

Local inference for locally stationary time series based on the empirical spectral measure

Dahlhaus, Rainer

Postprint / Postprint

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

www.peerproject.eu

Empfohlene Zitierung / Suggested Citation:

Dahlhaus, R. (2009). Local inference for locally stationary time series based on the empirical spectral measure. *Journal of Econometrics*, 151(2), 101-112. <https://doi.org/10.1016/j.jeconom.2009.03.002>

Nutzungsbedingungen:

Dieser Text wird unter dem "PEER Licence Agreement zur Verfügung" gestellt. Nähere Auskünfte zum PEER-Projekt finden Sie hier: <http://www.peerproject.eu> Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under the "PEER Licence Agreement". For more information regarding the PEER-project see: <http://www.peerproject.eu> This document is solely intended for your personal, non-commercial use. All of the copies of this document must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Accepted Manuscript

Local inference for locally stationary time series based on the empirical spectral measure

Rainer Dahlhaus

PII: S0304-4076(09)00075-X
DOI: [10.1016/j.jeconom.2009.03.002](https://doi.org/10.1016/j.jeconom.2009.03.002)
Reference: ECONOM 3185

To appear in: *Journal of Econometrics*

Received date: 24 October 2008

Accepted date: 13 March 2009

Please cite this article as: Dahlhaus, R., Local inference for locally stationary time series based on the empirical spectral measure. *Journal of Econometrics* (2009), doi:10.1016/j.jeconom.2009.03.002

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Local Inference for Locally Stationary Time Series Based on the Empirical Spectral Measure

Rainer Dahlhaus, Universität Heidelberg

October 24, 2008

Abstract

The time varying empirical spectral measure plays a major role in the treatment of inference problems for locally stationary processes. The properties of the empirical spectral measure and related statistics are studied - both when its index function is fixed or dependent on the sample size. In particular we prove a general central limit theorem. Several applications and examples are given including semiparametric Whittle estimation, local least squares estimation and spectral density estimation.

Keywords: Empirical spectral measure, asymptotic normality, locally stationary processes, non-stationary time series.

JEL-Code: C220,C140.

Address of the author: Rainer Dahlhaus, Institute of Applied Mathematics, University of Heidelberg, Im Neuenheimer Feld 294, 69120 Heidelberg, Germany. PHONE: +49 6221 544989, FAX: +49 6221 545331, EMAIL: dahlhaus@statlab.uni-heidelberg.de.

1 Introduction

In recent years several methods have been derived for locally stationary time series models, that is for models which can locally be approximated by stationary processes. Out of the large literature we mention the work of Priestley (1965) on oscillatory processes, Robinson (1989) on time varying regression models, Dahlhaus (1997) on locally stationary processes, Neumann and von Sachs (1997) on wavelet estimation of evolutionary spectra, Nason, von Sachs and Kroisandt (2000) on a wavelet-based model of evolutionary spectra, and more recent work such as Davis, Lee and Rodriguez-Yam (2005) on piecewise stationary processes, Fryzlewicz, Sapatinas and Subba Rao (2006) on locally stationary volatility estimation, Moulines, Priouret and Roueff (2005) on recursive estimation for time varying autoregressive processes and Sakiyama and Taniguchi (2004) on discriminant analysis for locally stationary processes.

In this paper we emphasize the relevance of the empirical spectral measure for locally stationary time series. Here a generalized form of the empirical spectral measure is used which depends on arbitrary index functions that may even depend on the sample size n . The empirical spectral measure for fixed index function has been studied in Dahlhaus and Polonik (2008) where a functional limit theorem and an exponential inequality have been established. For Gaussian processes some (sometimes stronger) results are contained in Dahlhaus and Polonik (2006). In this paper we investigate the asymptotic properties of the empirical spectral measure with index functions depending on n and some related statistics.

In Section 2 we introduce the empirical spectral measure and the empirical spectral process indexed by functions and give a central limit theorem with fixed index functions. In Section 3 we prove the corresponding result with index functions ϕ_n depending on n . In Section 4 related estimates such as sum estimates and shifted estimates are studied.

The paper contains many applications and smaller examples. An overview is given in Example 2.3 where also additional references may be found.

All proofs are delegated without further reference to Section 5.

2 The time varying empirical spectral measure

For completeness and for a deeper understanding we first present in this section the results on the time varying empirical spectral measure with fixed index functions (c.f. Dahlhaus and Polonik, 2008). The presentation includes all necessary assumptions since we will use them in later sections.

Locally stationary processes were introduced in Dahlhaus (1997) by using a time varying spectral representation. In contrast to this we use in this paper a time varying MA(∞)-representation and formulate the assumptions in the time domain. The assumptions are much weaker than for example in Dahlhaus (1997) since the parameter curves are allowed to have jumps (see Remark 2.8 below). As in nonparametric regression we rescale the functions in time to the unit interval in order to achieve a meaningful asymptotic theory.

Let

$$V(g) = \sup_{0 \leq x_0 < \dots < x_m \leq 1, m \in \mathbf{N}} \sum_{k=1}^m |g(x_k) - g(x_{k-1})| \quad (1)$$

be the total variation of a function g on $[0, 1]$, and for some $\kappa > 0$ let

$$\ell(j) := \begin{cases} 1, & |j| \leq 1 \\ |j| \log^{1+\kappa} |j|, & |j| > 1. \end{cases} \quad (2)$$

Assumption 2.1 (Locally stationary processes) *The sequence of stochastic processes $X_{t,n}$ ($t = 1, \dots, n$) is called a locally stationary process if $X_{t,n}$ has a representation*

$$X_{t,n} = \sum_{j=-\infty}^{\infty} a_{t,n}(j) \varepsilon_{t-j} \quad (3)$$

satisfying the following conditions:

$$\sup_{t,n} |a_{t,n}(j)| \leq \frac{K}{\ell(j)} \quad (\text{with } K \text{ not depending on } n), \quad (4)$$

and there exist functions $a(\cdot, j) : (0, 1] \rightarrow \mathbf{R}$ with

$$\sup_u |a(u, j)| \leq \frac{K}{\ell(j)}, \quad (5)$$

$$\sup_j \sum_{t=1}^n |a_{t,n}(j) - a(\frac{t}{n}, j)| \leq K, \quad (6)$$

$$V(a(\cdot, j)) \leq \frac{K}{\ell(j)}. \quad (7)$$

The ε_t are assumed to be independent and identically distributed with $E\varepsilon_t \equiv 0$ and $E\varepsilon_t^2 \equiv 1$. In addition we assume in this paper that all moments of ε_t exist. We set $\kappa_4 := \text{cum}_4(\varepsilon_t)$.

The existence of all moments is assumed for the central limit theorems and the exponential inequality which both are proved via the method of cumulants. Of course it would be very worthwhile to relax this assumption.

Let

$$f(u, \lambda) := \frac{1}{2\pi} |A(u, \lambda)|^2$$

with

$$A(u, \lambda) := \sum_{j=-\infty}^{\infty} a(u, j) \exp(-i\lambda j)$$

is the time varying spectral density, and

$$c(u, k) := \int_{-\pi}^{\pi} f(u, \lambda) \exp(i\lambda k) d\lambda = \sum_{j=-\infty}^{\infty} a(u, k+j) a(u, j) \quad (8)$$

is the time varying covariance of lag k at rescaled time u . Usually $\text{cov}(X_{t,n} X_{t+k,n}) = c(\frac{t}{n}, k) + O(n^{-1})$ (see Remark 2.8 below).

A simple example of a process $X_{t,n}$ which fulfills the above assumptions is $X_{t,n} = \phi(\frac{t}{n}) Y_t$ where $Y_t = \sum_j a(j) \varepsilon_{t-j}$ is stationary with $|a(j)| \leq K/\ell(j)$ and ϕ is of bounded variation. Another example are time varying ARMA (tvARMA) processes whose coefficient functions are of bounded variation (see Dahlhaus and Polonik, 2008, Proposition 2.3). Since we frequently use the tvAR - case in our examples below we state it explicitly.

Example 2.2 (tvAR-processes) Consider the system of difference equations

$$\sum_{j=0}^p \alpha_j(\frac{t}{n}) X_{t-j,n} = \sigma(\frac{t}{n}) \varepsilon_t \quad (9)$$

where ε_t are iid with $E\varepsilon_t = 0$, $E|\varepsilon_t| < \infty$, $\alpha_0(u) \equiv 1$ and $\alpha_j(u) = \alpha_j(0)$ for $u < 0$. If all $\alpha_j(\cdot)$ as well as $\sigma^2(\cdot)$ are of bounded variation and $\sum_{j=0}^p \alpha_j(u) z^j \neq 0$ for all u and all $0 < |z| \leq 1 + \delta$ for some $\delta > 0$ then there exists a solution of the form

$$X_{t,n} = \sum_{j=0}^{\infty} a_{t,n}(j) \varepsilon_{t-j}$$

which fulfills (4)-(7) of Assumption 2.1. The time varying spectral density is given by

$$f(u, \lambda) = \frac{\sigma^2(u)}{2\pi} \frac{1}{|\sum_{j=0}^p \alpha_j(u) \exp(i\lambda j)|^2}$$

(see Dahlhaus and Polonik, 2008, Proposition 2.3). □

The empirical spectral process is defined by

$$E_n(\phi) := \sqrt{n} \left(F_n(\phi) - F(\phi) \right)$$

where

$$F(\phi) := \int_0^1 \int_{-\pi}^{\pi} \phi(u, \lambda) f(u, \lambda) d\lambda du \quad (10)$$

is the generalized spectral measure and

$$F_n(\phi) := \frac{1}{n} \sum_{t=1}^n \int_{-\pi}^{\pi} \phi\left(\frac{t}{n}, \lambda\right) J_n\left(\frac{t}{n}, \lambda\right) d\lambda \quad (11)$$

the empirical spectral measure with the pre-periodogram

$$J_n\left(\frac{t}{n}, \lambda\right) := \frac{1}{2\pi} \sum_{k:1 \leq [t+1/2 \pm k/2] \leq n} X_{[t+1/2+k/2],n} X_{[t+1/2-k/2],n} \exp(-i\lambda k). \quad (12)$$

If $X_{[t+1/2+k/2],n} X_{[t+1/2-k/2],n}$ is regarded as a (raw-) estimate of $c(\frac{t}{n}, k)$ then $J_n(\frac{t}{n}, \lambda)$ can be regarded as a (raw-) estimate of $f(\frac{t}{n}, \lambda)$ - however, in order to become consistent $J_n(\frac{t}{n}, \lambda)$ needs to be smoothed in time and frequency direction. The pre-periodogram J_n was first defined by Neumann and von Sachs (1997).

Example 2.3 Many statistics occurring in the analysis of locally stationary time series are of the form $F_n(\phi)$. We discuss several examples in this paper. Here we give an overview (K_n always denotes a kernel function).

- | | | | |
|----|--|------------------------------|---|
| 1. | $\phi(u, \lambda) = K_n(u_0 - u) \cos(\lambda k)$ | local covariance estimator | Example 3.5 |
| 2. | $\phi(u, \lambda) = K_n(u_0 - u) K_n(\lambda_0 - \lambda)$ | spectral density estimator | Example 4.2 |
| 3. | $\phi(u, \lambda) = K_n(u_0 - u) \nabla f_\theta(u, \lambda)^{-1}$ | local Whittle estimator | Example 3.6 |
| 4. | $\phi(u, \lambda) \approx K_n(u_0 - u) \nabla f_\theta(u, \lambda)^{-1}$ | local least squares | Example 4.4 |
| 5. | $\phi(u, \lambda) = \nabla f_\theta(u, \lambda)^{-1}$ | parametric Whittle estimator | Example 3.7
in Dahlhaus and Polonik (2008) |
| 6. | $\phi(u, \lambda) = (I_{[0, u_0]}(u) - u_0) I_{[0, \lambda_0]}(\lambda)$ | testing stationarity | Example 2.7 |
| 7. | $\phi(u, \lambda) = \cos(\lambda k)$ | stationary covariance | Remark 2.5 |
| 8. | $\phi(u, \lambda) = \nabla f_\theta(\lambda)^{-1}$ | stationary Whittle estimator | Remark 2.5 |
| 9. | $\phi(u, \lambda) = K_n(\lambda_0 - \lambda)$ | stationary spectral density | Remark 2.5 |

More complex examples are nonparametric maximum likelihood estimation under shape restrictions (Dahlhaus and Polonik, 2006), model selection with a sieve estimator (Van Bellegem and Dahlhaus, 2006) and wavelet estimates (Dahlhaus and Neumann, 2001). Moreover $F_n(\phi)$ occurs with local polynomial fits (Kim, 2001; Jentsch, 2006) and several statistics suitable for goodness of fit testing (a goodness-of-fit test based on a different statistic can be found in Sergides and Paparoditis, 2007). These applications are quite involved.

We mention that examples 1. - 4. are examples with index functions ϕ_n depending on n . The asymptotic properties of $F_n(\phi_n)$ in this case is the main topic of this paper. \square

We now state a central limit theorem for $E_n(\phi)$ with index functions ϕ that do not vary with n . We use the assumption of bounded variation in both components of $\phi(u, \lambda)$. Besides the definition in (1) we need a definition in two dimensions. Let

$$V^2(\phi) = \sup \left\{ \sum_{j,k=1}^{\ell,m} |\phi(u_j, \lambda_k) - \phi(u_{j-1}, \lambda_k) - \phi(u_j, \lambda_{k-1}) + \phi(u_{j-1}, \lambda_{k-1})| : \right. \\ \left. 0 \leq u_0 < \dots < u_\ell \leq 1; -\pi \leq \lambda_0 < \dots < \lambda_m \leq \pi; \ell, m \in \mathbf{N} \right\}.$$

For simplicity we set

$$\|\phi\|_{\infty, V} := \sup_u V(\phi(u, \cdot)), \quad \|\phi\|_{V, \infty} := \sup_\lambda V(\phi(\cdot, \lambda)), \\ \|\phi\|_{V, V} := V^2(\phi) \quad \text{and} \quad \|\phi\|_{\infty, \infty} := \sup_{u, \lambda} |\phi(u, \lambda)|.$$

Theorem 2.4 *Suppose Assumption 2.1 holds and let ϕ_1, \dots, ϕ_d be functions with $\|\phi_j\|_{\infty, V}$, $\|\phi_j\|_{V, \infty}$, $\|\phi_j\|_{V, V}$ and $\|\phi_j\|_{\infty, \infty}$ being finite ($j = 1, \dots, d$). Then*

$$(E_n(\phi_j))_{j=1, \dots, d} \xrightarrow{\mathcal{D}} (E(\phi_j))_{j=1, \dots, d}$$

where $(E(\phi_j))_{j=1, \dots, d}$ is a Gaussian random vector with mean 0 and

$$\text{cov}(E(\phi_j), E(\phi_k)) = 2\pi \int_0^1 \int_{-\pi}^{\pi} \phi_j(u, \lambda) [\phi_k(u, \lambda) + \phi_k(u, -\lambda)] f^2(u, \lambda) d\lambda du \\ + \kappa_4 \int_0^1 \left(\int_{-\pi}^{\pi} \phi_j(u, \lambda_1) f(u, \lambda_1) d\lambda_1 \right) \left(\int_{-\pi}^{\pi} \phi_k(u, \lambda_2) f(u, \lambda_2) d\lambda_2 \right) du. \quad (13)$$

This result is proved in Dahlhaus and Polonik (2008), Theorem 2.5.

Remark 2.5 (Stationary processes/Model misspecification by stationary models)

The classical central limit theorem for the weighted periodogram in the stationary case can be obtained as a corollary: If $\phi(u, \lambda) = \tilde{\phi}(\lambda)$ is time-invariant then

$$F_n(\phi) = \int_{-\pi}^{\pi} \tilde{\phi}(\lambda) \frac{1}{n} \sum_{t=1}^n J_n\left(\frac{t}{n}, \lambda\right) d\lambda.$$

We have

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^n J_n\left(\frac{t}{n}, \lambda\right) &= \frac{1}{n} \sum_{t=1}^n \frac{1}{2\pi} \sum_{1 \leq [t+0.5+k/2], [t+0.5-k/2] \leq n} X_{[t+0.5+k/2], n} X_{[t+0.5-k/2], n} \exp(-i\lambda k) \\ &= \frac{1}{2\pi} \sum_{k=-(n-1)}^{n-1} \left(\frac{1}{n} \sum_{t=1}^{n-|k|} X_t X_{t+|k|} \right) \exp(-i\lambda k) \\ &= \frac{1}{2\pi n} \left| \sum_{s=1}^n X_s \exp(-i\lambda s) \right|^2 = I_n(\lambda) \end{aligned} \quad (14)$$

where $I_n(\lambda)$ is the classical periodogram. Therefore $F_n(\phi)$ is the classical spectral measure in the stationary case with the following applications:

- (i) $\phi(u, \lambda) = \tilde{\phi}(\lambda) = \cos \lambda k$ is the empirical covariance estimator of lag k ;
- (ii) $\phi(u, \lambda) = \tilde{\phi}(\lambda) = \frac{1}{4\pi} \nabla f_{\theta}^{-1}(\lambda)$ is the score function of the Whittle-likelihood.

Theorem 2.4 gives the asymptotic distribution for these examples - both in the stationary case and in the misspecified case where the true underlying process is only locally stationary. If $\phi(u, \lambda) = \tilde{\phi}(\lambda)$ is a kernel we obtain an estimate of the spectral density whose asymptotic distribution is a special case of Theorem 3.2 below (also in the misspecified case). \square

We now briefly mention an exponential inequality. Since this is a non-asymptotic result it holds regardless whether ϕ depends on n . Let

$$\rho_{2,n}(\phi) := \left(\frac{1}{n} \sum_{t=1}^n \int_{-\pi}^{\pi} \phi\left(\frac{t}{n}, \lambda\right)^2 d\lambda \right)^{1/2}. \quad (15)$$

Theorem 2.6 (Exponential inequality) *Suppose Assumption 2.1 holds with $E|\varepsilon_t|^k \leq C_{\varepsilon}^k$ for all $k \in \mathbf{N}$ for the ε_t . Then we have for all $\eta > 0$*

$$P\left(|\sqrt{n}(F_n(\phi) - \mathbf{E}F_n(\phi))| \geq \eta\right) \leq c_1 \exp\left(-c_2 \sqrt{\frac{\eta}{\rho_{2,n}(\phi)}}\right) \quad (16)$$

with some constants $c_1, c_2 > 0$ independent of n .

This result is proved in Dahlhaus and Polonik (2008), Theorem 2.7. There exist several variants of this result - for example in the Gaussian case it is possible to omit the $\sqrt{\cdot}$ in (16) or to prove a Bernstein-type inequality which is even stronger (cf. Dahlhaus and Polonik, 2006, Theorem 4.1).

Furthermore, a maximal inequality may be proved under conditions on the metric entropy of the corresponding function class. We refer to Dahlhaus and Polonik (2008), Theorem 2.9. Applications of the maximal inequality are for example uniform rates of convergence for different type of estimates. We omit details.

By using a maximal inequality tightness of the empirical spectral process may be proved leading to functional central limit theorem for the empirical spectral process where the process is indexed by a function class (cf. Dahlhaus and Polonik, 2008, Theorem 2.11). Furthermore a Glivenko Cantelli type result for the empirical spectral process can be obtained (Theorem 2.12).

Example 2.7 (Testing for stationarity) The idea for a test of stationarity is to test whether the time varying spectral density $f(u, \lambda)$ is constant in u . This is for example achieved by the test statistic

$$\sqrt{n} \sup_{u \in [0,1]} \sup_{\lambda \in [0,\pi]} \left| F_n(u, \lambda) - u F_n(1, \lambda) \right|$$

where

$$F_n(u, \lambda) := \frac{1}{n} \sum_{t=1}^{[un]} \int_0^\lambda J_n\left(\frac{t}{n}, \mu\right) d\mu$$

is an estimate of the integrated time frequency spectral density $F(u, \lambda) := \int_0^u \int_0^\lambda f(v, \mu) d\mu dv$, and

$$u F_n(1, \lambda) = u \int_0^\lambda I_n(\mu) d\mu$$

is the corresponding estimate of $F(u, \lambda)$ under the hypothesis of stationarity where $f(v, \mu) = f(\mu)$. Under the hypothesis of stationarity we have

$$F(u, \lambda) - u F(1, \lambda) = \int_0^1 \int_0^\lambda (I_{[0,u]}(v) - u) f(\mu) d\mu dv = 0$$

and therefore

$$\sqrt{n} \left(F_n(u, \lambda) - u F_n(1, \lambda) \right) = E_n(\phi_{u,\lambda})$$

with $\phi_{u,\lambda}(v, \mu) = (I_{[0,u]}(v) - u) I_{[0,\lambda]}(\mu)$. We now need functional convergence of $E_n(\phi_{u,\lambda})$. Convergence of the finite dimensional distributions follows from Theorem 2.4 above. Tightness and therefore the functional convergence follows from Theorem 2.11 of Dahlhaus and

Polonik (2008) (it is straightforward to check the conditions of this theorem on the metric entropy).

As a consequence we obtain under the null hypothesis

$$\sqrt{n} \left(F_n(u, \lambda) - u F_n(1, \lambda) \right)_{u \in [0,1], \lambda \in [0,\pi]} \xrightarrow{\mathcal{D}} E(u, \lambda)_{u \in [0,1], \lambda \in [0,\pi]}$$

where $E(u, \lambda)$ is a Gaussian process with mean 0 and covariance structure

$$\begin{aligned} \text{cov}(E(u_1, \lambda_1), E(u_2, \lambda_2)) &= 2\pi \int_0^1 (I_{[0,u_1]}(v) - u_1) (I_{[0,u_2]}(v) - u_2) \int_0^{\min\{\lambda_1, \lambda_2\}} f^2(\mu) d\mu dv \\ &\quad + \kappa_4 \int_0^1 (I_{[0,u_1]}(v) - u_1) (I_{[0,u_2]}(v) - u_2) \left(\int_0^{\lambda_1} f(\mu_1) d\mu_1 \right) \left(\int_0^{\lambda_2} f(\mu_2) d\mu_2 \right) dv \\ &= \left(\min\{u_1, u_2\} - u_1 u_2 \right) \left[2\pi \int_0^{\min\{\lambda_1, \lambda_2\}} f^2(\mu) d\mu + \kappa_4 \left(\int_0^{\lambda_1} f(\mu_1) d\mu_1 \right) \left(\int_0^{\lambda_2} f(\mu_2) d\mu_2 \right) \right]. \end{aligned}$$

If $\kappa_4 = 0$ (Gaussian case) and $f(\mu) = c$ this is the Kiefer-Müller process. However, for general f it is a difficult and unsolved task to calculate or estimate the limit distribution and in particular the distribution in (2.7). This may be done by transformations (like U_p - or T_p -type transforms) and/or by finding an adequate bootstrap method.

We conjecture that the same result can be obtained if we use instead of $F_n(u, \lambda)$ the more realistic sum statistic

$$F_n^\Sigma(u, \lambda) := \frac{2\pi}{n^2} \sum_{t=1}^{[un]} \sum_{j=1}^{[\lambda n]} J_n\left(\frac{t}{n}, \lambda_j\right)$$

with $\lambda_j = \frac{2\pi j}{n}$. In this case convergence of finite dimensional distributions follows from Theorem 4.1 below while the derivation of tightness still is an open problem.

We mention that Paparoditis (2006) has given a complete solution of this testing problem by using spectral density type statistics. \square

Remark 2.8 (The bounded variation assumption) The assumption of bounded variation in time direction (7) of Assumption 2.1 allows for jumps in the parameter curves. We recall that a function is of bounded variation if it is the difference of two bounded monotone functions. This implies that the function may have countably many jumps with absolutely summable jump-sizes. In between the function is continuous. Assumption 2.1 implies bounded variation for many other curves of interest. For example it has been proved in Proposition 5.4 of Dahlhaus and Polonik (2008) that the covariances $c(u, k)$ are of bounded variation in u with

$$V(c(\cdot, k)) \leq \frac{K}{\ell(k)}.$$

It can also be shown that $f(u, \lambda)$ is of bounded variation in u for each λ . Furthermore, Proposition 5.4 of Dahlhaus and Polonik (2008) implies

$$\sum_{t=1}^n \left| \text{cov}(X_{t,n} X_{t+k,n}) - c\left(\frac{t}{n}, k\right) \right| \leq K. \quad (17)$$

This result also reflects bounded variation. Heuristically it is due to the following:

$$\text{cov}(X_{t,n} X_{t+k,n}) - c\left(\frac{t}{n}, k\right) = \begin{cases} O\left(\frac{1}{n}\right), & \text{if } c(u, k) \text{ is continuous in } \left[\frac{t}{n}, \frac{t+k}{n}\right]; \\ O(1), & \text{if } c(u, k) \text{ has a jump in } \left[\frac{t}{n}, \frac{t+k}{n}\right]. \end{cases}$$

More precisely, it can e.g. be shown that $\text{cov}(X_{t,n} X_{t+k,n}) - c\left(\frac{t}{n}, k\right) = O(1/n)$ uniformly in k if all a_j fulfill the Lipschitz-continuity assumption

$$|a_j(u) - a_j(v)| \leq \frac{K |u - v|}{\ell(j)} \quad \text{for all } u, v \in \left[\frac{t}{n}, \frac{t+k}{n}\right]$$

(cf. Assumption 2.1). The result (17) also holds for $\text{cov}(X_{[t+1/2+k/2],n}, X_{[t+1/2-k/2],n})$.

We do not take full advantage of the bounded variation assumptions in the present paper. The main results Theorems 2.4, 3.2, 4.1 and 4.3 are proved under these weak assumptions while nearly all examples deal with local kernel-type-estimates which require additional smoothness assumptions. Possible applications are estimation of jump-locations and jump-sizes, testing for breakpoints (in particular the behaviour of tests under the alternative) etc. These applications need to be investigated in future work. \square

3 A central limit theorem with index functions depending on n

We now prove a central limit theorem for $F_n(\phi_n) - F(\phi_n)$ with index functions ϕ_n depending on n . In addition we extend the definitions of Section 2 to tapered data

$$X_{t,n}^{(h_n)} := h_n\left(\frac{t}{n}\right) \cdot X_{t,n}$$

where $h_n : (0, 1] \rightarrow [0, \infty)$ is a data taper (with $h_n(\cdot) = I_{(0,1]}(\cdot)$ being the nontapered case of Section 2). The main reason for introducing data-tapers is to include segment estimates - see the discussion below.

As before the empirical spectral measure is defined by

$$F_n(\phi) = F_n^{(h_n)}(\phi) := \frac{1}{n} \sum_{t=1}^n \int_{-\pi}^{\pi} \phi\left(\frac{t}{n}, \lambda\right) J_n^{(h_n)}\left(\frac{t}{n}, \lambda\right) d\lambda \quad (18)$$

now with the tapered pre-periodogram

$$J_n^{(h_n)}\left(\frac{t}{n}, \lambda\right) = \frac{1}{2\pi} \sum_{k: 1 \leq [t+1/2 \pm k/2] \leq n} X_{[t+1/2+k/2], n}^{(h_n)} X_{[t+1/2-k/2], n}^{(h_n)} \exp(-i\lambda k) \quad (19)$$

(we mention that in some cases a rescaling may be necessary for $J_n^{(h_n)}(u, \lambda)$ to become a pre-estimate of $f(u, \lambda)$ - an obvious example for this is $h_n(u) = (1/2) I_{(0,1]}(u)$).

$F(\phi)$ is the theoretical counterpart of $F_n(\phi)$

$$F(\phi) = F^{(h_n)}(\phi) := \int_0^1 h_n^2(u) \int_{-\pi}^{\pi} \phi(u, \lambda) f(u, \lambda) d\lambda du. \quad (20)$$

We stress that for $h_n(u) = I_{(0,1]}(u)$ all definitions are the same as in Section 2. This is one important special case.

Another case of importance is $h_n^{(u_0)}(\frac{t}{n}) := k\left(\frac{u_0 - t/n}{b_n}\right)$ with k having compact support on $[-\frac{1}{2}, \frac{1}{2}]$ and $b_n \rightarrow 0$. If in addition $\phi(u, \lambda) = \psi(\lambda)$ we obtain in this case as in (14)

$$\begin{aligned} F_n(\phi) &= \int_{-\pi}^{\pi} \psi(\lambda) \left(\frac{1}{n} \sum_{t=1}^n J_n^{(h_n)}\left(\frac{t}{n}, \lambda\right) \right) d\lambda \\ &= \int_{-\pi}^{\pi} \psi(\lambda) \frac{1}{2\pi n} \left| \sum_{s=1}^n X_s^{(h_n)} \exp(-i\lambda s) \right|^2 d\lambda = \frac{H_{2,n}}{n} \int_{-\pi}^{\pi} \psi(\lambda) I_n^{(h_n)}(\lambda) d\lambda \end{aligned} \quad (21)$$

with the tapered periodogram

$$I_n^{(h_n)}(\lambda) := \frac{1}{2\pi H_{2,n}} \left| \sum_{s=1}^n X_s^{(h_n)} \exp(-i\lambda s) \right|^2 \quad (22)$$

and $H_{2,n} := \sum_{t=1}^n h_n(t/n)^2$ (which is in this example proportional to $b_n n$). This means that F_n is a weighted mean of the tapered periodogram on the segment $(u_0 - b_n/2)n \leq t \leq (u_0 + b_n/2)n$. For $k(\cdot) = I_{[-\frac{1}{2}, \frac{1}{2}]}(\cdot)$ we have the classical nontapered periodogram.

The last example suggests to use $\frac{1}{H_{2,n}}$ instead of $\frac{1}{n}$ in (18) as a norming constant. However, this is not always the right choice (as can be seen from case (ii) in Remark 3.4).

We need the following assumption.

Assumption 3.1 The data taper $h_n : (0, 1] \rightarrow [0, \infty)$ fulfills $\sup_n V(h_n) \leq C$ and $\sup_{u,n} h_n(u) \leq C$ for some $C < \infty$. $\log h_n(\cdot)$ is assumed to be concave.

The assumption that $\log h_n(\cdot)$ is concave is very mild (note that even $\log(x^m)$ is concave).

It turns out that in the above situation the rate of converge of the empirical spectral measure becomes $\sqrt{n}/\rho_2^{(h_n)}(\phi_n)$ where

$$\rho_2^{(h_n)}(\phi) := \left(\int_0^1 h_n^4(u) \int_{-\pi}^{\pi} \phi(u, \lambda)^2 d\lambda du \right)^{1/2}.$$

Therefore we can embed this case into the situation treated in the last section by studying the convergence of

$$E_n^{(h_n)} \left(\frac{\phi_n}{\rho_2^{(h_n)}(\phi_n)} \right) = \frac{\sqrt{n}}{\rho_2^{(h_n)}(\phi_n)} \left(F_n(\phi_n) - F^{(h_n)}(\phi_n) \right).$$

Furthermore, let

$$\begin{aligned} c_E^{(h_n)}(\phi_j, \phi_k) &:= 2\pi \int_0^1 h_n^4(u) \int_{-\pi}^{\pi} \phi_j(u, \lambda) [\phi_k(u, \lambda) + \phi_k(u, -\lambda)] f^2(u, \lambda) d\lambda du \\ &+ \kappa_4 \int_0^1 h_n^4(u) \left(\int_{-\pi}^{\pi} \phi_j(u, \lambda_1) f(u, \lambda_1) d\lambda_1 \right) \left(\int_{-\pi}^{\pi} \phi_k(u, \lambda_2) f(u, \lambda_2) d\lambda_2 \right) du. \end{aligned} \quad (23)$$

Theorem 3.2 *Suppose Assumptions 2.1 and 3.1 hold. Furthermore, let $\phi_{n1}, \dots, \phi_{nd}$ be functions with*

$$\frac{1 + \|\phi_n\|_{\infty, \infty} + (\log n) \|\phi_n\|_{\infty, V} + \|\phi_n\|_{V, \infty} + (\log n) \|\phi_n\|_{V, V}}{\sqrt{n} \rho_2^{(h_n)}(\phi_n)} = o(1). \quad (24)$$

If the limit

$$\Sigma_{j,k} := \lim_{n \rightarrow \infty} \frac{c_E^{(h_n)}(\phi_{nj}, \phi_{nk})}{\rho_2^{(h_n)}(\phi_{nj}) \rho_2^{(h_n)}(\phi_{nk})} \quad (25)$$

exists for all $j, k = 1, \dots, d$ then

$$\left(\frac{\sqrt{n}}{\rho_2^{(h_n)}(\phi_{nj})} \left(F_n(\phi_{nj}) - F^{(h_n)}(\phi_{nj}) \right) \right)_{j=1, \dots, d} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma). \quad (26)$$

Remark 3.3

(i) (Bias) In addition there exists the bias term

$$\frac{\sqrt{n}}{\rho_2^{(h_n)}(\phi_n)} \left(F^{(h_n)}(\phi_n) - \lim_{n \rightarrow \infty} F^{(h_n)}(\phi_n) \right)$$

(see the examples below). The magnitude of this bias depends on the smoothness of the time varying spectral density. In this paper we usually require conditions such that this bias is of lower order. Of course it is an important task to study the relation between bias and variance and to investigate in that context the problem of adaptation, e.g. the choice of the bandwidth parameters in the examples below. However, this is beyond the scope of the present paper.

(ii) (Fixed index functions) Theorem 2.4 for fixed index functions ϕ_j follows from the above result as a corollary.

(iii) (Stationary case) With $\phi_n(u, \lambda) = \tilde{\psi}_n(\lambda)$ the result also covers the stationary case - e.g. for kernel estimates of the spectral density (see (21)).

Remark 3.4 (Typical applications) A typical application of this result is the case of (kernel type) local estimators which can be constructed by using kernels, data-tapers or a combination of both:

$$\begin{aligned} \text{(i)} \quad \phi_n(u, \lambda) &= \frac{1}{b_n} K\left(\frac{u_0 - u}{b_n}\right) \psi(\lambda) & h_n(\cdot) &= I_{(0,1]}(\cdot) \\ \text{(ii)} \quad \phi_n(u, \lambda) &= \frac{1}{b_n} K\left(\frac{u_0 - u}{b_n}\right) \psi(\lambda) & h_n(u) &= I_{[u_0 - b_n/2, u_0 + b_n/2]}(u) \\ \text{(iii)} \quad \phi_n(u, \lambda) &= \psi(\lambda) & h_n\left(\frac{t}{n}\right) &= k\left(\frac{u_0 - t/n}{b_n}\right) \end{aligned}$$

where $K(\cdot)$ and $k(\cdot)$ are kernel functions and b_n is the bandwidth. Dependent on the function $\psi(\lambda)$ this leads to different applications (see the examples below). If $K(\cdot) = k(\cdot)^2$ then the resulting estimates all have the same asymptotic properties - see below. However, they are numerically different and their finite sample behaviour needs to be investigated. For example the last estimate is a stationary estimate on a segment which uses the classical tapered periodogram (or the nontapered periodogram for $k(\cdot) = I_{[-1/2, 1/2]}(\cdot)$).

We now show how Theorem 3.2 leads to the asymptotic distribution for these estimates:

(i) We assume that K is of bounded variation with compact support on $[-\frac{1}{2}, \frac{1}{2}]$ and $b_n \rightarrow 0$, $b_n n \rightarrow \infty$. $\psi(\lambda)$ is also of bounded variation. Since $V\left(\frac{1}{b_n} K\left(\frac{u_0 - \cdot}{b_n}\right)\right) = O\left(\frac{1}{b_n}\right)$ it is obvious that $\|\phi_n\|_{\infty, V}$, $\|\phi_n\|_{V, \infty}$, $\|\phi_n\|_{V, V}$ and $\|\phi_n\|_{\infty, \infty}$ are of order $O\left(\frac{1}{b_n}\right)$. Furthermore,

$$\rho_2^{(h_n)}(\phi_n) = \rho_2(\phi_n) = \left(\frac{1}{b_n} \int K^2(x) dx \int |\psi(\lambda)|^2 d\lambda \right)^{1/2} \quad (27)$$

that is (24) is fulfilled if $nb \gg (\log n)^2$. For $f(\cdot, \lambda)$ continuous at u_0 we have

$$\begin{aligned}
 c_E^{(h_n)}(\phi_{nj}, \phi_{nk}) &= 2\pi \int_0^1 \frac{1}{b_n^2} K\left(\frac{u_0 - u}{b_n}\right)^2 \int_{-\pi}^{\pi} \psi_j(\lambda) [\psi_k(\lambda) + \psi_k(-\lambda)] f^2(u, \lambda) d\lambda du \\
 &\quad + \kappa_4 \int_0^1 \frac{1}{b_n^2} K\left(\frac{u_0 - u}{b_n}\right)^2 \left(\int_{-\pi}^{\pi} \psi_j(\lambda_1) f(u, \lambda_1) d\lambda_1 \right) \left(\int_{-\pi}^{\pi} \psi_k(\lambda_2) f(u, \lambda_2) d\lambda_2 \right) du \\
 &\sim \frac{1}{b_n} \int K^2(x) dx \left[2\pi \int_{-\pi}^{\pi} \psi_j(\lambda) [\psi_k(\lambda) + \psi_k(-\lambda)] f^2(u_0, \lambda) d\lambda \right. \\
 &\quad \left. + \kappa_4 \left(\int_{-\pi}^{\pi} \psi_j(\lambda_1) f(u_0, \lambda_1) d\lambda_1 \right) \left(\int_{-\pi}^{\pi} \psi_k(\lambda_2) f(u_0, \lambda_2) d\lambda_2 \right) \right] =: \frac{1}{b_n} \Gamma_{jk}
 \end{aligned}$$

that is (25) is also fulfilled and we obtain from (26) the central limit theorem

$$\sqrt{b_n n} \left(F_n(\phi_{nj}) - F^{(h_n)}(\phi_{nj}) \right)_{j=1, \dots, d} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Gamma). \quad (28)$$

One may also allow for different $u_0 = u_j$ and $u_0 = u_k$. Since

$$\int \frac{1}{b_n} K\left(\frac{u_j - u}{b_n}\right) \frac{1}{b_n} K\left(\frac{u_k - u}{b_n}\right) du = o\left(\frac{1}{b_n}\right)$$

for $u_j \neq u_k$ we obtain by straightforward calculations that the corresponding covariance structure tends to 0.

(ii) The additional taper $h_n(u) = I_{[u_0 - b_n/2, u_0 + b_n/2]}(u)$ implies that we use only data from the interval $[u_0 - b_n/2, u_0 + b_n/2]$. We obtain in this case

$$\begin{aligned}
 \rho_2^{(h_n)}(\phi_n) &= \left(\int_0^1 h_n^4(u) \frac{1}{b_n^2} K\left(\frac{u_0 - u}{b_n}\right)^2 du \int_{-\pi}^{\pi} |\psi(\lambda)|^2 d\lambda \right)^{1/2} \\
 &= \left(\int_0^1 \frac{1}{b_n^2} K\left(\frac{u_0 - u}{b_n}\right)^2 du \int_{-\pi}^{\pi} |\psi(\lambda)|^2 d\lambda \right)^{1/2},
 \end{aligned}$$

i.e. we have the same $\rho_2^{(h_n)}(\phi_n)$ as above. Furthermore, $c_E^{(h_n)}(\phi_n, \phi_n)$ is the same. Thus we obtain the same asymptotic distribution and the same rate of convergence.

(iii) If $K(\cdot) = k(\cdot)^2$ we obtain in this case

$$\frac{1}{b_n} \rho_2^{(h_n)}(\phi_n) = \left(\int_0^1 \frac{1}{b_n^2} K\left(\frac{u_0 - u}{b_n}\right)^2 du \int_{-\pi}^{\pi} |\psi(\lambda)|^2 d\lambda \right)^{1/2}$$

i.e. we obtain again the same expression. Furthermore, $\frac{1}{b_n^2} c_E^{(h_n)}(\phi_{nj}, \phi_{nk})$ is the same as $c_E^{(h_n)}(\phi_{nj}, \phi_{nk})$ above. Thus we have again the same asymptotic distribution and the same rate of convergence.

Example 3.5 (Local covariance estimator)

If we set $\psi(\lambda) = \cos(\lambda k)$ in the last remark then estimate (i) becomes the kernel covariance estimator of $c(u_0, k)$

$$\begin{aligned}\widehat{c}_n(u_0, k) &:= F_n(\phi_{n,u_0,k}) = \frac{1}{n} \sum_{t=1}^n \frac{1}{b_n} K\left(\frac{u_0 - t/n}{b_n}\right) \int_{-\pi}^{\pi} \cos(\lambda k) J_n\left(\frac{t}{n}, \lambda\right) d\lambda \\ &= \frac{1}{n} \sum_t \frac{1}{b_n} K\left(\frac{u_0 - t/n}{b_n}\right) X_{[t+1/2+k/2],n} X_{[t+1/2-k/2],n}.\end{aligned}\quad (29)$$

Apart from the bias (which is treated below) (28) gives

$$\sqrt{b_n n} \left(\widehat{c}_n(u_k, k) - c(u_k, k) \right)_{k=1,\dots,d} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Gamma) \quad (30)$$

with

$$\Gamma_{k,\ell} = \delta_{u_k, u_\ell} \int K^2(x) dx \left[2\pi \int_{-\pi}^{\pi} \cos(\lambda k) \cos(\lambda \ell) f(u_k, \lambda)^2 d\lambda + \kappa_4 c(u_k, k) c(u_k, \ell) \right].$$

By using Parseval's formula and $\cos(\lambda k) = \frac{1}{2}[\exp(i\lambda k) + \exp(-i\lambda k)]$ the term in the [...] - brackets is the same as $\sum_{j=-\infty}^{\infty} [c(u_k, j) c(u_k, j - k - \ell) + c(u_k, j) c(u_k, j - k + \ell)] + \kappa_4 c(u_k, k) c(u_k, \ell)$ which is the usual asymptotic variance of the covariance-estimator (cf. Fuller (1996), Theorem 6.4.1).

We finally have to treat the bias. We have

$$F(\phi_{n,u_0,k}) - c(u_0, k) = \int_0^1 \int_{-\pi}^{\pi} \frac{1}{b_n} K\left(\frac{u_0 - u}{b_n}\right) (f(u, \lambda) - f(u_0, \lambda)) \cos(\lambda k) d\lambda du = O(b_n^2)$$

if $f(u, \lambda)$ is differentiable in u with Lipschitz continuous derivative in some neighborhood of u_0 and the Lipschitz constant is uniform in λ . We therefore obtain (30) if $\sqrt{b_n n} b_n^2 \rightarrow 0$ (i.e. for $b_n \ll n^{-1/5}$).

The same follows for the estimates (ii) and (iii). We mention that the estimate (iii) with $\psi(\lambda) = \cos(\lambda k)$ is just the Fourier transform of the classical tapered periodogram. \square

Example 3.6 (Curve estimation by local Whittle-estimates)

Whittle estimates were introduced by Whittle (1953) for stationary processes and were generalized by Dahlhaus (2000) to locally stationary processes. Parametric Whittle estimates for locally stationary processes have been discussed in Dahlhaus (2000) and in combination with the empirical spectral measure in Dahlhaus and Polonik (2008), Example 3.1.

Here we consider semiparametric curve estimation where the time varying spectral density is of the form $f(u, \lambda) = f_{\theta(u)}(\lambda)$ with $\theta(u) \in \Theta \subseteq \mathbf{R}^d$ for all $u \in [0, 1]$. An example are tvAR-processes. Let

$$\hat{\theta}_n(u) := \operatorname{argmin}_{\theta \in \Theta} \mathcal{L}_n(u, \theta)$$

with

$$\mathcal{L}_n(u, \theta) := \frac{1}{4\pi} \frac{1}{n} \sum_{t=1}^n \frac{1}{b_n} K\left(\frac{u-t/n}{b_n}\right) \int_{-\pi}^{\pi} \left\{ \log 4\pi^2 f_{\theta}(\lambda) + \frac{J_n(\frac{t}{n}, \lambda)}{f_{\theta}(\lambda)} \right\} d\lambda. \quad (31)$$

In case of a tvAR(p)-process $\hat{\theta}_n(u)$ is the solution of the local Yule-Walker equations: Let $C_n(u) = (\hat{c}_n(u, 1), \dots, \hat{c}_n(u, p))'$ and $\Sigma_n(u) = \{\hat{c}_n(u, i-j)\}_{i,j=1,\dots,p}$ with the covariance estimator $\hat{c}_n(u, k)$ as in (29). If $\hat{\theta}_n(u) = (\hat{\alpha}_1(u), \dots, \hat{\alpha}_p(u), \hat{\sigma}^2(u))'$ then it is not difficult to show that

$$(\hat{\alpha}_1(u), \dots, \hat{\alpha}_p(u))' = -\Sigma_n(u)^{-1} C_n(u)$$

and

$$\hat{\sigma}^2(u) = \hat{c}_n(u, 0) + \sum_{k=1}^p \hat{\alpha}_k(u) \hat{c}_n(u, k).$$

The basic steps for proving asymptotic normality of $\hat{\theta}_n(u)$ are well known. We mention that the empirical spectral measure plays a key role in the proof since the difference of the likelihood and its first and second order derivatives can be represented with the spectral measure. Large parts of the proof (in particular consistency of $\hat{\theta}_n(u)$) follow in the same way as in Dahlhaus and Polonik (2008), Theorem 4.1 where a uniform rate of convergence for $\|\hat{\theta}_n(u) - \theta_0(u)\|_2$ has been derived. The final step of the proof is to show asymptotic normality of the score function. We have with $\nabla_i := \frac{\partial}{\partial \theta_i}$

$$\begin{aligned} \sqrt{b_n n} \nabla_i \mathcal{L}_n(u, \theta_0(u)) &= \sqrt{b_n n} \left(F_n(\phi_{n,u,i}) - F(\phi_{n,u,i}) \right) + O\left(\frac{1}{\sqrt{b_n n}}\right) \\ &+ \sqrt{b_n n} \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \frac{1}{b_n} K\left(\frac{u-v}{b_n}\right) (f_{\theta_0(v)}(\lambda) - f_{\theta_0(u)}(\lambda)) \nabla_i f_{\theta}^{-1}(\lambda)|_{\theta=\theta_0(u)} d\lambda dv \end{aligned} \quad (32)$$

where $\phi_{n,u,i}(v, \lambda) := \frac{1}{b_n} K\left(\frac{u-v}{b_n}\right) \frac{1}{4\pi} \nabla_i f_{\theta}^{-1}(\lambda)|_{\theta=\theta_0(u)}$ (for tvAR(p)-processes this score function is simpler - cf. (50)). If $\theta_0(u)$ is differentiable with Lipschitz continuous derivative in some neighborhood of u and $f_{\theta}(\lambda)$ is differentiable in θ with uniform Lipschitz continuous derivative in some neighborhood of $\theta_0(u)$ then the last term is of lower order provided $b_n \ll n^{-1/5}$.

As in Remark 3.4 we therefore obtain from Theorem 3.2 (cf.(27))

$$\sqrt{b_n n} \nabla \mathcal{L}_n(u, \theta_0(u)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \int K^2(x) dx (\Gamma(u) + \Lambda(u))\right) \quad (33)$$

with

$$\Gamma(u)_{k,\ell} := \frac{1}{4\pi} \int_{-\pi}^{\pi} \left\{ \nabla_k \log f_{\theta}(\lambda) \Big|_{\theta=\theta_0(u)} \nabla_{\ell} \log f_{\theta}(\lambda) \Big|_{\theta=\theta_0(u)} \right\} d\lambda$$

and

$$\Lambda(u)_{k,\ell} := \kappa_4 \left(\frac{1}{4\pi} \int_{-\pi}^{\pi} \nabla_k \log f_{\theta}(\lambda) \Big|_{\theta=\theta_0(u)} d\lambda \right) \left(\frac{1}{4\pi} \int_{-\pi}^{\pi} \nabla_{\ell} \log f_{\theta}(\lambda) \Big|_{\theta=\theta_0(u)} d\lambda \right).$$

Furthermore, under appropriate smoothness assumptions on f and θ , it can be shown that

$$\nabla^2 \mathcal{L}_n(u, \bar{\theta}_n) \xrightarrow{P} \Gamma(u) \quad \text{for every } \bar{\theta}_n \text{ with } \bar{\theta}_n \xrightarrow{P} \theta_0(u) \quad (34)$$

which implies with a Taylor expansion and consistency of $\hat{\theta}_n(u)$

$$\sqrt{b_n n} \left(\hat{\theta}_n(u) - \theta_0(u) \right) \xrightarrow{D} \mathcal{N} \left(0, \int K^2(x) dx \Gamma(u)^{-1} (\Gamma(u) + \Lambda(u)) \Gamma(u)^{-1} \right). \quad (35)$$

Similarly, it can be shown that $\sqrt{b_n n} \left(\hat{\theta}_n(u) - \theta_0(u) \right)$ and $\sqrt{b_n n} \left(\hat{\theta}_n(v) - \theta_0(v) \right)$ are jointly asymptotically normal with zero limit covariance.

The above estimator corresponds to case (i) in Remark 3.4. Case (iii) in Remark 3.4 leads instead to the classical tapered Whittle likelihood on the segment, that is

$$\tilde{\theta}_n(u) := \operatorname{argmin}_{\theta \in \Theta} \tilde{\mathcal{L}}_n(u, \theta)$$

with

$$\tilde{\mathcal{L}}_n(u, \theta) := \frac{1}{4\pi} \int_{-\pi}^{\pi} \left\{ \log 4\pi^2 f_{\theta}(\lambda) + \frac{I_n^{(h_{u,n})}(\lambda)}{f_{\theta}(\lambda)} \right\} d\lambda.$$

where $h_{u,n}(\frac{t}{n}) := k(\frac{u-t/n}{b_n})$. This estimate has the same asymptotic properties provided $k(\cdot)^2 = K(\cdot)$. We omit details. \square

4 Related estimators

Many estimators are ‘‘approximately’’ of the form discussed in Sections 2 and 3. In this section we derive the asymptotic properties of such estimators rigorously and give examples beyond the examples already discussed in the last sections.

4.1 Sum estimators

We first consider the empirical spectral process with the integral statistic $F_n(\phi)$ being replaced by the sum statistic

$$F_n^\Sigma(\phi) := \frac{2\pi}{n^2} \sum_{t=1}^n \sum_{j=1}^n \phi\left(\frac{t}{n}, \lambda_j\right) J_n^{(h_n)}\left(\frac{t}{n}, \lambda_j\right) \quad (36)$$

where $\lambda_j := \frac{2\pi j}{n}$. This estimator has already been used in Example 2.7 (testing for stationarity). The empirical spectral process in this case is defined by

$$E_n^\Sigma(\phi) := \sqrt{n} (F_n^\Sigma(\phi) - F(\phi)). \quad (37)$$

Theorem 4.1 (i) *Suppose Assumptions 2.1 and 3.1 hold. Furthermore, let ϕ_n be a function possibly depending on n . Then*

$$\frac{n}{\rho_2^{(h_n)}(\phi_n)^2} \mathbf{E}[F_n^\Sigma(\phi_n) - F_n(\phi_n)]^2 \leq K \frac{(\log n) \|\phi_n\|_{\infty, V}^2}{n \rho_2^{(h_n)}(\phi_n)^2}. \quad (38)$$

(ii) (Fixed index functions) *Thus under the assumptions of Theorem 2.4 $\sqrt{n}(F_n^\Sigma(\phi) - F_n(\phi)) \xrightarrow{P} 0$ and as a consequence*

$$(E_n^\Sigma(\phi_j))_{j=1, \dots, d} \xrightarrow{\mathcal{D}} (E(\phi_j))_{j=1, \dots, d}$$

where $(E(\phi_j))_{j=1, \dots, d}$ has the same Gaussian distribution as in Theorem 2.4.

(iii) (Varying index functions) *Thus under the assumptions of Theorem 3.2 $(E_n^\Sigma(\phi_j / \rho_2^{(h_n)}(\phi_{nj})))_{j=1, \dots, d}$ has the same asymptotic distribution as given in Theorem 3.2.*

Example 4.2 (Spectral density estimates)

Consider a kernel estimate of the time-varying spectral density

$$\hat{f}_n(u, \lambda) := \frac{2\pi}{n^2} \sum_{t=1}^n \sum_{j=1}^n \frac{1}{b_t} K_t\left(\frac{u - t/n}{b_t}\right) \frac{1}{b_f} K_f\left(\frac{\lambda - \lambda_j}{b_f}\right) J_n^{(h_n)}\left(\frac{t}{n}, \lambda_j\right) \quad (39)$$

with different bandwidths and kernels in frequency and time direction. We have

$$\widehat{f}_n(u, \lambda) = F_n^\Sigma(\phi_{n,u,\lambda})$$

with $\phi_{n,u,\lambda}(v, \mu) := \frac{1}{b_t} K_t\left(\frac{u-v}{b_t}\right) \frac{1}{b_f} K_f\left(\frac{\lambda-\mu}{b_f}\right)$. Below we deduce from Theorem 4.1 that

$$\sqrt{nb_t b_f} \left(\widehat{f}_n(u_i, \lambda_i) - f(u_i, \lambda_i) \right)_{i=1, \dots, d} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Gamma) \quad (40)$$

with $\Gamma_{i,j} = 2\pi f(u_1, \lambda_1)^2 \delta_{u_i, u_j} \delta_{\lambda_i, \lambda_j} \int K_t^2(x) dx \int K_f^2(x) dx$ (and an additional factor 2 if $\lambda_i = 0$). The situation is the same as in Remark 3.4 but in addition with a kernel in frequency direction, i.e. $\psi(\mu) = \frac{1}{b_f} K_f\left(\frac{\lambda-\mu}{b_f}\right)$. In particular we are having again 3 different choices for constructing estimates (by using kernels and data-tapers). The 3 estimates are numerically different but with the same asymptotic distribution. Besides (39) we mention the ordinary smoothed periodogram

$$\widehat{f}_n(u, \lambda) := \frac{2\pi}{n} \sum_{j=1}^n \frac{1}{b_f} K_f\left(\frac{\lambda - \lambda_j}{b_f}\right) I_n^{(h_{u,n})}(\lambda) \quad (41)$$

with data-taper $h_n\left(\frac{t}{n}\right) = k\left(\frac{u-t/n}{b_t}\right)$, $k^2(\cdot) = K_t(\cdot)$ and $H_{2,n} = \sum_{t=1}^n h_n^2\left(\frac{t}{n}\right) \approx b_t n$.

We now derive (40) by checking the assumptions of Theorem 4.1. We assume that the kernels K_t, K_f have compact support on $[-\frac{1}{2}, \frac{1}{2}]$ and are of bounded variation with $\int_{-1/2}^{1/2} x K(x) dx = 0$ and $\int_{-1/2}^{1/2} K(x) dx = 1$. Furthermore let $b_t, b_f \rightarrow 0$ and $nb_t b_f \gg (\log n)^2$ as $n \rightarrow \infty$. The kernel K_f is periodically extended (i.e. $K_f\left(\frac{x}{b_f}\right) = K_f\left(\frac{x \pm 2\pi}{b_f}\right)$ due to periodicity of f) to include estimates for frequencies around $\pm\pi$.

It is obvious that $\|\phi_{n,u,\lambda}\|_{\infty, V}$, $\|\phi_{n,u,\lambda}\|_{V, \infty}$, $\|\phi_{n,u,\lambda}\|_{V, V}$ and $\|\phi_{n,u,\lambda}\|_{\infty, \infty}$ are of order $O\left(\frac{1}{b_t b_f}\right)$. Furthermore,

$$\rho_2(\phi_{n,u,\lambda})^2 = \frac{1}{b_t b_f} \int K_t^2(x) dx \int K_f^2(x) dx$$

that is (24) is fulfilled if $nb_t b_f \gg (\log n)^2$. We now check (25). By using

$$\int \frac{1}{b_t} K_t\left(\frac{u_1 - v}{b_t}\right) \frac{1}{b_t} K_t\left(\frac{u_2 - v}{b_t}\right) dv = o\left(\frac{1}{b_t}\right)$$

for $u_1 \neq u_2$ (and the same for K_f) we obtain for continuous $f(u, \lambda)$ by straightforward calculations in the case $\lambda_1 \neq 0$

$$c_E^{(I(0,1))}(\phi_{n,u_1,\lambda_1}, \phi_{n,u_2,\lambda_2}) = 2\pi f(u_1, \lambda_1)^2 \delta_{u_1, u_2} \delta_{\lambda_1, \lambda_2} \frac{1}{b_t b_f} \int K_t^2(x) dx \int K_f^2(x) dx + o\left(\frac{1}{b_t b_f}\right)$$

(the term with κ_4 is of lower order) and therefore

$$\lim_{n \rightarrow \infty} \frac{c_E^{(I_{(0,1)})}(\phi_{n,u_1,\lambda_1}, \phi_{n,u_2,\lambda_2})}{\rho_2(\phi_{n,u_1,\lambda_1}) \rho_2(\phi_{n,u_2,\lambda_2})} = 2\pi f(u_1, \lambda_1)^2 \delta_{u_1, u_2} \delta_{\lambda_1, \lambda_2}.$$

If $\lambda_1 = 0$ the same holds with a factor 2 on the right hand side. Thus we obtain

$$\sqrt{nb_t b_f} \left(\widehat{f}_n(u_i, \lambda_i) - F(\phi_{n,u_i,\lambda_i}) \right)_{i=1,\dots,d} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Gamma)$$

with Γ as above. We now treat the bias. We have

$$F(\phi_{n,u,\lambda}) - f(u, \lambda) = \int_0^1 \int_{-\pi}^{\pi} \frac{1}{b_t} K_t\left(\frac{u-v}{b_t}\right) \frac{1}{b_f} K_f\left(\frac{\lambda-\mu}{b_f}\right) (f(v, \mu) - f(u, \lambda)) d\mu du.$$

Suppose $f(u, \lambda)$ is twice differentiable in u and λ with uniformly bounded derivatives $\frac{\partial^2}{\partial u^2} f(u, \lambda)$, $\frac{\partial^2}{\partial u \partial \lambda} f(u, \lambda)$ and $\frac{\partial^2}{\partial \lambda^2} f(u, \lambda)$. Then a Taylor expansion yields $F(\phi_{n,u,\lambda}) - f(u, \lambda) = O_p(b_t^2 + b_f^2)$ and therefore the CLT in (40) holds if $\sqrt{nb_t b_f}(b_t^2 + b_f^2) \rightarrow 0$ (e.g. for $b_t = b_f \ll n^{-1/6}$). \square

4.2 Shifted estimates

By using Parseval's equality $F_n(\phi)$ can be written in the form

$$F_n(\phi) = \frac{1}{2\pi n} \sum_{t=1}^n \sum_k \widehat{\phi}\left(\frac{t}{n}, -k\right) X_{[t+1/2+k/2],n}^{(h_n)} X_{[t+1/2-k/2],n}^{(h_n)} \quad (42)$$

with

$$\widehat{\phi}(u, k) := \int_{-\pi}^{\pi} \phi(u, \lambda) \exp(i\lambda k) d\lambda. \quad (43)$$

There exist several quadratic forms which are stochastically equivalent to $F_n(\phi)$, for example the statistic

$$\frac{1}{2\pi n} \sum_{t=1}^n \sum_{k: 1 \leq t-k \leq n} \widehat{\phi}\left(\frac{t}{n}, -k\right) X_{t,n}^{(h_n)} X_{t-k,n}^{(h_n)}. \quad (44)$$

This is equal to

$$F_n^*(\phi) = \frac{1}{2\pi n} \sum_{t=1}^n \sum_k \widehat{\phi}_{t,-k} X_{[t+1/2+k/2],n}^{(h_n)} X_{[t+1/2-k/2],n}^{(h_n)}$$

with $\hat{\phi}_{t,-k} = \hat{\phi}(\frac{[t+1/2+k/2]}{n}, -k)$ which implies

$$\sum_t |\hat{\phi}_{t,k} - \hat{\phi}(\frac{t}{n}, k)| \leq K \min\{|k| + 1, n\} V(\hat{\phi}(\cdot, k)) \quad \text{and} \quad \sup_t |\hat{\phi}_{t,k}| \leq K \sup_u |\hat{\phi}(u, k)|.$$

By using Lemma 5.1 we therefore have

$$\sum_t |\hat{\phi}_{t,k} - \hat{\phi}(\frac{t}{n}, k)| \leq K \frac{\min\{|k|, n\}}{|k|} \|\phi\|_{V,V} \quad (k \neq 0), \quad \sum_t |\hat{\phi}_{t,0} - \hat{\phi}(\frac{t}{n}, 0)| \leq K \|\phi\|_{V,\infty}, \quad (45)$$

and

$$\sup_t |\hat{\phi}_{t,k}| \leq \frac{K}{|k|} \|\phi\|_{\infty,V} \quad (k \neq 0), \quad \sup_t |\hat{\phi}_{t,0}| \leq K \|\phi\|_{\infty,\infty}. \quad (46)$$

We will use these technical assumptions since they cover several other cases as well.

Let

$$E_n^*(\phi) := \sqrt{n}(F_n^*(\phi) - F(\phi)) \quad (47)$$

be the corresponding empirical spectral process.

Theorem 4.3 (i) *Suppose Assumptions 2.1 and 3.1 hold. Furthermore, let ϕ_n be a function possibly depending on n with Fourier-transform $\hat{\phi}_n(u, k)$, and $\hat{\phi}_{n,t,k}$ be complex numbers such that (45) and (46) hold (with ϕ replaced by ϕ_n and K independent of n) and*

$$F_n^*(\phi_n) := \frac{1}{2\pi n} \sum_{t=1}^n \sum_k \hat{\phi}_{n,t,-k} X_{[t+1/2+k/2],n}^{(h_n)} X_{[t+1/2-k/2],n}^{(h_n)}. \quad (48)$$

Then

$$\begin{aligned} & \frac{n}{\rho_2^{(h_n)}(\phi_n)^2} \mathbf{E}[F_n^*(\phi_n) - F_n(\phi_n)]^2 \\ & \leq K \frac{(1 + \|\phi_n\|_{\infty,\infty}) \|\phi_n\|_{V,\infty} + (1 + \log n \|\phi_n\|_{\infty,V}) \|\phi_n\|_{V,V}}{n \rho_2^{(h_n)}(\phi_n)^2}. \end{aligned}$$

(ii) (Fixed index functions) *Thus under the assumptions of Theorem 2.4 $\sqrt{n}(F_n^*(\phi) - F_n(\phi)) \xrightarrow{P} 0$ and as a consequence*

$$(E_n^*(\phi_j))_{j=1,\dots,d} \xrightarrow{\mathcal{D}} (E(\phi_j))_{j=1,\dots,d}$$

where $(E(\phi_j))_{j=1,\dots,d}$ has the same Gaussian distribution as in Theorem 2.4.

(iii) (Varying index functions) *Thus under the assumptions of Theorem 3.2*

$(E_n^(\phi_j / \rho_2^{(h_n)}(\phi_{nj}))_{j=1,\dots,d}$ has the same asymptotic distribution as given in Theorem 3.2.*

Example 4.4 (Local least squares estimates)

Suppose we want to fit a tvAR(p)-model

$$X_{t,n} + \sum_{j=1}^p \alpha_j \left(\frac{t}{n}\right) X_{t-j,n} = \sigma \left(\frac{t}{n}\right) \varepsilon_t$$

by the local least squares estimate

$$(\hat{\alpha}_1^*(u), \dots, \hat{\alpha}_p^*(u))' := \operatorname{argmin}_{\alpha_j} \frac{1}{n} \sum_{t=1}^n \frac{1}{b_n} K\left(\frac{u-t/n}{b_n}\right) \left(X_{t,n} + \sum_{j=1}^p \alpha_j X_{t-j,n}\right)^2$$

and

$$\hat{\sigma}^{*2}(u) := \frac{1}{n} \sum_{t=1}^n \frac{1}{b_n} K\left(\frac{u-t/n}{b_n}\right) \left(X_{t,n} + \sum_{j=1}^p \hat{\alpha}_j^*(u) X_{t-j,n}\right)^2 / \frac{1}{n} \sum_{t=1}^n \frac{1}{b_n} K\left(\frac{u-t/n}{b_n}\right).$$

Both equations can be summarized with $\theta(u) = (\alpha_1(u), \dots, \alpha_p(u), \sigma^2(u))'$ to

$$\hat{\theta}_n^*(u) := \operatorname{argmin}_{\theta} \mathcal{L}_n^*(u, \theta)$$

with

$$\mathcal{L}_n^*(u, \theta) = \frac{1}{n} \sum_{t=1}^n \frac{1}{b_n} K\left(\frac{u-t/n}{b_n}\right) \frac{1}{2} \left[\log \sigma^2 + \frac{1}{\sigma^2} \left(X_{t,n} + \sum_{j=1}^p \alpha_j X_{t-j,n}\right)^2 \right]. \quad (49)$$

As in Example 3.6 the asymptotic distribution is determined by the score function $\nabla \mathcal{L}_n^*(u, \theta_0(u))$. Below we show that this is of the form $F_n^*(\phi_n)$ with (45) and (46) while the score function of the local Whittle-estimator from Example 3.6 is the corresponding $F_n(\phi_n)$ from (42). We have in the special situation of tvAR(p)-processes due to Kolmogorov's equation (cf. (32))

$$\nabla_i \mathcal{L}_n(u, \theta_0(u)) = \frac{1}{n} \sum_{t=1}^n \frac{1}{b_n} K\left(\frac{u-t/n}{b_n}\right) \frac{1}{2} (\nabla_i \log \sigma^2)_{|\theta=\theta_0(u)} + F_n(\phi_{n,u,i}) \quad (50)$$

with $\phi_{n,u,i}(v, \lambda) := \frac{1}{b_n} K\left(\frac{u-v}{b_n}\right) \frac{1}{4\pi} \nabla_i f_{\theta}^{-1}(\lambda)_{|\theta=\theta_0(u)}$. The first term is the same as in $\nabla_i \mathcal{L}_n^*(u, \theta_0(u))$. We now calculate $F_n(\phi_{n,u,i})$. The time varying spectral density is given by $f(u, \lambda) = f_{\theta(u)}(\lambda)$ with

$$f_{\theta}(\lambda) = \frac{\sigma^2}{2\pi} \frac{1}{|1 + \sum_{j=1}^p \alpha_j \exp(i\lambda j)|^2}.$$

Therefore we have with $\alpha_0 = 1$, $\alpha_\ell = \alpha_{0,\ell}(u)$ ($\ell = 1, \dots, p$), $\alpha_\ell = 0$ ($\ell \notin \{0, \dots, p\}$) and $\sigma^2 = \sigma_0^2(u)$ for $i = 1, \dots, p$

$$\begin{aligned}\hat{\phi}_{n,u,i}(v, -k) &= \frac{1}{b_n} K\left(\frac{u-v}{b_n}\right) \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{\partial}{\partial \alpha_i} f_{\theta}^{-1}(\lambda)|_{\theta=\theta_0(u)} \exp(-i\lambda k) d\lambda \\ &= \frac{1}{b_n} K\left(\frac{u-v}{b_n}\right) \frac{1}{2\sigma^2} \int_{-\pi}^{\pi} \left(\sum_{j=0}^p \alpha_j [\exp(i\lambda(j-i)) + \exp(i\lambda(i-j))] \right) \exp(-i\lambda k) d\lambda \\ &= \frac{1}{b_n} K\left(\frac{u-v}{b_n}\right) \frac{\pi}{\sigma^2} [\alpha_{i-k} + \alpha_{i+k}].\end{aligned}$$

Thus for $i = 1, \dots, p$

$$\begin{aligned}\nabla_i \mathcal{L}_n(u, \theta_0(u)) &= F_n(\phi_{n,u,i}) = \frac{1}{2\pi n} \sum_t \sum_k \hat{\phi}_{n,u,i}\left(\frac{t}{n}, -k\right) X_{[t+1/2+k/2],n} X_{[t+1/2-k/2],n} \\ &= \frac{1}{2\pi n} \sum_t \sum_k \frac{1}{b_n} K\left(\frac{u-\frac{t}{n}}{b_n}\right) \frac{\pi}{\sigma^2} [\alpha_{i-k} + \alpha_{i+k}] X_{[t+1/2+k/2],n} X_{[t+1/2-k/2],n}.\end{aligned}$$

For the local least squares estimate we have for $i = 1, \dots, p$

$$\begin{aligned}\nabla_i \mathcal{L}_n^*(u, \theta_0(u)) &= \frac{1}{n} \sum_t \frac{1}{b_n} K\left(\frac{u-\frac{t}{n}}{b_n}\right) \frac{1}{\sigma^2} \left(\sum_{j=0}^p \alpha_j X_{t-j,n} \right) X_{t-i,n} \\ &= \frac{1}{n} \sum_t \frac{1}{b_n} K\left(\frac{u-\frac{t}{n}}{b_n}\right) \frac{1}{\sigma^2} \sum_k \alpha_{i+k} X_{t-i,n} X_{t-i-k,n} \\ &= \frac{1}{n} \sum_t \sum_k \frac{1}{b_n} K\left(\frac{u-\frac{[t+i+1/2+k/2]}{n}}{b_n}\right) \frac{1}{\sigma^2} \alpha_{i+k} X_{[t+1/2+k/2],n} X_{[t+1/2-k/2],n} \\ &= \frac{1}{2\pi n} \sum_t \sum_k \left[\frac{1}{b_n} K\left(\frac{u-\frac{[t+i+1/2-k/2]}{n}}{b_n}\right) \frac{\pi \alpha_{i-k}}{\sigma^2} + \right. \\ &\quad \left. + \frac{1}{b_n} K\left(\frac{u-\frac{[t+i+1/2+k/2]}{n}}{b_n}\right) \frac{\pi \alpha_{i+k}}{\sigma^2} \right] X_{[t+1/2+k/2],n} X_{[t+1/2-k/2],n} \\ &=: \frac{1}{2\pi n} \sum_{t=1}^n \sum_k \hat{\phi}_{n,u,i,t,-k} X_{[t+1/2+k/2],n} X_{[t+1/2-k/2],n} =: F_n^*(\phi_{n,u,i}).\end{aligned}$$

It is now straightforward to check (45) and (46). Thus Theorem 4.3 implies the equivalence of $F_n(\phi_{n,u,i})$ and $F_n^*(\phi_{n,u,i})$. In a similar way this follows for $F_n(\phi_{n,u,0})$ and $F_n^*(\phi_{n,u,0})$ (the score with respect to σ^2). As a consequence the convergence result (33) also holds for $\nabla \mathcal{L}_n^*(u, \theta_0(u))$. In addition convergence of the second derivative as in (34) needs to be proved for \mathcal{L}_n^* which again follows from the convergence of the empirical spectral measure. This implies that the convergence result (35) also holds for $\hat{\theta}_n^*(u)$. \square

Example 4.5 (Symmetrized estimator)

We have

$$J_n\left(\frac{t}{n}, \lambda\right) = \frac{1}{2\pi} \sum_{k \text{ even}} X_{t+k/2} X_{t-k/2} \exp(-i\lambda k) \\ + \frac{1}{2\pi} \sum_{k \text{ odd}} X_{t+(k+1)/2} X_{t-(k-1)/2} \exp(-i\lambda k).$$

Of course it is more natural to use a ‘symmetrized pre-periodogram’ with the second summand being

$$\frac{1}{2\pi} \sum_{k \text{ odd}} \frac{1}{2} [X_{t+(k+1)/2} X_{t-(k-1)/2} + X_{t+(k-1)/2} X_{t-(k+1)/2}] \exp(-i\lambda k)$$

instead. It is easy to see that the resulting estimator is of the form $F_n^*(\phi)$ with

$$\hat{\phi}_{t,k} = \begin{cases} \hat{\phi}\left(\frac{t}{n}, k\right), & k \text{ even} \\ \frac{1}{2} [\hat{\phi}\left(\frac{t}{n}, k\right) + \hat{\phi}\left(\frac{t+1}{n}, k\right)], & k \text{ odd,} \end{cases}$$

and it is easy to check assumptions (45) and (46) for this $\hat{\phi}_{t,k}$. Theorem 4.3 now implies that all results also hold with the symmetrized pre-periodogram.

5 Proofs

Lemma 5.1 *If $\|\phi\|_{\infty, V} < \infty$ then we have for $k \neq 0$*

$$\sup_u |\hat{\phi}(u, k)| \leq \frac{K}{|k|} \|\phi\|_{\infty, V}, \quad V(\hat{\phi}(\cdot, k)) \leq \frac{K}{|k|} \|\phi\|_{V, V}$$

and

$$\sup_u |\hat{\phi}(u, 0)| \leq 2\pi \|\phi\|_{\infty, \infty}, \quad V(\hat{\phi}(\cdot, 0)) \leq 2\pi \|\phi\|_{V, \infty}.$$

PROOF. We have for $k \neq 0$

$$\hat{\phi}(u, k) = \int_0^{2\pi} \frac{\exp(-ik\lambda) - 1}{ik} \phi_R(u, d\lambda)$$

where $\phi_R(u, d\lambda)$ is the signed measure corresponding to $\phi_R(u, \lambda) := \lim_{\mu \downarrow \lambda} \phi(u, \mu)$ (since ϕ is of bounded variation in λ the limit exists; for the same reason $\phi_R(u, d\lambda)$ is a signed measure). This implies $\sup_u |\hat{\phi}(u, k)| \leq \frac{K}{|k|} \sup_u V(\phi(u, \cdot)) = \frac{K}{|k|} \|\phi\|_{\infty, V}$. The other results follow similarly. \square

5.2 (Proof of Theorem 3.2)

The result is established via the method of cumulants. In particular we want to adapt cumulant calculations in Dahlhaus and Polonik (2008) where the cumulants have been calculated under stronger smoothness assumptions. The trick is to smooth $\phi_n(u, \lambda)$ in λ -direction and to prove asymptotic normality instead for the resulting sequence of approximations: Let $k(x) := \frac{1}{\sqrt{2\pi}} \exp\{-\frac{1}{2}x^2\}$ be the Gaussian kernel, $k_b(x) := \frac{1}{b} k(\frac{x}{b})$ and

$$\phi_n^*(u, \lambda) := \int_{-\infty}^{\infty} k_b(\lambda - \mu) \phi_n(u, \mu) d\mu$$

with $b = b_n \rightarrow 0$ as $n \rightarrow \infty$ (where $\phi_n(u, \mu) = 0$ for $|\mu| > \pi$). We have

$$\widehat{\phi}_n^*(u, k) = \widehat{\phi}_n(u, k) \widehat{k}_b(k) \quad (51)$$

with $\widehat{\phi}_n(u, k)$ being the Fourier transform as in (43) and $\widehat{k}_b(k) = \exp(-k^2 b^2 / 2)$. Thus we obtain from Lemma 5.7(i) in Dahlhaus and Polonik (2008) and Lemma 5.1

$$\begin{aligned} n \operatorname{var} \left[F_n \left(\frac{\phi_n^*}{\rho_2^{(h_n)}(\phi_n)} \right) - F_n \left(\frac{\phi_n}{\rho_2^{(h_n)}(\phi_n)} \right) \right] &\leq \frac{K}{n \rho_2^{(h_n)}(\phi_n)^2} \sum_{t=1}^n \sum_{k=-\infty}^{\infty} \left(\widehat{\phi}_n^* \left(\frac{t}{n}, k \right) - \widehat{\phi}_n \left(\frac{t}{n}, k \right) \right)^2 \\ &\leq \frac{K \|\phi_n\|_{\infty, V}^2}{\rho_2^{(h_n)}(\phi_n)^2} \sum_k \frac{[\exp(-k^2 b^2 / 2) - 1]^2}{k^2}. \end{aligned} \quad (52)$$

Since $|1 - \exp(-k^2 b^2 / 2)| \leq \min\{1, \frac{k^2 b^2}{2}\}$ the sum is bounded by

$$K \sum_{|k| \leq 1/b} \frac{k^2 b^4}{4} + K \sum_{|k| > 1/b} \frac{1}{k^2} = O(b)$$

which implies

$$\sqrt{n} \left[\left\{ F_n \left(\frac{\phi_n^*}{\rho_2^{(h_n)}(\phi_n)} \right) - \mathbf{E} F_n \left(\frac{\phi_n^*}{\rho_2^{(h_n)}(\phi_n)} \right) \right\} - \left\{ F_n \left(\frac{\phi_n}{\rho_2^{(h_n)}(\phi_n)} \right) - \mathbf{E} F_n \left(\frac{\phi_n}{\rho_2^{(h_n)}(\phi_n)} \right) \right\} \right] \xrightarrow{P} 0 \quad (53)$$

if $\frac{b \|\phi_n\|_{\infty, V}^2}{\rho_2^{(h_n)}(\phi_n)^2} = o(1)$, e.g. for $b = 1/n$.

We now prove a CLT for $\sqrt{n} [F_n(\phi_{nj}^* / \rho_2^{(h_n)}(\phi_{nj})) - \mathbf{E} F_n(\phi_{nj}^* / \rho_2^{(h_n)}(\phi_{nj}))]_{j=1, \dots, d}$ by applying Lemma 5.6(i) and Lemma 5.7(i) of Dahlhaus and Polonik (2008). We obtain from (51) and Lemma 5.1 with $\tilde{\phi}(k) := \max\{\sup_u |\hat{\phi}(u, k)|, \sup_u |\hat{\phi}(u, -k)|\}$

$$\sup_k \frac{|k| \tilde{\phi}_n^*(k)}{\rho_2^{(h_n)}(\phi_n)} \leq K \frac{\|\phi_n\|_{\infty, V}}{\rho_2^{(h_n)}(\phi_n)} \quad \text{and} \quad \sup_k \frac{|k| V(\hat{\phi}_n^*(\cdot, k))}{\rho_2^{(h_n)}(\phi_n)} \leq K \frac{\|\phi_n\|_{V, V}}{\rho_2^{(h_n)}(\phi_n)}.$$

Furthermore, we have with $\rho_\infty(\phi) := \sum_{k=-\infty}^{\infty} \tilde{\phi}(k)$

$$\begin{aligned} \rho_\infty\left(\frac{\phi_n^*}{\rho_2^{(h_n)}(\phi_n)}\right) &\leq \frac{2\pi\|\phi_n\|_{\infty, \infty}}{\rho_2^{(h_n)}(\phi_n)} + K \frac{\|\phi_n\|_{\infty, V}}{\rho_2^{(h_n)}(\phi_n)} \sum_{k=1}^{\infty} \frac{1}{|k|} \exp(-k^2 b^2/2) \\ &\leq K \frac{\|\phi_n\|_{\infty, \infty} + \log(b^{-1})\|\phi_n\|_{\infty, V}}{\rho_2^{(h_n)}(\phi_n)}. \end{aligned}$$

For $b = 1/n$ (24) implies that the above terms are all of order $o(\sqrt{n})$. In addition

$$\rho_{2,n}(\phi)^2 \leq \rho_2(\phi)^2 + \frac{1}{n} \rho_\infty(\phi) \sup_k V(\hat{\phi}(\cdot, k))$$

leading to

$$\rho_{2,n}^{(h_n)}\left(\frac{\phi_n^*}{\rho_2^{(h_n)}(\phi_n)}\right)^2 \leq \frac{\rho_2^{(h_n)}(\phi_n^*)^2}{\rho_2^{(h_n)}(\phi_n)^2} + K \frac{\|\phi_n\|_{\infty, \infty} + \log(b^{-1})\|\phi_n\|_{\infty, V}}{\sqrt{n} \rho_2^{(h_n)}(\phi_n)} \frac{\|\phi_n\|_{V, V}}{\sqrt{n} \rho_2^{(h_n)}(\phi_n)} \leq K.$$

Therefore the remainder term R_n in Lemma 5.6(i) and the higher cumulants in Lemma 5.7(i) in Dahlhaus and Polonik (2008) converge to zero. With (25) this implies the convergence of all cumulants and therefore

$$\sqrt{n} \left[F_n\left(\frac{\phi_{nj}^*}{\rho_2^{(h_n)}(\phi_{nj})}\right) - \mathbf{E} F_n\left(\frac{\phi_{nj}^*}{\rho_2^{(h_n)}(\phi_{nj})}\right) \right]_{j=1, \dots, d} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma)$$

and with (53) the same CLT for $\sqrt{n} [F_n(\phi_{nj}/\rho_2^{(h_n)}(\phi_{nj})) - \mathbf{E} F_n(\phi_{nj}/\rho_2^{(h_n)}(\phi_{nj}))]_{j=1, \dots, d}$.

From Lemma 5.5(i) in Dahlhaus and Polonik (2008) we obtain with $\ell(k)$ as in (2) and Lemma 5.1

$$\begin{aligned} \sqrt{n} \left| \mathbf{E} F_n\left(\frac{\phi_n}{\rho_2^{(h_n)}(\phi_n)}\right) - F\left(\frac{\phi_n}{\rho_2^{(h_n)}(\phi_n)}\right) \right| \\ \leq K \frac{\|\phi_n\|_{\infty, \infty} + (\log n) \|\phi_n\|_{\infty, V} + \|\phi_n\|_{V, \infty} + \|\phi_n\|_{V, V}}{\sqrt{n} \rho_2^{(h_n)}(\phi_n)} = o(1) \end{aligned}$$

which finally proves Theorem 3.2. \square

We now prove the properties of the sum estimator $F_n^\Sigma(\phi)$ from Section 4.1. It is easy to show that

$$F_n^\Sigma(\phi) = \frac{1}{2\pi n} \sum_{t=1}^n \sum_k \hat{\phi}_\Sigma\left(\frac{t}{n}, -k\right) X_{[t+1/2+k/2],n}^{(h_n)} X_{[t+1/2-k/2],n}^{(h_n)} \quad (54)$$

with

$$\hat{\phi}_\Sigma\left(\frac{t}{n}, -k\right) := \frac{2\pi}{n} \sum_{j=1}^n \phi\left(\frac{t}{n}, \lambda_j\right) \exp\{-i\lambda_j k\}.$$

Recall that

$$F_n(\phi) = \frac{1}{2\pi n} \sum_{t=1}^n \sum_k \hat{\phi}\left(\frac{t}{n}, -k\right) X_{[t+1/2+k/2],n}^{(h_n)} X_{[t+1/2-k/2],n}^{(h_n)}.$$

We need the following lemma.

Lemma 5.3 *We have for $|k| < n$*

$$\max_t \left| \hat{\phi}_\Sigma\left(\frac{t}{n}, -k\right) - \hat{\phi}\left(\frac{t}{n}, -k\right) \right| \leq \frac{K}{n - |k|} \|\phi\|_{\infty, V}.$$

PROOF. The proof can e.g. be found in Dahlhaus (1985), Lemma 3.1. \square

5.4 (Proof of Theorem 4.1)

(i) As in the proof of Lemma 5.7(i) in Dahlhaus and Polonik (2008) we obtain (with $\hat{\phi}(\frac{t}{n}, -k)$ replaced by $\hat{\phi}_\Sigma(\frac{t}{n}, -k) - \hat{\phi}(\frac{t}{n}, -k)$)

$$\begin{aligned} n \operatorname{var} [F_n^\Sigma(\phi) - F_n(\phi)] &\leq \frac{K}{n} \sum_{t_1, t_2} \sum_{k_1, k_2} \left| \hat{\phi}_\Sigma\left(\frac{t_1}{n}, -k_1\right) - \hat{\phi}\left(\frac{t_1}{n}, -k_1\right) \right| \left| \hat{\phi}_\Sigma\left(\frac{t_2}{n}, -k_2\right) - \hat{\phi}\left(\frac{t_2}{n}, -k_2\right) \right| \\ &\times \left[\sum_{i,j} \frac{1}{\ell(t_1^+ - i)\ell(t_1^- - j)\ell(t_2^+ - i)\ell(t_2^- - j)} + \sum_{i,j} \frac{1}{\ell(t_1^+ - i)\ell(t_1^- - j)\ell(t_2^+ - j)\ell(t_2^- - i)} \right. \\ &\left. + \sum_i \frac{1}{\ell(t_1^+ - i)\ell(t_1^- - i)\ell(t_2^+ - i)\ell(t_2^- - i)} \right]. \end{aligned} \quad (55)$$

With Lemma 5.3 and the relation

$$\sum_{j=-\infty}^{\infty} \frac{1}{\ell(k+j)} \frac{1}{\ell(j)} \leq \frac{K}{\ell(k)} \quad (56)$$

we now obtain as an upper bound for the first term in the $[\dots]$ - brackets in (55)

$$\frac{K}{n} \|\phi\|_{\infty, V}^2 \sum_{k_1, k_2} \frac{1}{n - |k_1|} \frac{1}{n - |k_2|} \sum_{t_1, t_2} \frac{1}{\ell(t_1 - t_2 + (k_1 - k_2)/2) \ell(t_1 - t_2 - (k_1 - k_2)/2)}.$$

The substitution $s = t_1 - t_2$ and the fact that there are at most $n - |k_2|$ values of t_2 (due to the restriction $1 \leq [t_2 + 1/2 \pm k_2/2] \leq n$, see (12)) lead to the bound

$$\frac{K}{n} \|\phi\|_{\infty, V}^2 \sum_{k_1, k_2} \frac{1}{n - |k_1|} \sum_s \frac{1}{\ell(s + (k_1 - k_2)/2) \ell(s - (k_1 - k_2)/2)} \leq \frac{K \log n}{n} \|\phi\|_{\infty, V}^2.$$

The second term in the $[\dots]$ - brackets in (55) can be estimated in the same way and the third term obviously is smaller than the first one.

Furthermore, we have

$$\mathbf{E}F_n(\phi) = \frac{1}{2\pi n} \sum_{t, |k| \leq n} \hat{\phi}\left(\frac{t}{n}, -k\right) \text{cov}\left(X_{[t+1/2+k/2], n}^{(h_n)}, X_{[t+1/2-k/2], n}^{(h_n)}\right). \quad (57)$$

Since

$$\sup_t |\text{cov}(X_{t, n}^{(h_n)}, X_{t+k, n}^{(h_n)})| \leq \frac{K}{\ell(k)} \quad (58)$$

(c.f. Dahlhaus and Polonik (2008), Proposition 5.4) we obtain from (54) and Lemma 5.3

$$\mathbf{E}F_n^\Sigma(\phi) - \mathbf{E}F_n(\phi) \leq K \|\phi\|_{\infty, V} \frac{1}{n} \sum_{t=1}^n \sum_{k: 1 \leq [t+1/2 \pm k/2] \leq n} \frac{1}{n - |k|} \frac{1}{\ell(k)} \leq \frac{K}{n} \|\phi\|_{\infty, V} \quad (59)$$

which implies assertion (i). (ii) and (iii) are immediate consequences. \square

5.5 (Proof of Theorem 4.3)

(i) We set $\phi = \phi_n$ for short. We obtain from Lemma 5.7(i) in Dahlhaus and Polonik (2008) (with $\hat{\phi}_{t, k} - \hat{\phi}(\frac{t}{n}, k)$ instead of $\hat{\phi}(\frac{t}{n}, k)$), (45), (46) and Lemma 5.1

$$n \text{var}[F_n^*(\phi) - F_n(\phi)] \leq \frac{K}{n} \sum_{t=1}^n \sum_{k=-\infty}^{\infty} (\hat{\phi}_{t, k} - \hat{\phi}(\frac{t}{n}, k))^2 \quad (60)$$

$$\leq \frac{K}{n} \left(\|\phi\|_{\infty, \infty} \|\phi\|_{V, \infty} + \|\phi\|_{\infty, V} \|\phi\|_{V, V} \sum_{k=1}^{\infty} \frac{\min\{|k|, n\}}{k^2} \right) \quad (61)$$

$$\leq K \frac{\|\phi\|_{\infty, \infty} \|\phi\|_{V, \infty} + \log n \|\phi\|_{\infty, V} \|\phi\|_{V, V}}{n}. \quad (62)$$

Furthermore, we obtain from (57), (58) and (46)

$$n [\mathbf{E}F_n^*(\phi) - \mathbf{E}F_n(\phi)]^2 \leq \frac{K}{n} \left[\sum_k \frac{1}{\ell(k)} \sum_t |\hat{\phi}_{t,k} - \hat{\phi}(\frac{t}{n}, k)| \right]^2 \leq K \frac{\|\phi\|_{V,\infty} + \|\phi\|_{V,V}}{n} \quad (63)$$

that is the result.(ii) and (iii) are immediate consequences. \square

Acknowledgement: The author is very grateful to two anonymous referees whose comments have led to substantial improvements.

This work has been supported by the Deutsche Forschungsgemeinschaft (DA 187/15-1).

References

- Dahlhaus, R., 1985, On the asymptotic distribution of Bartlett's U_p -statistic. *Journal of Time Series Analysis* **6** 213-227.
- Dahlhaus, R., 1997, Fitting time series models to nonstationary processes. *Annals of Statistics* **25** 1-37.
- Dahlhaus, R., 2000, A likelihood approximation for locally stationary processes. *Annals of Statistics* **28** 1762-1794. .
- Dahlhaus, R. and M.H. Neumann, 2001, Locally adaptive fitting of semiparametric models to nonstationary time series. *Stochastic Processes and their Applications* **91** 277-308.
- Dahlhaus, R. and W. Polonik, 2006, Nonparametric quasi-maximum likelihood estimation for Gaussian locally stationary processes. *Annals of Statistics* **34** 2790-2824.
- Dahlhaus, R. and W. Polonik, 2008, Empirical spectral processes for locally stationary time series. *Bernoulli*, to appear.
- Davis, R.A. and Lee, T., and G. Rodriguez-Yam, 2005, Structural Break Estimation for Nonstationary Time Series Models. *Journal of the American Statistical Association* **101** 223 - 239.
- Fryzlewicz, P., Sapatinas, T. and S. Subba Rao, 2006, A Haar-Fisz technique for locally stationary volatility estimation. *Biometrika* **93** 687-704.
- Fuller, W.A., 1996, *Introduction to Statistical Time Series*. Sec. Ed. John Wiley & Sons, New York.
- Jentsch, C., 2006, Asymptotik eines nicht-parametrischen Kernschätzers für zeitvariable autoregressive Prozesse. Diploma thesis, University of Braunschweig.
- Kim, W., 2001, Nonparametric kernel estimation of evolutionary autoregressive processes. Discussion paper 103. Sonderforschungsbereich 373, Berlin.
- Moulines, E., Priouret, P. and F. Roueff, 2005, On recursive estimation for locally stationary time varying autoregressive processes. *Annals of Statistics* **33** 2610-2654.
- Nason, G. P., von Sachs, R. and G. Kroisandt, 2000, Wavelet processes and adaptive estimation of evolutionary wavelet spectra. *Journal of the Royal Statistical Society: Series B* **62** 271-292.

- Neumann, M.H. and R. von Sachs, 1997, Wavelet thresholding in anisotropic function classes and applications to adaptive estimation of evolutionary spectra. *Annals of Statistics* **25** 38 - 76.
- Paparoditis, E., 2006, Testing temporal constancy of the spectral structure of a time series. Technical Report 15/2006, University of Cyprus, Nicosia.
- Priestley, M.B., 1965, Evolutionary spectra and non-stationary processes. *Journal of the Royal Statistical Society: Series B* **27** 204-237.
- Robinson, P. M., 1989, Nonparametric estimation of time-varying parameters. In: *Statistical Analysis and Forecasting of Economic Structural Change* (P. Hackl, ed.) 253-264. Springer, Berlin.
- Sakiyama, K. and M. Taniguchi, 2004, Discriminant analysis for locally stationary processes. *Journal of Multivariate Analysis* **90** 282-300.
- Sergides, M. and E. Paparoditis, 2007, Frequency domain tests of semiparametric hypotheses for locally stationary processes. Preprint, University of Cyprus, Nicosia.
- Van Bellegem, S. and R. Dahlhaus, 2006, Semiparametric estimation by model selection for locally stationary processes. *Journal of the Royal Statistical Society: Series B* **68** 721-746.
- Whittle, P., 1953, Estimation and information in stationary time series. *Arkiv för Matematik* **2** 423-434.