, where $\delta_n = m_n/(2n)$. Then, the test statistic proposed is given by

$$T_n = \frac{1}{N} \sum_{s=1}^N \int_{-\pi}^{\pi} Q_n^2(u_s, \lambda) d\lambda, \quad (2)$$

where

$$Q_n(u, \lambda) = \frac{1}{m_n} \sum_j K_b(\lambda - \lambda_j) \left(\frac{I_{m_n}(u, \lambda_j)}{\hat{g}(\lambda_j)} - 1 \right).$$

Notice that $Q_n(u, \lambda)$ is a nonparametric estimator of the local mean function $E(I_{m_n}(u, \lambda)/g(\lambda) - 1)$ which is expected to be close to the zero function if H_0 is true and b is the local bandwidth used to estimate this function. Furthermore, T_n is a time average of the L_2 distances of the estimator $Q_n(u_s, \cdot)$, $s = 1, 2, \dots, n$, to the zero function. A modification of this test based on the supremum of $\int_{-\pi}^{\pi} Q_n^2(u, \lambda) d\lambda$ evaluated over all values of u using rolling time windows has been considered by Paparoditis (2010).

It has been shown in Paparoditis (2009) that if $h \sim n^{-\lambda}$, $\lambda \in (3/20, 1/3)$, $b \sim m_n^{-\lambda}$, $N = n/m_n$ and $m_n = n^\delta$, for $\delta \in (\delta_1, \delta_2)$ with $\delta_1 = \max\{1/(3 - \lambda), \lambda/(1 - \lambda)\}$ and $\delta_2 = \min\{(8\lambda - 1)/(1 - \lambda), (1 - 2\lambda)/(1 - \lambda)\}$, then

$$m_n \sqrt{Nb} T_n - \mu_n \xrightarrow{D} N(0, v^2), \quad (3)$$

where

$$\mu_n = \sqrt{Nb^{-1}} \int_{-\pi}^{\pi} K^2(x) dx + \sqrt{Nb} \frac{1}{4\pi} \int_{-2\pi}^{2\pi} (K \star K)(y) dy,$$

$$v^2 = 2\pi^{-2} \int_{-2\pi}^{2\pi} (K \star K)^2(y) dy$$

and $K \star K$ denotes the convolution of the kernel K . Notice that, due to the normalization of the local periodogram $I_n(u, \lambda)$ by the global spectral density estimator $\hat{g}(\lambda)$, the centering sequence μ_n and the limiting variance v^2 do not depend on any unknown parameters or characteristics of the underlying process. Based on (3), a level α test for stationarity can be obtained by rejecting the null hypothesis whenever

$$\frac{m_n \sqrt{Nb} T_n - \mu_n}{v} \geq u_{1-\alpha}, \quad (4)$$

where $u_{1-\alpha}$ denotes the $1 - \alpha$ quantile of the standard normal distribution.

2.3 L_2 tests based on integrated local periodograms

A different approach to test stationarity has been proposed by Dette et al. (2011). The authors consider the L_2 distance

$$\begin{aligned} D^2 &= \int_0^1 \int_{-\pi}^{\pi} (f(u, \lambda) - g(\lambda))^2 d\lambda du \\ &= \int_0^1 \int_{-\pi}^{\pi} f(u, \lambda)^2 d\lambda du - \int_{-\pi}^{\pi} g^2(\lambda) d\lambda, \end{aligned} \quad (5)$$

which differs essentially from the quantity (2) in that the non-normalized difference $f(u, \lambda) - g(\lambda)$ is estimated by integrating instead of smoothing. Normalizing in this context is done later on by dividing through the standard deviation of the corresponding estimator. Observing that $D^2 = 0$ under the null hypothesis, the testing approach proposed is based on the estimator

$$\hat{D}_n^2 = \pi \hat{F}_{1,n} - 2\pi \hat{F}_{2,n},$$

of D^2 , where

$$\hat{F}_{1,n} = \frac{1}{n} \sum_{s=1}^N \sum_{j=1}^{\lfloor \frac{m_n}{2} \rfloor} I_{m_n}(u_s, \lambda_j)^2$$

and

$$\hat{F}_{2,n} = \frac{1}{m_n} \sum_{j=1}^{\lfloor \frac{m_n}{2} \rfloor} \left(\frac{1}{N} \sum_{s=1}^N I_{m_n}(u_s, \lambda_j) \right)^2.$$

Concerning the grow rate of m_n it is assumed that $m_n = n^\delta$ for some $\delta \in (1/2, 3/4)$. Dette et al. (2011) showed that, as $n \rightarrow \infty$,

$$\sqrt{n} \left(\hat{D}_n^2 - D^2 + \frac{2\pi m_n}{n} \hat{F}_{1,n} \right) \xrightarrow{D} N(0, \tau^2) \quad (6)$$

holds true, where under the null hypothesis, the expression for the variance τ^2 reduces to $\tau_{H_0}^2 = 4\pi \int_{-\pi}^{\pi} g^4(\lambda) d\lambda$. Note that $2\pi m_n/n\hat{F}_{1,n}$ appearing in (6) is just an estimator of a bias term which occurs in the asymptotics of $\hat{F}_{2,n}$. Dette et al. (2011) proposed to use

$$\hat{\tau}_{H_0}^2 = \frac{4\pi^2}{6n} \sum_{s=1}^N \sum_{j=1}^{\lfloor \frac{m_n}{2} \rfloor} I_{m_n}^4(u_s, \lambda_j)$$

as an estimator of $\tau_{H_0}^2$ and obtained under the null hypothesis, that,

$$\sqrt{n} \left(\hat{D}_n^2 + \frac{2\pi m_n}{n} \hat{F}_{1,n} \right) / \hat{\tau}_{H_0} \xrightarrow{D} N(0, 1). \quad (7)$$

A level α test for stationarity is then given by rejecting the null hypothesis whenever

$$\sqrt{n} \frac{\hat{D}_n^2 + 2\pi m_n/n\hat{F}_{1,n}}{\hat{\tau}_{H_0}} \geq u_{1-\alpha}, \quad (8)$$

where, as before, $u_{1-\alpha}$ denotes the $1 - \alpha$ quantile of the standard normal distribution. Notice that compared to the testing approach proposed in Paparoditis (2009) and since this test is based on integrating local periodograms, it does not involve any kernel smoothing and avoids, therefore, the need of selecting smoothing parameters.

2.4 Kolmogorov-Smirnov type tests

A different approach to test second order stationarity was proposed by Dahlhaus (2009) and Preuß et al. (2012). The starting point for this test of stationarity is the Kolmogorov-Smirnov type distance

$$D_{KS} = \sup_{(v, \omega) \in [0, 1]^2} \left| \int_0^v \int_0^{\pi\omega} f(u, \lambda) d\lambda du - v \int_0^{\pi\omega} g(\lambda) d\lambda \right| \quad (9)$$

and the observation that $D_{KS} = 0$ under the null hypothesis. An estimator of D_{KS} is given by

$$\hat{D}_{KS,n} = \sup_{v, \omega \in [0, 1]} |\hat{D}_n(v, \omega)|,$$

where

$$\hat{D}_n(v, \omega) = \frac{2\pi}{n} \sum_{s=1}^{\lfloor vN \rfloor} \sum_{j=1}^{\lfloor \omega \frac{m_n}{2} \rfloor} I_{m_n}(u_s, \lambda_j) - \frac{\lfloor vN \rfloor}{N} \frac{2\pi}{n} \sum_{s=1}^N \sum_{j=1}^{\lfloor \omega \frac{m_n}{2} \rfloor} I_{m_n}(u_s, \lambda_j).$$

A similar Test based on a slightly different estimator of D_{KS} has been considered by Dahlhaus (2009). Preuß et al. (2012) showed that, on the space $l_\infty([0, 1]^2)$, the process $\hat{G}_n(v, \omega) = \sqrt{n}(\hat{D}(v, \omega) - \mathbb{E}(\hat{D}(v, \omega)))$ converges weakly to a Gaussian process $G(v, \omega)$ with mean zero and an appropriate covariance structure, which reduces under the null hypothesis to

$$\text{Cov}(G(v_1, \omega_1), G(v_2, \omega_2)) = 2\pi(\min(v_1, v_2) - v_1 v_2) \int_0^{\pi \min(\omega_1, \omega_2)} g^2(\lambda) d\lambda. \quad (10)$$

We denote with $G_{H_0}(v, \omega)$ the process which corresponds to $G(v, \omega)$ under H_0 , i.e., the Gaussian process with mean function zero and covariance structure as specified in (10). If we write $q_{1-\alpha}(G_0)$ for the $(1 - \alpha)$ quantile of

$$G_0 = \sup_{v, \omega \in [0,1]} |G_{H_0}(v, \omega)|, \quad (11)$$

then a decision rule is given by rejecting the null hypothesis of stationarity whenever

$$\hat{D}_{KS,n} \geq q_{1-\alpha}(G_0). \quad (12)$$

Notice that the quantiles of G_0 are not known and depend on characteristics of the underlying process. Preuß et al. (2012) showed that critical values of this test can be approximated sufficiently good using appropriate bootstrap methods.

3 Behavior for local alternatives

3.1 Preliminaries and Assumptions

As mentioned in the Introduction, our aim is to develop a framework that enables the investigation of local power properties of the different tests, that is their power behavior for sequences $\{\mathbf{X}_n, n \in \mathbb{N}\}$ of locally stationary processes, that “converge” to a (limiting) stationary process at some appropriate rate as n increases to infinity. For this we consider processes within the class described by (1) and which satisfy the following assumptions.

Assumption 1. *Let $\mathbf{X}_n = \{X_{t,n}, t = 1, 2, \dots, n\}_{n \in \mathbb{N}}$ be a triangular array of random variables, where*

$$X_{t,n} = \sum_{l=-\infty}^{\infty} a_n(t/n, l) \varepsilon_{t-l}, \quad (13)$$

the ε_t 's are independent, standard Gaussian random variables and the sequence of functions $a_n(\cdot, l) : [0, 1] \rightarrow \mathbb{R}$, are for every $n \in \mathbb{N}$ twice continuously differentiable and satisfy

$$\sup_{u \in [0,1]} \sum_l |l| |a_n(u, l)| \leq C, \quad \sup_{u \in [0,1]} \sum_l \left| \frac{\partial}{\partial u} |l| a_n(u, l) \right| \leq C, \quad \sup_{u \in [0,1]} \sum_l \left| \frac{\partial^2}{\partial u^2} a_n(u, l) \right| \leq C, \quad (14)$$

for some constant $C < \infty$ independent of n .

Depending on the particular specification of the sequence of coefficient functions $a_n(u, j)$, $n \in \mathbb{N}$, in (13), different types of local stationary alternatives can be generated. We focus in the following on two particular cases.

3.2 Global in time local alternatives to stationarity

Consider the case where the sequences of functions $a_n(\cdot, l)$ in (13) are given by

$$a_n(u, l) = a_0(l)(1 + c_n b(u, l)), \quad n \in \mathbb{N}, \quad (15)$$

where $c_n = n^{-\kappa}$ for some $\kappa > 0$ and $b(\cdot, l) : [0, 1] \rightarrow \mathbb{R}$ are twice continuously differentiable functions. We suppose that c_n and $b(u, l)$ are chosen in way which ensures validity of (14) and that there exists a $l_0 \in \mathbb{N}$ such that $u \mapsto b(u, l_0)$ is not constant. Notice that the process $\{\mathbf{X}_n\}$ with coefficient functions as in (15) possess a local spectral density $f_n(u, \lambda)$ satisfying

$$\begin{aligned}
f_n(u, \lambda) &= \frac{1}{2\pi} \sum_{l, m=-\infty}^{\infty} a_n(u, l) a_n(u, m) \exp(-i\lambda(l - m)) \\
&= \frac{1}{2\pi} \sum_{l, m=-\infty}^{\infty} a_0(l) a_0(m) \exp(-i\lambda(l - m)) \\
&\quad + \frac{c_n}{2\pi} \sum_{l, m=-\infty}^{\infty} a_0(l) a_0(m) (b(u, l) + b(u, m)) \exp(-i\lambda(l - m)) \\
&\quad + \frac{c_n^2}{2\pi} \sum_{l, m=-\infty}^{\infty} a_0(l) a_0(m) b(u, l) b(u, m) \exp(-i\lambda(l - m)) \\
&\sim f(\lambda) + c_n f_a(u, \lambda), \quad \text{as } n \rightarrow \infty,
\end{aligned} \tag{16}$$

where

$$f(\lambda) = \frac{1}{2\pi} \sum_{l, m=-\infty}^{\infty} a_0(l) a_0(m) \exp(-i\lambda(l - m)) \tag{17}$$

and

$$f_a(u, \lambda) = \frac{1}{2\pi} \sum_{l, m=-\infty}^{\infty} a_0(l) a_0(m) (b(u, l) + b(u, m)) \exp(-i\lambda(l - m)). \tag{18}$$

Roughly speaking, $f(\lambda)$ corresponds to the “stationary part” of the time varying spectral density $f_n(u, \lambda)$ while $f_a(u, \lambda)$ stands for the dominating “non stationary part” that describes the deviation from stationarity. Notice that the speed of convergence of the sequence of locally stationary processes with corresponding local spectral densities $f_n(u, \lambda)$, $n \in \mathbb{N}$, to a stationary process with spectral density $f(\lambda)$ is controlled by the sequence c_n , $n \in \mathbb{N}$. Our aim is to identify the maximal rate at which c_n can converge to zero, that is at which $f_n(u, \lambda)$ converges to $f(\lambda)$, but such that the tests considered have power bounded away from the level and from unity. Write

$$A = \int_0^1 \int_{-\pi}^{\pi} \left(f_a(u, \lambda) - \int_0^1 f_a(w, \lambda) dw \right)^2 d\lambda du > 0$$

for the L_2 distance between $f_a(u, \lambda)$ and its best approximation through a stationary spectral density $\int_0^1 f_a(w, \lambda) dw$ and

$$A^* = \int_0^1 \int_{-\pi}^{\pi} \left([f_a(u, \lambda) - \int_0^1 f_a(w, \lambda) dw] / f(\lambda) \right)^2 d\lambda du > 0,$$

for the distance between the corresponding normalized functions. The following results are then obtained for the different tests considered.

Theorem 3.1 *Suppose that Assumption 1 is true and that the sequence $a_n(u, l)$ satisfies (15). Assume further that $h \sim n^{-\lambda}$, $\lambda \in (3/20, 1/3)$, $b \sim m_n^{-\lambda}$, $N = n/m_n$ and $m_n = n^\delta$ for $\delta \in (1/4, \tilde{\delta})$, where $\tilde{\delta} = \min\{(8\lambda - 1)/(1 - \lambda), (1 - 2\lambda)/(1 - \lambda)\}$. Then, as $n \rightarrow \infty$,*

$$P(\text{decision rule (4) rejects } H_0) = \begin{cases} 1 & \text{if } \kappa < 1/4(1 + \delta(1 - \lambda)) \\ \Phi(A^*/v - u_{1-\alpha}) & \text{if } \kappa = 1/4(1 + \delta(1 - \lambda)) \\ \alpha & \text{if } \kappa > 1/4(1 + \delta(1 - \lambda)) \end{cases}$$

Theorem 3.2 *Suppose that Assumption 1 is true and that the sequence $a_n(u, l)$ satisfies (15). In addition we assume that $\delta \in (1/2, 3/4)$ holds. Then, as $n \rightarrow \infty$,*

$$P(\text{decision rule (8) rejects } H_0) = \begin{cases} 1 & \text{if } \kappa < 1/4 \\ \Phi(A/\tau_1 - u_{1-\alpha}) & \text{if } \kappa = 1/4 \\ \alpha & \text{if } \kappa > 1/4 \end{cases}$$

with $\tau_1 = 4\pi \int_{-\pi}^{\pi} f(\lambda)^4 d\lambda$.

Theorem 3.3 *Suppose the assumptions of Theorem 3.2 are satisfied. Then, as $n \rightarrow \infty$,*

$$P(\text{decision rule (12) rejects } H_0) = \begin{cases} 1 & \text{if } \kappa < 1/2 \\ P(\sup_{v, \omega} |G(v, \omega)| > q_{1-\alpha}(G_0)) & \text{if } \kappa = 1/2 \\ \alpha & \text{if } \kappa > 1/2 \end{cases}$$

where $q_{1-\alpha}(G_0)$ denotes the $(1 - \alpha)$ quantile of the process G_0 [cf. (11)] and (v, ω) is a Gaussian process on $[0, 1]^2$ with covariance structure given in (10) [with $g(\lambda)$ replaced by $f(\lambda)$] and mean

$$\mu(v, \omega) = \left(\int_0^v \int_0^{\pi\omega} f_a(u, \lambda) d\lambda du - v \int_0^{\pi\omega} \int_0^1 f_a(u, \lambda) dud\lambda \right). \quad (19)$$

According to Theorem 3.1, the rate c_n at which the sequence of local spectral densities $f_n(u, \lambda)$ is allowed to converge to the stationary spectral density $f(\lambda)$ in order for the test (4) to have a power bounded away from the level and from unity is affected differently by the averaging in the time domain and the smoothing in the frequency domain. In particular, increasing δ increases the rate c_n at which local deviations from the null of stationarity can be detected. On the other hand, smoothing in the frequency domain, which is controlled by the parameter λ , decreases c_n by the factor $\delta\lambda/4$. Interestingly enough, this last property is common for other L_2 -based tests using kernel smoothed estimators of the functions involved; cf. Bickel and Rosenblatt (1973) for the case of density testing and Härdle and Mammen (1993) for testing the regression function. Notice that under the assumptions made, it is always true that $1/4 < 1/4 + \delta(1 - \lambda)/4 < 1/2$. Now, if we compare this with the results of Theorem 3.2, it can be observed that the non-smoothed L_2 test (8) performs worse than the smoothed test for this class of local alternatives. More specifically, the rate at which this test detects deviations from stationarity is smaller than the corresponding rate of the

smoothed L_2 test. However, by Theorem 3.3, both testing approaches based on the L_2 distance, are outperformed by the Kolmogorov-Smirnov type test (12), which for the same class of global in time locally stationary alternatives, can detect deviations from the null with asymptotic power of one whenever the convergence rate of the deviations from the null is less than the parametric rate $n^{-1/2}$, that is if $\kappa < 1/2$.

Notice that for the class of local alternatives considered in this section and in contrast to test (4), the probability of rejecting the null hypothesis of the two other testing procedures given in (8) and (12) can not be improved by the choice of the time-window parameter δ . Furthermore, compared to Theorem 3.1 and Theorem 3.2, the case where $\kappa = 1/2$ in Theorem 3.3 is less clear in the sense that it is not obvious if the probability of rejecting the null hypothesis is smaller, equal or bigger than the level α . We conjecture, however, that this probability will be larger than the level α since both limiting Gaussian processes $G(v, \omega)$ and $G_{H_0}(v, \omega)$ have the same covariance structure and differ only with respect to their mean function (recall that $G_0 = \sup_{v, \omega} |G_{H_0}(v, \omega)|$). While it is true that under the null hypothesis $\mathbb{E}(G_{H_0}(v, \omega)) = 0$ for all $v, \omega \in [0, 1]$, under the alternative, the mean function $\mu(v, \omega)$ is strictly positive or negative on a Lebesgue set having positive measure. This will imply that the probability that test (12) rejects the null hypothesis will be larger than α .

3.3 Time localized local alternatives to stationarity

The local deviations from the null of stationarity investigated in the previous section can be thought as being of rather global nature, in the sense that they do not necessarily become more concentrated around any particular time regions or time points as the sample size increases. We, therefore, refer to these local alternatives as Pitman-type local alternatives. However, since departures from stationarity which are rather localized in time, also form a relevant class of alternatives, we consider in the following a class of locally stationary alternatives that possesses this kind of departures from stationarity. Toward this goal, we consider the case where the sequence of functions $a_n(\cdot, l)$ is given by

$$a_n(u, l) = a_0(l)(1 + c_n b((u - u_0)/\gamma_n, l)), \quad (20)$$

where $u_0 \in (0, 1)$ is some fixed time point, $c_n = n^{-\kappa}$ and $\gamma_n = n^{-\zeta}$ for some $\kappa, \zeta > 0$. Here $b(\cdot, l)$ is a function on the whole real line the absolute value of which is integrable, and we assume that (14) remains valid. In addition, we suppose the existence of an $l_0 \in \mathbb{N}$ such that $z \mapsto b(z, l_0)$ is not constant. Notice that, since $\gamma_n \rightarrow 0$ as $n \rightarrow \infty$, the deviations from the null hypothesis produced by this type of local alternatives, become more concentrated around the time point u_0 as n increases to infinity. In this sense, such deviations can be considered as being more localized in time compared to those generated by the Pitman-type alternatives discussed in Section 3.2. Now, similarly to the previously considered alternatives, we obtain

$$f_n(u, \lambda) \sim f(\lambda) + c_n f_{b, u_0, \gamma_n}(u, \lambda) \quad \text{as } n \rightarrow \infty, \quad (21)$$

where $f(\lambda)$ is given by (17) and

$$f_{b, u_0, \gamma_n}(u, \lambda) = \frac{1}{2\pi} \sum_{l, m=-\infty}^{\infty} a_0(l) a_0(m) \left(b((u_0 - u)/\gamma_n, l) + b((u_0 - u)/\gamma_n, m) \right) \exp(-i\lambda(l - m)).$$

We define

$$f_b(z, \lambda) = \frac{1}{2\pi} \sum_{l, m=-\infty}^{\infty} a_0(l)a_0(m) (b(z, l) + b(z, m)) \exp(-i\lambda(l - m))$$

and using the substitution $z = (u_0 - u)/\gamma_n$, we obtain that,

$$\int_0^1 f_n(u, \lambda) du \sim f(\lambda) + c_n \gamma_n \int_{(u_0-1)/\gamma_n}^{u_0/\gamma_n} f_b(z, \lambda) dz.$$

If we now assume that $z \mapsto f_b(z, \lambda)$ is integrable for every λ , this directly implies

$$\left| \int_0^1 f_n(u, \lambda) du - f(\lambda) \right| \leq K c_n \gamma_n \quad (22)$$

for some constant $K < \infty$. Thus, the best (in the L_2 sense) stationary approximation of $f_n(u, \lambda)$, that is $g_n(\lambda) = \int_0^1 f_n(u, \lambda) du$, converges to the stationary part $f(\lambda)$ at a rate which is completely described by the sum of κ and ζ . We set

$$B = \int_{-\infty}^{\infty} \int_{-\pi}^{\pi} f_b(z, \lambda)^2 d\lambda dz$$

and

$$B^* = \int_{-\infty}^{\infty} \int_{-\pi}^{\pi} (f_b(z, \lambda)/f(\lambda))^2 d\lambda dz$$

for the squared L_2 norm of the corresponding non-normalized and normalized functions. Then, for the tests based on the L_2 distance, the following results are obtained.

Theorem 3.4 *Under the assumptions of Theorem 3.1 and if the sequence $a_n(u, l)$, $n \in \mathbb{N}$, satisfies (20), then, as $n \rightarrow \infty$,*

$$P(\text{decision rule (4) rejects } H_0) = \begin{cases} 1 & \text{if } 2\kappa + \zeta < 1/2(1 + \delta(1 - \lambda)) \\ \Phi(B^*/v - u_{1-\alpha}) & \text{if } 2\kappa + \zeta = 1/2(1 + \delta(1 - \lambda)) \\ \alpha & \text{if } 2\kappa + \zeta > 1/2(1 + \delta(1 - \lambda)) \end{cases}$$

Theorem 3.5 *Suppose that the assumptions of Theorem 3.2 hold and the sequence $a_n(u, l)$, $n \in \mathbb{N}$, satisfies (20). Then, as $n \rightarrow \infty$,*

$$P(\text{decision rule (8) rejects } H_0) = \begin{cases} 1 & \text{if } 2\kappa + \zeta < 1/2 \\ \Phi(B/\tau_2 - u_{1-\alpha}) & \text{if } 2\kappa + \zeta = 1/2 \\ \alpha & \text{if } 2\kappa + \zeta > 1/2 \end{cases}$$

with $\tau_2 = 4\pi \int_{-\pi}^{\pi} f(\lambda)^4 d\lambda$.

According to Theorem 3.4, and for the class of local alternatives considered in this section, the test based on the L_2 distance of the smoothed and normalized local periodograms turns out to be quite powerful in detecting deviations from stationarity. In particular, and since $2\kappa + \zeta = 1/2(1 + \delta(1 - \lambda)) > 1/2$, for δ and λ in the allowed range of values, we get the interesting result that, this test has nontrivial power even if the sequence of local spectral densities $f_n(\cdot, \cdot)$ is allowed to converge to the stationary spectral density $f(\cdot)$ at a rate which is faster than the parametric rate $n^{-1/2}$.

Note that, in contrast to the Pitman-type alternatives, the power for the case $2\kappa + \zeta = 1/2$ does not depend on the variation of $f_b(u, \lambda)$ in u but only on its magnitude for both tests (4) and (8). To get an intuition for the reason behind this behavior, consider the testing procedure (8) and notice that

$$\begin{aligned} & \int_0^1 \int_{-\pi}^{\pi} \left(f_n(u, \lambda) - \int_0^1 f_n(w, \lambda) dw \right)^2 d\lambda du \\ & \sim c_n^2 \gamma_n \int_{-\infty}^{\infty} \int_{-\pi}^{\pi} f_b(z, \lambda)^2 d\lambda dz - \int_{-\pi}^{\pi} \left(c_n \gamma_n \int_{-\infty}^{\infty} f_b(z, \lambda) dz \right)^2 d\lambda, \end{aligned}$$

which directly yields that it is the first summand of the right hand side of the last expression above which is the dominating one; see also the proof of Theorem 3.5 for more details. In this step it is essential that the absolute value of $z \mapsto f_b(z, \lambda)$ is integrable on the whole real line for every frequency λ , and, therefore, the case where $z \mapsto f_b(z, \lambda)$ equals some nonzero constant is not included in the above theorem.

The next theorem states the behavior of the Kolmogorov-Smirnov type test for the class of time localized local alternatives.

Theorem 3.6 *Suppose that the assumptions of Theorem 3.5 are satisfied. Then, as $n \rightarrow \infty$,*

$$P(\text{decision rule (12) rejects } H_0) = \begin{cases} 1 & \text{if } \kappa + \zeta < 1/2 \\ \alpha & \text{if } \kappa + \zeta > 1/2. \end{cases}$$

If $\kappa + \zeta = 1/2$, we have

$$\sqrt{n} \left(\hat{D}_n(v, \omega) - \mathbb{E}(\hat{D}(v, \omega)) \right)_{v, \omega \in [0, 1]^2} \Rightarrow (G_{H_0}(v, \omega))_{v, \omega \in [0, 1]^2}$$

where G_{H_0} is the Gaussian process which was used in the definition of G_0 (see (11) with $g(\lambda)$ replaced by $f(\lambda)$) and

$$\sqrt{n} \mathbb{E}(\hat{D}_n(v, \omega)) \sim \left(\int_0^{\pi\omega} w_{u_0}(v, \lambda) d\lambda \right) = \mu_{u_0}(v, \omega) \quad \text{for every } v, \omega \text{ as } n \rightarrow \infty,$$

with

$$w_{u_0}(v, \lambda) = \begin{cases} -v \int_{-\infty}^{\infty} f_b(z, \lambda) dz & \text{if } v < u_0 \\ \int_0^{\infty} f_b(z, \lambda) dz - v \int_{-\infty}^{\infty} f_b(z, \lambda) dz & \text{if } v = u_0 \\ (1 - v) \int_{-\infty}^{\infty} f_b(z, \lambda) dz & \text{if } v > u_0. \end{cases}$$

By the same argumentation as in the discussion following Theorem 3.3, we conjecture that in the case $\kappa + \zeta = 1/2$ the test (12) has a power between α and one. However, if we compare Theorem 3.5 and Theorem 3.6, it is observed that, since $\kappa > 0$, the Kolmogorov-Smirnov type test performs always better than the test (8) in detecting deviations from the null for the class of time localized, local alternatives. If in addition $\kappa + \zeta > 1/2$, it follows from Theorem 3.4 and Theorem 3.6 that the smoothed L_2 test has an asymptotic power larger than α for choices of the parameters δ and λ in the allowed range of values (cf. also the discussion after Theorem 3.4) while in this case, the test (12) has no power at all.

4 Simulated Examples

4.1 Preliminaries

We provide in this section some simulation results which demonstrate that our theoretical findings can, in fact, be observed also in finite sample situations. In order to obtain a fair comparison between the three different testing approaches considered and to avoid problems associated with the quality of the asymptotic Gaussian approximations involved, we estimate the quantiles of the distribution of the test statistics under the null hypothesis by means of an AR-sieve bootstrap procedure; cf. Kreiss (1988). In our context, this bootstrap procedure works by generating pseudo-time series using an autoregressive process of order p fitted to the entire stretch of data and i.i.d. innovations obtained from the empirical distribution function of the centered residuals of the autoregressive fit. It is well known, that this bootstrap procedure works theoretically for a large class of stochastic processes by letting the autoregressive order $p = p(n)$ tending to infinity with the sample size n . For the problem of testing for stationarity considered in this paper, theoretical justification of such an AR-sieve bootstrap is given by the fact that for all three testing procedures considered, the limiting distribution of the corresponding test statistics under the null, depends only on the second order characteristics of the hypothesized stationary process, i.e., on $g(\lambda)$. These second order characteristics, however, are successfully mimicked by the AR-sieve bootstrap; see Kreiss et al. (2011) for more details. Preuß et al. (2012) gave a formal proof of this statement for the case of the test (8).

Now, in order to implement the testing approaches considered in this paper, we have to choose additionally to the window length m_n also the order p of the autoregressive process used in the bootstrap procedure. Furthermore, for the L_2 test based on kernel smoothed quantities, the bandwidths b and h should be chosen as well. Concerning the smoothing bandwidth h , we follow the recommendations of Paparoditis (2009) and choose this parameter as the minimizer of the objective function

$$CV(h) = \frac{1}{L_n} \sum_{j=1}^{L_n} \left\{ \log(\hat{g}^{-j}(\omega_j)) + \frac{I_n(\omega_j)}{\hat{g}^{-j}(\omega_j)} \right\},$$

where $\hat{g}^{-j}(\omega_j)$ is the leave-out- j version of the estimator $\hat{g}_h(\omega_j)$ given by $\hat{g}^{-j}(\omega_j) = n^{-1} \times \sum_{s \in L_{n,j}} K_h(\omega_j - \omega_s) I_n(\omega_s)$ and $L_{n,j} = \{s : -L_n \leq s \leq L_n \text{ and } j - s \neq \pm j \text{ mod } L_n\}$. For h chosen, the bandwidth

b is selected as $b = h(n/m_n)^{0.2}$. The autoregressive order p used in the simulations, is chosen by minimizing the AIC criterion (see Akaike (1973)), that is

$$\hat{p} = \operatorname{argmin}_p \frac{1}{n} \sum_{k=1}^{\lfloor n/2 \rfloor} \left(\log(f_{\hat{\theta}(p)}(\lambda_{k,n})) + \frac{I_T^X(\lambda_{k,n})}{f_{\hat{\theta}(p)}(\lambda_{k,T})} \right) + p/n$$

in the context of stationary processes using Whittle's likelihood function; cf. Whittle (1951). Rules for choosing the time window parameter m_n are less objective and for this we decided to set this parameter equal to $\lfloor n/8 \rfloor$ throughout this section; see also Preuß et al. (2012) for similar choices in a comprehensive simulation study.

4.2 Numerical Results

We simulated time series data stemming from the following two models

$$X_{t,n} = \sigma_n(t/n)Z_t, \tag{23}$$

and

$$X_{t,n} = a_n(t/n)X_{t-1,n} + Z_t, \tag{24}$$

where

$$\begin{aligned} \sigma_n(u) &= (0.5 + n^{-0.45} \times 1.5u), \\ a_n(u) &= 0.5n^{-0.05} \exp(-n^{0.5}(u - 0.5)^2) \sin(4\pi u) \end{aligned}$$

and Z_t is a sequence of independent, standard normal distributed random variables. While the first process lies in the class of global in time locally stationary alternatives considered in Section 3.2, the second process corresponds to the more localized alternatives investigated in Section 3.3. Notice that the choice of the sequences of locally stationary processes considered in the simulation study and the rate of their convergence to the corresponding stationary processes is selected in order to illustrate our theoretical findings and to demonstrate clearly the differences in the behavior of the tests considered. For all three tests considered, we calculated the empirical rejection frequencies at a 5% level based on 500 simulation runs where for each run, 200 bootstrap replications have been generated. The results obtained for six different sample sizes n are summarized in Figure 1 .

For the case of model (23), the non smoothed L_2 -based test has very small power, while the power of the smoothed version of the L_2 test, first increases up to a rejection frequency of 0.142 for $n = 512$ and then decreases. For the Kolmogorov-Smirnov type test, however, the rejection frequencies are strictly increasing as the sample size n increases. Notice that this is in line with the results of Theorem 3.3 according to which we will expect the power of this test to converge to unity for the case of model (23). Furthermore, the convergence of the time varying variance function $\sigma_n(u)$ to the constant value 0.5 seems to be too fast so that the power of the other two testing approaches approaches the level α as n increases.

The picture is, however, different if we consider the results obtained for the more time localized alternative described by model (24). In this case, the performance of both, the non-smoothed L_2

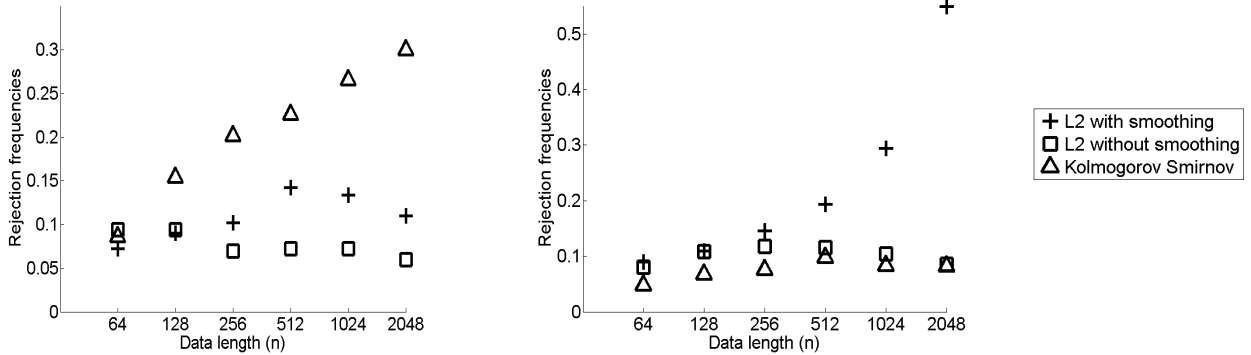


Figure 1: *Rejection frequencies of all three test at a 5% level for model (23) (left panel) and model (24) (right panel).*

based test and the Kolmogorov-Smirnov type test, is quite poor and the power of these tests seem to be close to the level α for large values of the sample size n . This pose behavior is due to the fast convergence rate of the particular locally stationary alternative. On the other hand, the smoothed L_2 test performs rather well in this case. In fact, as it can be seen form Figure 1, the rejection frequencies of this test increase steadily as the time series length n increases.

5 Conclusions

In this paper we have developed a framework for investigating local power properties of some recently proposed, frequency domain-based tests of second order stationarity in time series analysis. Two different classes of locally stationary, local alternatives to stationarity have been considered. The first one causes more global in time deviations from stationarity while the second one allows for more time localized deviations from the null. The tests considered, evaluate in a different manner differences between global and local spectral characteristics under the null and under the alternative. Some recently proposed, frequency-domain based tests for stationarity are considered which evaluate either the L_2 distance between local and global estimates of the spectral density or the Kolmogorov-Smirnov type distance between local and global estimates of the spectral distribution function. Our findings highlight several interesting features and differences between the tests considered.

Concerning the class of global in time deviations from stationarity, the Kolmogorov-Smirnov based test is more powerful and detects deviations from the null even if this type of local alternatives converge to the null at the parametric rate $n^{-1/2}$. For this class of local alternatives, tests based on the L_2 distance estimated by smoothing or by integrating local periodograms are less powerful although the test based on smoothing local periodograms turns out to be to be more powerful than the one based on integrating local periodograms. Things are different if one considers more time localized alternatives to stationarity. For this class of local alternatives, the tests based on the Kolmogorov-Smirnov or on the L_2 distance of non smoothed local periodograms, are less powerful

compared to the test based on the L_2 distance of smoothed estimators. In fact, this test turns out to be very powerful for this class of local alternatives and it is able to detect deviations from the null even if these local alternatives converge to the null at a rate which is faster than the parametric one.

Acknowledgements. This work has been supported in part by the Collaborative Research Center “Statistical modeling of nonlinear dynamic processes” (SFB 823, Teilprojekt C1) of the German Research Foundation (DFG) and by a University of Cyprus Research Grant. Parts of it were written while the second author was a Visiting Fellow at the London School of Economics which hospitality is gratefully acknowledged.

Appendix: Proofs

Proof of Theorem 3.1: As in the proof of Lemma 6.3 in Paparoditis (2009), it suffices to consider $m_n\sqrt{N}b\tilde{T}_n - \mu_n(K)$, where

$$\tilde{T}_n = N^{-1} \sum_{s=1}^N \int_{-\pi}^{\pi} \left(m_n^{-1} \sum_j K_b(\lambda - \lambda_j) (I_{m_n}(u_s, \lambda_j) / g_n(\lambda_j) - 1) \right)^2 d\lambda,$$

and $g_n(\lambda) = f(\lambda) + c_n \int_0^1 f_a(u, \lambda) du + O(c_n^2)$. We first fix some notation. Let $X_t^{(0)} = \sum_{j=-\infty}^{\infty} a_0(j) \varepsilon_{t-j}$, $Y_{t,n} = \sum_{j=-\infty}^{\infty} a_0(j) b(t/n, j) \varepsilon_{t-j}$, $d_{m_n}^{(0)}(u, \lambda) = (2\pi m_n)^{-1/2} \sum_{t=1}^{m_n} X_{t+[un]-m_n/2}^{(0)} \exp\{-i\lambda t\}$, $d_{m_n}^{(Y)}(u, \lambda) = (2\pi m_n)^{-1/2} \sum_{t=1}^{m_n} Y_{t+[un]-m_n/2,n} \exp\{-i\lambda t\}$, $I_{m_n}^{(0)}(u, \lambda) = |d_{m_n}^{(0)}(u, \lambda)|^2$ and $I_{m_n}^{(Y)}(u, \lambda) = |d_{m_n}^{(Y)}(u, \lambda)|^2$. We then have

$$\begin{aligned} m_n\sqrt{N}b\tilde{T}_n - \mu_n(K) &= m_n b^{1/2} N^{-1/2} \sum_{s=1}^n \int_{-\pi}^{\pi} \left(L_{0,n}(u_s, \lambda) + L_{1,n}(u_s, \lambda) + L_{2,n}(u_s, \lambda) \right. \\ &\quad \left. + L_{3,n}(u_s, \lambda) + O(c_n^2) \right)^2 d\lambda - \mu_n(K) \end{aligned}$$

where

$$L_{0,n}(u_s, \lambda) = m_n^{-1} \sum_j K_b(\lambda - \lambda_j) l_{n,j} [I_{m_n}^{(0)}(u_s, \lambda_j) / f(\lambda_j) - 1],$$

$$L_{1,n}(u_s, \lambda) = c_n m_n^{-1} \sum_j K_b(\lambda - \lambda_j) l_{n,j} [D_{m_n}(u_s, \lambda_j) - f_a(u_s, \lambda_j)] / f(\lambda_j),$$

$$L_{2,n}(u_s, \lambda) = c_n^2 m_n^{-1} \sum_j K_b(\lambda - \lambda_j) l_{n,j} I_{m_n}^{(Y)}(u_s, \lambda_j) / f(\lambda_j),$$

and

$$L_{3,n}(u_s, \lambda) = c_n m_n^{-1} \sum_j K_b(\lambda - \lambda_j) l_{n,j} [f_a(u_s, \lambda_j) - \int_0^1 f_a(w, \lambda_j) dw] / f(\lambda_j),$$

with $l_{n,j} = (1 + c_n \tilde{v}(\lambda_j) + O(c_n^2))^{-1}$, $\tilde{v}(\lambda_j) = \int_0^1 (f_a(u, \lambda_j) / f(\lambda_j)) du$ and

$$D_{m_n}(u, \lambda_j) = d_{m_n}^{(0)}(u, \lambda_j) d_{m_n}^{(Y)}(u, -\lambda_j) + d_{m_n}^{(0)}(u, -\lambda_j) d_{m_n}^{(Y)}(u, \lambda_j).$$

Now, for $c_n = m_n^{-1/2}b^{-1/4}N^{-1/4}$ it follows by straightforward calculations under the assumptions made, that,

$$m_n b^{1/2} N^{-1/2} \sum_{s=1}^N \int_{-\pi}^{\pi} \left(L_{1,n}(u_s, \lambda) + L_{2,n}(u_s, \lambda) \right)^2 d\lambda = O_P(c_n) + O_P(c_n^4 m_n (N b)^{1/2}) \rightarrow 0,$$

$$m_n b^{1/2} N^{-1/2} \sum_{s=1}^N \int_{-\pi}^{\pi} L_{0,n}(u_s, \lambda) \left(L_{1,n}(u_s, \lambda) + L_{2,n}(u_s, \lambda) \right) d\lambda = O_P(c_n N^{1/2} b^{-1/2}) + O_P(c_n^2 m_n b^{1/2}) \rightarrow 0,$$

$$m_n b^{1/2} N^{-1/2} \sum_{s=1}^N \int_{-\pi}^{\pi} L_{0,n}(u_s, \lambda) L_{3,n}(u_s, \lambda) d\lambda = O_P(m_n^{1/2} c_n) \rightarrow 0$$

and

$$m_n b^{1/2} N^{-1/2} \sum_{s=1}^N \int_{-\pi}^{\pi} L_{3,n}(u_s, \lambda) \left(L_{1,n}(u_s, \lambda) + L_{2,n}(u_s, \lambda) \right) d\lambda = O_P(c_n^2 m_n b N^{-1/2}) \rightarrow 0.$$

Thus

$$\begin{aligned} m_n \sqrt{N b} \tilde{T}_n - \mu_n(K) &= m_n b^{1/2} N^{-1/2} \sum_{s=1}^N \int_{-\pi}^{\pi} L_{0,n}^2(u_s, \lambda) d\lambda - \mu_n(K) \\ &\quad + m_n b^{1/2} N^{-1/2} \sum_{s=1}^N \int_{-\pi}^{\pi} L_{3,n}^2(u_s, \lambda) d\lambda + o_P(1), \end{aligned}$$

from which the assertion of the theorem follows since

$$m_n b^{1/2} N^{-1/2} \sum_{s=1}^N \int_{-\pi}^{\pi} L_{0,n}^2(u_s, \lambda) d\lambda - \mu_n(K) \xrightarrow{D} N(0, v^2),$$

as in Theorem 3.1 of Paparoditis (2009) and

$$m_n b^{1/2} N^{-1/2} \sum_{s=1}^N \int_{-\pi}^{\pi} L_{3,n}^2(u_s, \lambda) d\lambda \rightarrow \int_0^1 \int_{-\pi}^{\pi} \left(\frac{f_a(u, \lambda) - \int_0^1 f_a(w, \lambda) dw}{f(\lambda)} \right)^2 d\lambda du$$

□

Proof of Theorem 3.2: By using (14) we can proceed as in the proof of Theorem 2 in Dette et al. (2011) to obtain that the expression

$$\sqrt{n} \left(\hat{D}_n^2 + \frac{2\pi m_n}{n} \hat{F}_{1,n} - \int_0^1 \int_{-\pi}^{\pi} \left(f_n(u, \lambda) - \int_0^1 f_n(w, \lambda) dw \right)^2 d\lambda du \right) \quad (25)$$

converges to a normal distribution with mean zero and variance τ_1 . If we use the decomposition (16) we obtain

$$\sqrt{n} \int_0^1 \int_{-\pi}^{\pi} \left(f_n(u, \lambda) - \int_0^1 f_n(w, \lambda) dw \right)^2 d\lambda du \sim \sqrt{n} c_n^2 A$$

which directly yields the claim of Theorem 3.2. \square

Proof of Theorem 3.3: We proceed similarly to the above proof of Theorem 3.2. By combining (14) with the reasoning in the proof of Theorem 2.1 in Preuß et al. (2012) it follows that

$$\sqrt{n} \left(\hat{D}_n(v, \omega) - \mathbb{E}(\hat{D}_n(v, \omega)) \right) \quad (26)$$

converges [weakly in $l_\infty([0, 1]^2)$] to a zero mean Gaussian process with covariance structure given by (10) [note that here it is $g(\lambda) = f(\lambda)$]. Therefore the claim follows, as in the proof of Theorem 3.2, by calculating the expectation. By proceeding as in the proof of equation (5.3) in Preuß et al. (2012) we obtain

$$\sqrt{n} \mathbb{E}(\hat{D}_n(v, \omega)) = \sqrt{n} \left(\int_0^v \int_0^{\pi\omega} f_n(u, \lambda) d\lambda du - v \int_0^{\pi\omega} \int_0^1 f_n(u, \lambda) du d\lambda \right) + o(1) \quad (27)$$

which then yields the claim by employing (16) as in the proof of proof of Theorem 3.2. \square

Proof of Theorem 3.4: As in the proof of Theorem 3.1 it suffices to consider $m_n \sqrt{N} b \tilde{T}_n - \mu_n(K)$, where

$$\tilde{T}_n = N^{-1} \sum_{s=1}^N \int_{-\pi}^{\pi} \left(m_n^{-1} \sum_j K_b(\lambda - \lambda_j) (I_{m_n}(u_s, \lambda_j) / g_n(\lambda_j) - 1) \right)^2 d\lambda,$$

and $g_n(\lambda) = f(\lambda) + c_n \gamma_n \int_{(u_0-1)/\gamma_n}^{u_0/\gamma_n} f_b(u, \lambda) du + o(c_n \gamma_n)$. We adopt the notation of the proof of Theorem 3.1 with $Y_{t,n}$ replaced by $Y_{t,n} = c_n \sum_{j=-\infty}^{\infty} b((u - u_0)/\gamma_n, j) a_0(j) \varepsilon_{t-j}$. We then have

$$\begin{aligned} m_n \sqrt{N} b \tilde{T}_n - \mu_n(K) &= m_n b^{1/2} N^{-1/2} \sum_{s=1}^n \int_{-\pi}^{\pi} \left(L_{0,n}(u_s, \lambda) + L_{1,n}(u_s, \lambda) + L_{2,n}(u_s, \lambda) \right. \\ &\quad \left. + L_{3,n}(u_s, \lambda) + O(c_n^2) \right)^2 d\lambda - \mu_n(K) \end{aligned}$$

where $L_{0,n}(u_s, \lambda)$ and $L_{2,n}(u_s, \lambda)$ are defined as in the proof of Theorem 3.1 with $l_{n,j}$ replaced by $\tilde{l}_{n,j} = (1 + c_n \gamma_n \bar{v}(\lambda_j))^{-1}$, $\bar{v}(\lambda) = \int_{(u_0-1)/\gamma_n}^{u_0/\gamma_n} f_b(z, \lambda) dz / f(\lambda)$,

$$L_{1,n}(u_s, \lambda) = c_n m_n^{-1} \sum_j K_b(\lambda - \lambda_j) \tilde{l}_{n,j} [D_{m_n}(u_s, \lambda_j) - f_{b, u_0, \gamma_n}(u_s, \lambda_j)] / f(\lambda_j),$$

and

$$L_{3,n}(u_s, \lambda) = c_n m_n^{-1} \sum_j K_b(\lambda - \lambda_j) \tilde{l}_{n,j} [f_{b, u_0, \gamma_n}(u_s, \lambda_j) - c_n \gamma_n \int_{(u_0-1)/\gamma_n}^{u_0/\gamma_n} f_b(w, \lambda_j) dw] / f(\lambda_j).$$

Now, for $m_n b^{1/2} N^{1/2} c_n^2 \gamma_n \rightarrow 1$, it follows along the same lines as the proof of Theorem 3.1 that

$$\begin{aligned} m_n \sqrt{N} b \tilde{T}_n - \mu_n(K) &= m_n b^{1/2} N^{-1/2} \sum_{s=1}^n \int_{-\pi}^{\pi} L_{0,n}^2(u_s, \lambda) d\lambda - \mu_n(K) \\ &\quad + m_n b^{1/2} N^{-1/2} \sum_{s=1}^n \int_{-\pi}^{\pi} L_{3,n}^2(u_s, \lambda) d\lambda + o_P(1), \end{aligned}$$

from which the desired results follows because, as in Theorem 3.1 of Paparoditis (2009),

$$m_n b^{1/2} N^{-1/2} \sum_{s=1}^n \int_{-\pi}^{\pi} L_{0,n}^2(u_s, \lambda) d\lambda - \mu_n(K) \xrightarrow{D} N(0, v^2),$$

and, by straightforward calculations,

$$\begin{aligned} m_n \sqrt{b} N^{-1/2} \sum_{s=1}^n \int_{-\pi}^{\pi} L_{3,n}^2(u_s, \lambda) d\lambda &= m_n \sqrt{Nb} c_2^2 \gamma_n \int_{(u_0-1)/\gamma_n}^{u_0/\gamma_n} \int_{-\pi}^{\pi} f_b(z, \lambda) / f(\lambda)^2 d\lambda dz + o(1) \\ &\rightarrow \int_{-\infty}^{\infty} \int_{-\pi}^{\pi} (f_b(z, \lambda) / f(\lambda))^2 d\lambda dz, \end{aligned}$$

as $n \rightarrow \infty$. □

Proof of Theorem 3.5: As in the proof of Theorem 3.2 one validates asymptotic normality of (25) with the arguments from Dette et al. (2011), so we can again restrict ourselves to a calculation of the mean. The decomposition (21) yields

$$\begin{aligned} &\sqrt{n} \int_0^1 \int_{-\pi}^{\pi} \left(f_n(u, \lambda) - \int_0^1 f_n(w, \lambda) dw \right)^2 d\lambda du \\ &\sim \int_0^1 \int_{-\pi}^{\pi} c_n^2 f_{b,u_0,\gamma_n}(u, \lambda)^2 d\lambda du - \int_{-\pi}^{\pi} \left(c_n \int_0^1 f_{b,u_0,\gamma_n}(w, \lambda) dw \right)^2 d\lambda. \end{aligned}$$

By substituting $z = (u_0 - u) / \gamma_n$ we get that this term equals

$$\int_{(u_0-1)/\gamma_n}^{u_0/\gamma_n} \int_{-\pi}^{\pi} c_n^2 \gamma_n f_b(z, \lambda)^2 d\lambda dz - \int_{-\pi}^{\pi} \left(c_n \gamma_n \int_{(u_0-1)/\gamma_n}^{u_0/\gamma_n} f_b(z, \lambda)^2 dz \right)^2 d\lambda \sim c_n^2 \gamma_n B,$$

which then yields the claim of Theorem 3.5. □

Proof of Theorem 3.6: Again we only consider the expectation and obtain with the same substitution as in the above proof of Theorem 3.5 that

$$\begin{aligned} &\sqrt{n} \left(\int_0^v \int_0^{\pi\omega} f_n(u, \lambda) d\lambda du - v \int_0^{\pi\omega} \int_0^1 f_n(u, \lambda) dud\lambda \right) \\ &\sim \int_{(u_0-v)/\gamma_n}^{u_0/\gamma_n} \int_0^{\pi\omega} c_n \gamma_n f_b(z, \lambda) d\lambda dz - v \int_0^{\pi\omega} \int_{(u_0-1)/\gamma_n}^{u_0/\gamma_n} c_n \gamma_n f_b(z, \lambda) dz d\lambda. \end{aligned}$$

This concludes the proof with (27). □

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. *Second International Symposium on Information Theory (B.N. Petrov and F. Csaki, eds.)*, Budapest, Akademia Kiado, 267-281.

- Bickel, P. and Rosenblatt, M. (1973). On some global measures of the deviation of density function estimates. *Annals of Statistics*, 1, 1071-1095.
- Dahlhaus, R. (1997). Fitting time series models to non stationary processes. *Annals of Statistics*, 25, 1-37
- Dahlhaus, R. (2009). Local Inference for locally stationary time series based on the empirical spectral measure. *Journal of Econometrics*, 151, 101-112.
- Dahlhaus, R. (2012). Locally Stationary Processes. *Handbook of Statistics*, 30.
- Dahlhaus, R. and Polonik, W. (2009). Empirical spectral processes for locally stationary time series. *Bernoulli*, 15, 1-39.
- Davis, R. A., Lee, T. C. M. and Rodriguez-Yam, G. A. (2006). Structural breaks estimation for non stationary time series models. *Journal of the American Statistical Association*, 101, 223-239.
- Dette, H. and Preuß, P. and Vetter, M. (2011). A measure of stationarity in locally stationary processes with applications to testing. *Journal of the American Statistical Association*, 106(495), 1113-1124.
- Dwivedi, J. and Subba Rao, S. (2010). A test for second order stationarity based on the discrete Fourier transform. *Journal of Time Series Analysis*, 32, 68-91.
- Granger, C. W. J. (1964). Spectral analysis of economic time series. In *Princeton Studies in Mathematical Economics I* (association with M. Hatanaka). NJ: Princeton Univ. Press.
- Giraitis, L. and Leipus, R. (1992). Testing and estimating in the change-point problem of the spectral function. *Lithuanian Mathematical Journal*, 32, 15-29.
- Härdle, W. and Mammen, E. (1993). Comparing nonparametric versus parametric regression fits. *Annals of Statistics*, 21, 1926-1947.
- Kreiss, J.-P. (1988). Asymptotic statistical inference for a class of stochastic processes. *Habilitationschrift, Fachbereich Mathematik, Universität Hamburg*
- Kreiss, J.-P., Paparoditis, E. and D. N. Politis (2011). On the range of validity of the autoregressive sieve bootstrap. *Annals of Statistics*, 39(4), 2103-2130.
- Nason, G. P., von Sachs, R. and Kroisandt, G. (2000). Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum. *Journal of the Royal Statistical Society, Series B*, 62, 271-292.
- Ombao, H., von Sachs, R. and Guo, W. (2005). SLEX analysis of multivariate non stationary time series. *Journal of the American Statistical Association*, 100, 519-531.
- Paparoditis, E. (2009). Testing temporal constancy of the spectral structure of a time series. *Bernoulli*, 15, 1190–1221.

- Papadoditis, E. (2010). Validating stationarity assumptions in time series analysis by rolling local periodograms. *Journal of the American Statistical Association*, 105, 839-851.
- Picard, D. (1985). Testing and estimating change-points in time series. *Advances in Applied Probability*, 17, 841-867.
- Priestley, M. B. (1965). Evolutionary spectra and non-stationary processes. *Journal of the Royal Statistical Society, Series B*, 27, 204-237.
- Priestley, M. B. (1988). Non-linear and non-stationary time series analysis. *London: Academic Press*.
- Priestley, M. B. and Subba Rao, T. (1969). A test for non-stationarity of time series. *Journal of the Royal Statistical Society, Series B*, 31, 140-149.
- Preuß, P. and Vetter, M. and Dette, H. (2012). A test for stationarity based on empirical processes. *to appear in Bernoulli*.
- von Sachs, R. and Neumann, M. H. (2000). A wavelet-based test for stationarity. *Journal of Time Series Analysis*, 21, 597-613.
- Whittle, P. (1951). Hypothesis Testing in Time Series Analysis. *HUppsala: Almqvist and Wiksell*.
- Whittle, P. (1952). Some results in time series analysis. *HSkand. Aktuarietidskr.*, 35, 48-60.

