Empirical spectral processes for locally stationary time series

Rainer Dahlhaus*
Institut für Angewandte Mathematik
Universität Heidelberg
Im Neuenheimer Feld 294
69120 Heidelberg
Germany

Wolfgang Polonik
Department of Statistics
University of California
Davis, CA 95616-8705
USA

January 21, 2007

Abstract

A time varying empirical spectral process indexed by classes of functions is defined for locally stationary time series. We derive weak convergence in a function space, prove a maximal exponential inequality and a Glivenko-Cantelli type convergence result. The results use conditions based on metric entropy of the index class. In contrast to related earlier work no Gaussian assumption is made. As applications quasi likelihood estimation, goodness of fit testing and inference under model misspecification are discussed. In an extended application uniform rates of convergence are derived for local Whittle-estimates of the parameter curves of locally stationary time series models.

Running title: Empirical spectral processes for time series.

Keywords: Empirical spectral process, asymptotic normality, quadratic forms, locally stationary processes, nonstationary time series.

AMS 2000 Mathematics Subject Classification: Primary 62M10; secondary 62E20.

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, 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 and Sakiyama and Taniguchi (2004)on discriminant analysis for locally stationary processes.

In this paper we emphasize the relevance of the empirical spectral process for locally stationary time series. During the last decade the theory of empirical processes has developed considerably and the number of statistical problems approached utilizing concepts from empirical process theory is steeply increasing. In this paper we show how large parts of the existing methodology on empirical processes can fruitfully be used for time series analysis of locally stationary processes. In our setup the role of the empirical distribution of iid data is taken over by the empirical time-varying spectral measure. This generalizes a similar approach for stationary time series, c.f. Dahlhaus (1988), Mikosch and Norvaisa (1997), Fay and Soulier (2001). An overview over these methods and some references to the existing literature on empirical process techniques in other settings may be found in Dahlhaus and Polonik (2002).

In Section 2 we introduce the empirical spectral process indexed by classes of functions, derive its convergence including a functional central limit theorem, prove a maximal exponential inequality and a Glivenko-Cantelli type convergence result. These results use conditions based on the metric entropy of the index class.

The empirical spectral process plays a key role in many statistical applications. In Section 3 we briefly discuss parametric quasi likelihood estimation, nonparametric quasi likelihood estimation, inference under model misspecification by stationary models an local estimates. An extended application is given in Section 4 where uniform rates of convergence are derived for local Whittle-estimates of the parameter curves of locally stationary time series models.

Although our concept is based on empirical process techniques in the frequency domain, there exist many applications in the time domain. Section 3 and Section 4 contain many examples in the time domain - particularly with time varying ARMA - models.

The empirical spectral process for locally stationary processes has also briefly been considered in Dahlhaus and Polonik (2006) in the special context of nonparametric estimation. In comparison to that paper we consider here also the case of non-Gaussian processes and use weaker assumptions on the underlying process. We mention that the assumptions on the underlying process are very week allowing for jumps in the parameter curves by assuming bounded variation instead of continuity in time direction.

All proofs are delegated without further reference to Section 5.

2 The time varying empirical spectral process

In this section we define the empirical spectral process and derive its properties including a functional central limit theorem and a maximal exponential inequality.

2.1 Locally stationary processes

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. As in nonparametric regression we rescale the functions in time to the unit interval in order to achieve a meaningful asymptotic theory. The following definition of locally stationarity is the same as used in Dahlhaus and Polonik (2006). It is more general than for example in Dahlhaus (1997) since the parameter curves are allowed to have jumps.

Let

$$V(g) = \sup\Big\{\sum_{k=1}^{m} |g(x_k) - g(x_{k-1})| : 0 \le x_o < \ldots < x_m \le 1,\, m \in \mathbf{N}\Big\} \tag{1}$$

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

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

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

$$X_{t,n} = \sum_{j=-\infty}^{\infty} a_{t,n}(j)\, \varepsilon_{t-j} \tag{2}$$

satisfying the following conditions:

$$\sup_t |a_{t,n}(j)| \le \frac{K}{\ell(j)} \quad \textit{(with } K \textit{ not depending on } n\textit{)}, \tag{3}$$

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

$$\sup_u |a(u,j)| \le \frac{K}{\ell(j)}, \tag{4}$$

$$\sup_j \sum_{t=1}^{n} |a_{t,n}(j) - a(\frac{t}{n}, j)| \le K, \tag{5}$$

$$V(a(\cdot, j)) \le \frac{K}{\ell(j)}. \tag{6}$$

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 $\kappa_4 := cum_4(\varepsilon_t)$.

The above assumptions are discussed in Remark 2.12 below.

Definition 2.2 (Time varying spectral density and covariance) Let $X_{t,n}$ be a locally stationary process. The function

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

with

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

is the time varying spectral density, and

$$c(u,k) := \int_{-\pi}^{\pi} f(u,\lambda) \exp(i\lambda k) d\lambda = \sum_{j=-\infty}^{\infty} a(u,k+j) a(u,j) \tag{7}$$

is the time varying covariance of lag k at rescaled time u.

For a deeper understanding of the time varying covariance see also Proposition 5.4. 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 $Y_t = \Sigma_j\, a(j)\varepsilon_{t-j}$ is stationary with $|a(j)| \le K/\ell(j)$ and ϕ is of bounded variation. From the following proposition it follows, that time varying ARMA (tvARMA) models whose coefficient functions are of bounded variation are locally stationary in the above sense. The result is proved in Section 6.

Proposition 2.3 (tvARMA) *Consider the system of difference equations*

$$\sum_{j=0}^{p} \alpha_j(\frac{t}{n}) X_{t-j,n} = \sum_{k=0}^{q} \beta_k(\frac{t}{n}) \sigma(\frac{t-k}{n}) \varepsilon_{t-k} \tag{8}$$

where ε_t are iid with $E\varepsilon_t = 0$, $E|\varepsilon_t| < \infty$, $\alpha_0(u) \equiv \beta_0(u) \equiv 1$ and $\alpha_j(u) = \alpha_j(0)$, $\beta_k(u) = \beta_k(0)$ for $u < 0$. If all $\alpha_j(\cdot)$ and $\beta_k(\cdot)$ 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 (3)-(6) of Definition 2.1. The time varying spectral density is given by

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

2.2 Convergence of the empirical spectral process

The empirical spectral process is defined by

$$E_n(\phi) = \sqrt{n} \Big(F_n(\phi) - F(\phi) \Big) \tag{9}$$

where

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

and

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

with the pre-periodogram

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

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).

Many statistics occurring in the analysis of nonstationary time series can be written as a functional of $F_n(\phi)$. Several examples are discussed in Section 3 and Section 4.

We first prove a central limit theorem for $E_n(\phi)$ under the assumption, that we have bounded variation in both components of $\phi(u,\lambda)$. Besides the definition in (1) we need a definition in two dimensions. Let

$$V^2(\phi) = \sup\Big\{ \sum_{j,k=1}^{\ell,m} |\phi(u_j,\lambda_k) - \phi(u_{j-1},\lambda_k) - \phi(u_j,\lambda_{k-1}) + \phi(u_{j-1},\lambda_{k-1})| \; : \\ 0 \le u_o < \ldots < u_\ell \le 1;\; -\pi \le \lambda_o < \ldots < \lambda_m \le \pi;\; \ell, m \in \mathbf{N}\Big\}.$$

For simplicity we set

$$\|\phi\|_{\infty,V} := \sup_u V\big(\phi(u,\cdot)\big), \quad \|\phi\|_{V,\infty} := \sup_\lambda V\big(\phi(\cdot,\lambda)\big),$$

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

Theorem 2.4 *Let $X_{t,n}$ be a locally stationary process and $\phi_1,\ldots,\phi_k$ functions with $\|\phi_j\|_{\infty,V}$, $\|\phi_j\|_{V,\infty}$, $\|\phi_j\|_{V,V}$ and $\|\phi_j\|_{\infty,\infty}$ being finite ($j = 1,\ldots,k$). Then*

$$\big(E_n(\phi_j)\big)_{j=1,\ldots,k} \xrightarrow{\mathcal{D}} \big(E(\phi_j)\big)_{j=1,\ldots,k}$$

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

$$\operatorname{cov}\big(E(\phi_j),E(\phi_k)\big) = 2\pi \int_0^1 \int_{-\pi}^{\pi} \phi_j(u,\lambda)\,[\phi_k(u,\lambda) + \phi_k(u,-\lambda)]\, f^2(u,\lambda)\, d\lambda\, du \\ + \kappa_4 \int_0^1 \Big(\int_{-\pi}^{\pi} \phi_j(u,\lambda_1) f(u,\lambda_1)\, d\lambda_1 \Big)\Big(\int_{-\pi}^{\pi} \phi_k(u,\lambda_2) f(u,\lambda_2)\, d\lambda_2 \Big)\, du.$$

Remark 2.5 (i) We mention that Theorem 5.3 contains a similar statement under a different set of conditions which is obtained as a by-product from our calculations. Furthermore we mention that Theorem 2.4 also holds if a data-taper is used, i.e. if $F_n(\phi)$ and $F(\phi)$ are defined as in (42) and (44) (in that case we in addition need Assumption 5.1 and $\operatorname{cov}\big(E(\phi_j), E(\phi_k)\big)$ must be replaced by $c_E^h(\phi_j,\phi_k)$ as defined in Theorem 5.3). For simplicity we consider the tapered case only in Section 5.

(ii) In contrast to earlier results (cf. Dahlhaus and Neumann, 2001, Lemma 2.1) the assumptions on $\phi(u,\lambda)$ and $f(u,\lambda)$ are very weak. In particular we allow for non-continuous behavior.

(iii) In the stationary case where $\phi_j(u,\lambda) = \tilde{\phi}_j(\lambda)$ and $f(u,\lambda) = \tilde{f}(\lambda)$ this is the classical central limit theorem for the weighted periodogram (see Example 3.3 below).

(iv) The limit behavior for complex valued ϕ_j can easily be derived from Theorem 2.4 by considering the real and imaginary part separately.

In Theorem 2.10 a functional central limit theorem indexed by function spaces and in Theorem 2.11 a Glivenko-Cantelli type theorem are proved. The central ingredient of their proofs will be an exponential inequality for the empirical spectral process and a maximal inequality derived in the next section.

2.3 A maximal exponential inequality

Let

$$\rho_2(\phi) := \Big(\int_0^1 \int_{-\pi}^{\pi} \phi(u,\lambda)^2 d\lambda\, du\Big)^{1/2}, \qquad \rho_{2,n}(\phi) := \Big(\frac{1}{n}\sum_{t=1}^{n}\int_{-\pi}^{\pi} \phi(\frac{t}{n},\lambda)^2 d\lambda\Big)^{1/2} \tag{13}$$

and

$$\tilde{E}_n(\phi) := \sqrt{n}\,\big(F_n(\phi) - \mathbf{E}F_n(\phi)\big). \tag{14}$$

Theorem 2.6 (Exponential inequality) *Let $X_{t,n}$ be a locally stationary process with $E|\varepsilon_t|^k \le C_\varepsilon^k$ for all $k \in \mathbf{N}$ for the ε_t from Definition 2.1. Then we have for all $\eta > 0$*

$$P(|\tilde{E}_n(\phi)| \ge \eta) \le c_1 \exp\Big(-c_2\sqrt{\frac{\eta}{\rho_{2,n}(\phi)}}\Big) \tag{15}$$

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

Remark 2.7 (i) In the Gaussian case it is possible to omit the $\sqrt{\cdot}$ in (15) or to prove a Bernstein-type inequality which is even stronger (cf. Dahlhaus and Polonik, 2006, Theorem 4.1).

(ii) To treat the bias $\mathbf{E}F_n(\phi) - F(\phi)$ we set $F^+(\phi) := \frac{1}{n}\sum_{t=1}^{n}\int_{-\pi}^{\pi} \phi(\frac{t}{n},\lambda)\, f(\frac{t}{n},\lambda)\, d\lambda$. Then we have (for a proof see (5.9))

$$\sqrt{n}\,|\mathbf{E}F_n(\phi) - F^+(\phi)| \;\le\; K\,\rho_{2,n}(\phi), \tag{16}$$

$$\sqrt{n}\,|F^+(\phi) - F(\phi)| \;\le\; \frac{K}{\sqrt{n}}\Big(\|\phi\|_{V,\infty} + \|\phi\|_{\infty,\infty}\Big), \tag{17}$$

and

$$\rho_{2,n}(\phi)^2 \le \rho_2(\phi)^2 + \frac{4\pi}{n}\|\phi\|_{V,\infty}\,\|\phi\|_{\infty,\infty}, \tag{18}$$

leading for example to the exponential inequality

$$P(|E_n(\phi)| \ge \eta) \le c_1' \exp\Big\{-c_2'\frac{\eta^{1/2}}{\big(\rho_2(\phi) + \frac{1}{\sqrt{n}}\|\phi\|_{V,\infty} + \frac{1}{\sqrt{n}}\|\phi\|_{\infty,\infty}\big)^{1/2}}\Big\}. \tag{19}$$

An alternative inequality used later is

$$\sqrt{n}\,|\mathbf{E}F_n(\phi) - F^+(\phi)| \;\le\; K\,\frac{\log n}{\sqrt{n}}\,\|\phi\|_{\infty,V} + \frac{K}{\sqrt{n}}\,\|\phi\|_{\infty,\infty}. \tag{20}$$

The above exponential inequality is the core of the proof of the following result which then leads to stochastic equicontinuity of the empirical spectral process. Analogously to standard empirical process theory, stochastic equicontinuity is crucial for proving tightness.

As for the standard empirical process the results for the function indexed empirical spectral process $\big(E_n(\phi), \phi \in \Phi\big)$, are derived under conditions on the richness of Φ measured by the metric entropy. For each $\epsilon > 0$, the *covering number* of Φ with respect to the metric ρ_2 is defined by

$$N(\epsilon, \Phi, \rho_2) = \inf \{n \geq 1 : \exists\, \phi_1, \ldots, \phi_n \in \Phi \ \text{ such that} \\ \forall\, \phi \in \Phi \ \ \exists\, 1 \leq i \leq n \text{ with } \rho_2(\phi, \phi_i) \leq \epsilon\},$$

and the *metric entropy* of Φ with respect to ρ_2 by

$$H(\epsilon, \Phi, \rho_2) = \log N(\epsilon, \Phi, \rho_2). \tag{21}$$

For technical reasons we assume that

$$H(\epsilon, \Phi, \rho_2) \leq \tilde{H}_\Phi(\epsilon)$$

with $\tilde{H}_\Phi(\cdot)$ continuous and strictly monotonically decreasing.

We mention that in standard empirical process theory often so-called bracketing numbers are used instead. Here, we do not use bracketing covering numbers since the empirical spectral process is not monotone in ϕ.

Let

$$\tau_{\infty,V} := \sup_{\phi\in\Phi} \|\phi\|_{\infty,V}, \quad \tau_{V,\infty} := \sup_{\phi\in\Phi} \|\phi\|_{V,\infty},$$

$$\tau_{V,V} := \sup_{\phi\in\Phi} \|\phi\|_{V,V} \quad \text{and} \quad \tau_{\infty,\infty} := \sup_{\phi\in\Phi} \|\phi\|_{\infty,\infty}.$$

Measurability: In order to avoid further technical assumptions the following results assume measurability of all random quantities without further mentioning it.

Theorem 2.8 (Maximal inequality) *Let $X_{t,n}$ be a locally stationary process with $E|\varepsilon_t|^k \leq C_\varepsilon^k$ for all $k \in \mathbf{N}$ for the ε_t from Definition 2.1. Suppose that Φ is such that $\tau_{\infty,V}$, $\tau_{V,\infty}$, $\tau_{V,V}$, and $\tau_{\infty,\infty}$ are finite and*

$$\sup_{\phi\in\Phi} \rho_2(\phi) \leq \tau_2 < \infty.$$

Let $L = \max\{K_1,\, K_2, K\}$ where $K_1, K_2 > 0$ are the constants from Lemma 5.13 and K is the constant from (17) and (20).

Then there exists a set B_n (see (74)) with $\lim_{n\to\infty} P(B_n) = 1$ such that for all η satisfying

$$\eta \geq 26\,L\ \max\{\tau_{\infty,V}\,, \tau_{V,\infty}, \tau_{V,V}\,,\ \tau_{\infty,\infty}\}\,\frac{(\log n)^3}{\sqrt{n}} \tag{22}$$

and

$$\eta \geq \frac{72}{c_2^2}\int_0^\alpha \tilde{H}_\Phi(s)^2 ds \quad \text{ with } \quad \alpha := \tilde{H}_\Phi^{-1}\Big(\frac{c_2}{4}\sqrt{\frac{\eta}{\tau_2}}\,\Big), \tag{23}$$

we have

$$P\Big(\sup_{\phi\in\Phi}|\tilde{E}_n(\phi)| > \eta\,,\ B_n\Big) \leq 3\,c_1\ \exp\Big\{-\frac{c_2}{4}\sqrt{\frac{\eta}{\tau_2}}\ \Big\} \tag{24}$$

and

$$P\Big(\sup_{\phi\in\Phi}|E_n(\phi)| > \eta\,,\ B_n\Big) \leq 3\,c_1\ \exp\Big\{-\frac{c_2}{4}\sqrt{\frac{\eta}{\tau_2}}\ \Big\}. \tag{25}$$

where $c_1, c_2 > 0$ are the constants from (15).

Remark 2.9 (i) A maximal inequality for $\{\tilde{E}_n(\phi), \phi\in\Phi\}$ assuming Gaussian innovations can be found in Dahlhaus and Polonik (2006). The additional Gaussian assumption enables a weakening of the crucial assumption (23), essentially replacing $\int_0^\alpha \tilde{H}_\Phi^2(s)\,ds$ by $\int_0^\alpha \tilde{H}_\Phi(s)\,ds$. Furthermore, the resulting exponential inequality is stronger, replacing $\sqrt{\frac{\eta}{\tau}}$ in the exponent of (24) by $\frac{\eta}{\tau}$. On the other hand the proof in the present non-Gaussian case is much more complicated.

(ii) The restriction to the set B_n has several advantages. First it allows for replacing $\rho_{2,n}(\phi)$ by $\rho_2(\phi)$ (more precisely by $\tau_2 = \sup_{\phi\in\Phi}\rho_2(\phi)$) which makes the results much simpler. Furthermore, the extra terms due to the bias as in (19) can be avoided. For many results the set B_n means no restriction, in particular if the probability of an event is calculated as for equicontinuity.

2.4 A functional central limit theorem and a GC-type result

The two Theorems 2.4 and 2.8 are the main ingredients for deriving the following weak convergence result for the process $\{E_n(\phi);\ \phi\in\Phi\}$ in the space $\ell^\infty(\Phi)$ of uniformly bounded (real-valued) functions on Φ, i.e. with $\|g\|_\Phi := \sup_{\phi\in\Phi}|g(\phi)|$ we have $\ell^\infty(\Phi) = \{g : \Phi\to \mathbf{R};\ \|g\|_\Phi < \infty\}$.

Theorem 2.10 (Functional limit theorem) *Let $X_{t,n}$ be a locally stationary process with $E|\varepsilon_t|^k \leq C_\varepsilon^k$ for all $k\in\mathbf{N}$ for the ε_t from Definition 2.1. Furthermore let Φ be such that $\tau_{\infty,V}$, $\tau_{V,\infty}$, $\tau_{V,V}$, $\tau_{\infty,\infty}$ and $\sup_{\phi\in\Phi}\rho_2(\phi)$ are finite. If in addition*

$$\int_0^1 \tilde{H}_\Phi(s)^2\,ds < \infty, \tag{26}$$

then we have as $n \to \infty$ that

$$E_n(\cdot) \to E(\cdot) \qquad \textit{weakly in } \ell_\infty(\Phi)$$

where $\{ E(\phi),\, \phi \in \Phi \}$ denotes a tight, mean zero Gaussian process with covariance structure as given in Theorem 2.4.

Weak convergence in the above theorem means that $\mathbf{E}^*\alpha(E_n) \to \mathbf{E}\alpha(E)$ as $n \to \infty$ for every bounded, continuous real-valued function α on $\ell^\infty(\Phi)$ equipped with the supremum norm where $\mathbf{E}^*$ denotes outer expectation. This Hoffman-Jørgensen type formulation of weak convergence avoids measurability considerations for the process $\{E_n(\phi),\, \phi \in \Phi\}$. Measurability of E_n might become problematic in particular if Φ is not separable. Nevertheless, this notion of weak convergence admits applications of useful probabilistic tools such as continuous mapping theorems. For more details we refer to van der Vaart and Wellner (1996).

Finally we mention our conjecture that it should be possible to prove another version of the above central limit theorem under much weaker moment assumptions on the ε_t if the class Φ is smaller (by avoiding the use of the maximal inequality for proving equicontinuity).

As another application of the maximal inequality we now prove a Glivenko-Cantelli type theorem for the empirical spectral process. Here we allow for a class $\Phi = \Phi_n$ which may be increasing with n. We set $\tau_{\infty,V}^{(n)} := \sup_{\phi \in \Phi_n} \|\phi\|_{\infty,V}$ etc.

Theorem 2.11 *Let $X_{t,n}$ be a locally stationary process with $E|\varepsilon_t|^k \le C_\varepsilon^k$ for all $k \in \mathbf{N}$ for the ε_t from Definition 2.1. Suppose that Φ_n is such that $\tau_{\infty,V}^{(n)}$, $\tau_{V,\infty}^{(n)}$, $\tau_{V,V}^{(n)}$, and $\tau_{\infty,\infty}^{(n)}$ are of order $o(\frac{n}{\log^3 n})$, $\sup_{\phi \in \Phi_n} \rho_2(\phi) \le \tau_2^{(n)} = o(\sqrt{n})$ and*

$$\int_0^1 \tilde{H}_{\Phi_n}(s)^2 \, ds = o(\sqrt{n}) \, .$$

Then

$$\sup_{\phi \in \Phi_n} |F_n(\phi) - F(\phi)| = \sup_{\phi \in \Phi_n} |\frac{1}{\sqrt{n}} E_n(\phi)| \xrightarrow{P} 0 \, .$$

Remark 2.12 (Discussion of Definition 2.1)

(i) The rather complicated construction with different coefficients $a_{t,n}(j)$ and $a(\frac{t}{n}, j)$ is necessary since we need on the one hand a certain smoothness in time direction (guaranteed by bounded variation of the functions $a(u,j)$) and a class which is rich enough to cover interesting examples. For instance Proposition 2.3 implies that the process

$X_{t,n} = \phi(\frac{t}{n})X_{t-1,n} + \varepsilon_t$ has a representation of the form (1). However, the proof of Proposition 2.3 reveals that this $X_{t,n}$ does not have a representation of the form

$$X_{t,n} = \sum_{j=-\infty}^{\infty} a(\frac{t}{n}, j)\varepsilon_{t-j}. \tag{27}$$

(ii) The conditions may give the impression that only homoscedastic innovations ε_t are allowed. This is not true since a time varying scaling factor of the ε_t may be included in the $a_{t,n}(j)$. An example are the tvARMA-models from Proposition 2.3.

(iii) The time varying MA(∞)-representation (2) can easily be transformed into a time varying spectral representation as used e.g. in Dahlhaus (1997). If the ε_t are assumed to be stationary then there exists a Cramér representation

$$\varepsilon_t = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \exp(i\lambda t)\, d\xi(\lambda)$$

where $\xi(\lambda)$ is a process with mean 0 and orthonormal increments (cf. Brillinger, 1981). Let

$$A_{t,n}(\lambda) := \sum_{j=-\infty}^{\infty} a_{t,n}(j)\exp(-i\lambda j). \tag{28}$$

Then

$$X_{t,n} = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \exp(i\lambda t) A_{t,n}(\lambda) d\xi(\lambda). \tag{29}$$

If (5) is replaced by the stronger condition

$$\sup_t |a_{t,n}(j) - a(\frac{t}{n}, j)| \leq \frac{K}{n\,\ell(j)}$$

then it follows

$$\sup_{t,\lambda} |A_{t,n}(\lambda) - A(\frac{t}{n}, \lambda)| \leq K n^{-1} \tag{30}$$

which was assumed in Dahlhaus (1997). Conversely, if we start with (29) and (30) then the conditions of Definition 2.1 can be derived from adequate smoothness conditions on $A(u, \lambda)$.

3 Applications

In this section we give several examples for the statistic $F_n(\phi)$. In all cases the results from Section 2 can be applied. As a nontrivial application the uniform convergence of local Whittle estimates is proved in the next section.

Example 3.1 (Parametric quasi likelihood estimation)
In Dahlhaus (2000) it has been shown that

$$\mathcal{L}_n(\theta) := \frac{1}{4\pi}\frac{1}{n}\sum_{t=1}^{n}\int_{-\pi}^{\pi}\Big\{\log 4\pi^2 f_\theta(\frac{t}{n},\lambda) + \frac{J_n(\frac{t}{n},\lambda)}{f_\theta(\frac{t}{n},\lambda)}\Big\}\,d\lambda \tag{31}$$

is an approximation to $-\log$ Gaussian likelihood of a locally stationary process. The above likelihood is a generalization of the Whittle-likelihood (Whittle, 1953) to locally stationary processes. An example for a locally stationary process with finite dimensional parameter θ is the tvARMA process from Proposition 2.3 with coefficient functions being polynomials in time. Proving the asymptotic properties of

$$\widehat{\theta}_n := \operatorname*{argmin}_{\theta\in\Theta} \mathcal{L}_n(\theta)$$

is simplified a lot by using the above properties of the empirical spectral process. We give a brief sketch. Let

$$\mathcal{L}(\theta) := \frac{1}{4\pi}\int_0^1\int_{-\pi}^{\pi}\Big\{\log 4\pi^2 f_\theta(u,\lambda) + \frac{f(u,\lambda)}{f_\theta(u,\lambda)}\Big\}\,d\lambda\,du \tag{32}$$

be (up to a constant) the asymptotic Kullback-Leibler divergence between the true process and the fitted model (cf. Dahlhaus, 1996a, Theorem 3.4 ff) and

$$\theta_0 := \operatorname*{argmin}_{\theta\in\Theta} \mathcal{L}(\theta)$$

the best approximating parameter from Θ (this is the true parameter if the model is correctly specified). We have

$$\mathcal{L}_n(\theta) - \mathcal{L}(\theta) = \frac{1}{\sqrt{n}}\,E_n(\frac{1}{4\pi}f_\theta^{-1}) + R_{log}(f_\theta)$$

with

$$R_{log}(f_\theta) := \frac{1}{4\pi}\int_{-\pi}^{\pi}\Big[\frac{1}{n}\sum_{t=1}^{n}\log f_\theta(\frac{t}{n},\lambda) - \int_0^1 \log f_\theta(u,\lambda)\,du\Big]\,d\lambda. \tag{33}$$

Thus, ignoring the R_{log}-term, uniform convergence follows from the Glivenko-Cantelli type Theorem 2.11. The R_{log}-term can be treated as in Dahlhaus and Polonik (2006), Lemma A.2. If Θ is compact and the minimum θ_0 is unique this implies consistency of $\widehat{\theta}_n$. Furthermore,

$$\sqrt{n}\,\nabla\mathcal{L}_n(\theta_0) = E_n(\frac{1}{4\pi}\nabla f_\theta^{-1})$$

and

$$\nabla^2 \mathcal{L}_n(\theta) = \frac{1}{\sqrt{n}}\, E_n\Big(\frac{1}{4\pi}\nabla^2 f_\theta^{-1}\Big) + \frac{1}{n}\sum_{t=1}^{n}\frac{1}{4\pi}\int_{-\pi}^{\pi}\Big(\nabla \log f_\theta(\frac{t}{n},\lambda)\Big)\Big(\nabla \log f_\theta(\frac{t}{n},\lambda)\Big)'\, d\lambda.$$

The first term of $\nabla^2 \mathcal{L}_n(\theta)$ converges uniformly to 0 while the second term converges for $\theta \to \theta_0$ to the Fisher information matrix. The usual Taylor expansion then gives a central limit theorem for $\sqrt{n}(\widehat{\theta}_n - \theta_0)$. For details about the result and examples we refer to Dahlhaus (2000), Theorem 3.1, whose proof is greatly simplified by using the above arguments. Furthermore, the present assumptions are weaker. Due to (36) below the result also covers the misspecified stationary case where the stationary Whittle-likelihood is used with a stationary model but the true process is only locally stationary.

Another application of the empirical spectral process is model selection for Whittle estimates. In Van Bellegem and Dahlhaus (2006) a model selection criterion for semiparametric model selection has been derived. Furthermore, an upper bound for the risk has been proved by using the exponential inequality for the empirical spectral process.

Example 3.2 (Nonparametric quasi likelihood estimation)
In Dahlhaus and Polonik (2006) we have considered the corresponding nonparametric estimator

$$\widehat{f}_n = \underset{g\in\mathcal{F}}{\operatorname{argmin}}\, \mathcal{L}_n(g)$$

with

$$\mathcal{L}_n(g) = \frac{1}{n}\sum_{t=1}^{n}\frac{1}{4\pi}\int_{-\pi}^{\pi}\{\log g(\frac{t}{n},\lambda) + \frac{J_n(\frac{t}{n},\lambda)}{g(\frac{t}{n},\lambda)}\}\, d\lambda. \tag{34}$$

where the contrast functional now is minimized over an “infinite dimensional” target class $\mathcal{F}$ of spectral densities whose complexity is characterized by metric entropy conditions. The optimal rate of convergence has been derived for sieve estimates ín the Gaussian case by using a 'peeling device' and 'chaining' together with an exponential inequality similar to the one in Theorem 2.6 (the exponential inequality is stronger due to the additional Gaussian assumption). It is an open problem whether the optimal rates of convergence can also be achieved for full nonparametric maximum likelihood estimates or (as in the present paper) without the assumption of Gaussianity.

Example 3.3 (Stationary processes/Model misspecification by stationary models)
We start by showing how several classical results for the stationary case can be obtained from the results above. Let $\phi(u,\lambda) = \tilde{\phi}(\lambda)$ be time invariant. Then

$$F_n(\phi) = \int_{-\pi}^{\pi}\tilde{\phi}(\lambda)\frac{1}{n}\sum_{t=1}^{n} J_n(\frac{t}{n},\lambda)\, d\lambda. \tag{35}$$

However we have

$$\frac{1}{n}\sum_{t=1}^{n} J_n(\frac{t}{n},\lambda) = \frac{1}{n}\sum_{t=1}^{n}\frac{1}{2\pi}\sum\nolimits_{k \atop 1\le[t+0.5+k/2],[t+0.5-k/2]\le n} X_{[t+0.5+k/2],n}X_{[t+0.5-k/2],n}\exp(-i\lambda k)$$

$$= \frac{1}{2\pi}\sum_{k=-(n-1)}^{n-1}\left(\frac{1}{n}\sum_{t=1}^{n-|k|} X_t X_{t+|k|}\right)\exp(-i\lambda k)$$

$$= \frac{1}{2\pi n}|\sum_{s=1}^{n} X_s \exp(-i\lambda s)|^2 = I_n(\lambda) \tag{36}$$

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

(i) $\phi(u,\lambda) = \tilde{\phi}(\lambda) = I_{[0,\mu]}(\lambda)$ gives the empirical spectral measure;
(ii) $\phi(u,\lambda) = \tilde{\phi}(\lambda) = \frac{1}{4\pi}\nabla f_\theta^{-1}(\lambda)$ is the score function of the Whittle-likelihood (similar to Example 3.1 above);
(iii) $\phi(u,\lambda) = \tilde{\phi}(\lambda) = \cos\lambda k$ is the empirical covariance estimator of lag k.

Theorem 2.4 gives in all cases the asymptotic distribution - both in the stationary case and in the misspecified case where the true underlying process is only locally stationary. In case (i) Theorem 2.10 leads to a functional central theorem on $C[0,\pi]$ with the supremum-norm. If $\tilde{\phi}(\lambda)$ is a kernel we obtain a kernel estimate of the spectral density (see the remark below).

Example 3.4 (Local estimates) There is an even larger variety of local estimates - some of them are listed below. The asymptotic distribution of these estimates is not covered by Theorem 2.4 since the function $\phi(u,\lambda)$ depends on n in this case. However, in all cases the uniform rate of convergence of these estimators may be derived by using the maximal inequality in Theorem 2.8. A detailed example is given in the next section where the uniform rate of convergence of local Whittle estimates is derived ((iii) below).

For a short overview let $k_n(x) = \frac{1}{b_n}K(\frac{x}{b_n})$ be some kernel with bandwidth b_n. Then

(i) $\phi(u,\lambda) = k_n(u-u_0)\,k_n(\lambda-\lambda_0)$ gives an estimator of the time varying spectral density $f(u_0,\lambda_0)$;
(ii) $\phi(u,\lambda) = k_n(u-u_0)\,\cos\lambda k$ gives a local estimator of the covariance function $c(u_0,k)$;
(iii) $\phi(u,\lambda) = k_n(u-u_0)\,\frac{1}{4\pi}\nabla f_\theta^{-1}(\lambda)$ is the score function of the local Whittle-estimator of the parameter curve $\theta(u_0)$.

4 Uniform convergence of local Whittle estimates

We now study kernel estimates for parameter curves of locally stationary processes and derive uniform consistency from the Glivenko-Cantelli type Theorem 2.11 and a uniform rate of convergence from the maximal inequality in Theorem 2.8. We investigate locally stationary processes 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 is the tvARMA-process from Proposition 2.3.

Let

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

with

$$\mathcal{L}_n(u,\theta) := \frac{1}{4\pi}\frac{1}{n}\sum_{t=1}^{n} \frac{1}{b_n} K\Big(\frac{u-t/n}{b_n}\Big) \int_{-\pi}^{\pi} \Big\{ \log 4\pi^2 f_\theta(\lambda) + \frac{J_n(\frac{t}{n},\lambda)}{f_\theta(\lambda)} \Big\}\, d\lambda. \tag{37}$$

We assume that the kernel K has compact support on $[-\frac{1}{2},\frac{1}{2}]$ and is 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 $b_n \to 0$ and $nb_n \to \infty$ as $n \to \infty$.

In case of a tvAR(p)-process $\widehat{\theta}_n(u)$ is the solution of the local Yule-Walker equations: Let $\widehat{c}_n(u,k) := \frac{1}{n}\sum_t \frac{1}{b_n} K\Big(\frac{u-t/n}{b_n}\Big)\, X_{[t+1/2+k/2],n} X_{[t+1/2-k/2],n}$ (cf. Proposition 5.4), $C_n(u) = (\widehat{c}_n(u,1),\ldots,\widehat{c}_n(u,p))'$ and $\Sigma_n(u) = \{\widehat{c}_n(u,i{-}j)\}_{i,j=1,\ldots,p}$. If $\widehat{\theta}_n(u) = (\widehat{\alpha}_1(u),\ldots,\widehat{\alpha}_p(u),\widehat{\sigma}^2(u))'$ then it is not difficult to show that

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

and

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

We now derive a uniform rate of convergence for $\widehat{\theta}_n(u)$.

Theorem 4.1 *Let $X_{t,n}$ be a locally stationary process with $E|\varepsilon_t|^k \le C_\varepsilon^k$ for all $k \in \mathbf{N}$ for the ε_t from Definition 2.1 and time varying spectral density $f(u,\lambda) = f_{\theta_0(u)}(\lambda)$. Suppose*

(i) θ is identifiable from f_θ (i.e. $f_\theta(\lambda) = f_{\theta'}(\lambda)$ for all λ implies $\theta = \theta'$) and $\theta_0(u)$ lies in the interior of the compact parameter space $\Theta \subseteq \mathbf{R}^d$ for all u;
(ii) $\theta_0(u)$ is twice differentiable with Lipschitz continuous second derivative;
(iii) $f_\theta(\lambda)$ is twice differentiable in θ with Lipschitz continuous second derivative;
(iv) $f_\theta^{-1}(\lambda)$ and the components of $\nabla f_\theta^{-1}(\lambda)$, $\nabla^2 f_\theta^{-1}(\lambda)$ are uniformly bounded in λ and θ;
(v) the minimal eigenvalue of $I(\theta) := \frac{1}{4\pi}\int_{-\pi}^{\pi} \Big(\nabla \log f_\theta(\lambda)\Big)\Big(\nabla \log f_\theta(\lambda)\Big)' d\lambda$ is bounded from

below uniformly in θ.

Then we have for $b_n n >> (\log n)^6$

$$\sup_{u \in [b_n/2\,,\,1-b_n/2]} \big\|\widehat{\theta}_n(u) - \theta_0(u)\big\|_2 = O_p\Big(\frac{1}{\sqrt{b_n n}} + b_n^2\Big)\,,$$

that is for $b_n \sim n^{-1/5}$ we obtain the uniform rate $O_p(n^{-2/5})$.

REMARKS. (i) Instead of assumptions (ii) and (iii) we only need second order differentiability with Lipschitz continuous second order derivative of $f_{\theta_0(u)}(\lambda)$ in u. This enters the following proof only in the estimation of the second summand of (41).
(ii) We conjecture that a similar results holds in case of model misspecification where the model spectral density $f_{\theta_0(u)}(\lambda)$ is only an approximation to the true spectral density $f(u,\lambda)$.
(iii) For tvAR(p)-processes it follows from Moulines, Priouret and Roueff (2005) that this is the optimal rate of convergence.

PROOF. We note at the beginning that the difficult parts of the following proof are handled by using the empirical spectral process and applying Theorems 2.8 and 2.11. We start by proving consistency. We have with

$$\mathcal{L}(u,\theta) := \frac{1}{4\pi}\int_{-\pi}^{\pi}\Big\{\log 4\pi^2 f_\theta(\lambda) + \frac{f(u,\lambda)}{f_\theta(\lambda)}\Big\}\,d\lambda$$

$$\begin{aligned}\mathcal{L}_n(u,\theta) - \mathcal{L}(u,\theta) = {} & \frac{1}{\sqrt{n}}\,E_n\Big(\frac{1}{b_n}K(\frac{u-\cdot}{b_n})\,\frac{1}{4\pi}f_\theta^{-1}\Big)\\ & + \frac{1}{4\pi}\int_{-\pi}^{\pi}\int_0^1 \frac{1}{b_n}\,K\Big(\frac{u-v}{b_n}\Big)\frac{f(v,\lambda)-f(u,\lambda)}{f_\theta(\lambda)}\,dv\,d\lambda\\ & + \frac{1}{4\pi}\int_{-\pi}^{\pi}\log 4\pi^2 f_\theta(\lambda)\,d\lambda\,\Big(\frac{1}{n}\sum_{t=1}^{n}\frac{1}{b_n}\,K\Big(\frac{u-t/n}{b_n}\Big)-1\Big).\end{aligned}$$

We now apply Theorem 2.11 with $\Phi_n = \big\{\frac{1}{b_n}K(\frac{u-\cdot}{b_n})\,\frac{1}{4\pi}f_\theta^{-1}(\,\cdot\,)\,\big|\,u\in[b_n/2,1-b_n/2],\theta\in\Theta\big\}$. It is straightforward to show that $N(\varepsilon,\Phi_n,\rho_2) = K\big/\big(b_n^{(d+4)/2}\varepsilon^{d+2}\big)$ i.e $\int_0^1 \tilde{H}_{\Phi_n}(s)^2\,ds = O((\log b_n)^2)$.

Furthermore $\tau_{\infty,V}^{(n)}$, $\tau_{V,\infty}^{(n)}$, $\tau_{V,V}^{(n)}$, and $\tau_{\infty,\infty}^{(n)}$ are of order $O(b_n^{-1})$ and $\tau_2^{(n)} = \sup_{\phi\in\Phi_n}\rho_2(\phi) = O\big(b_n^{-1/2}\big)$. Thus for $b_n n \geq \log^4 n$ Theorem 2.11 implies

$$\sup_{u\in[b_n/2\,,\,1-b_n/2]}\ \sup_{\theta\in\Theta}\big|\mathcal{L}_n(u,\theta)-\mathcal{L}(u,\theta)\big| \xrightarrow{P} 0\,. \tag{38}$$

The identifiability condition implies that the minimum $\theta_0(u)$ of $\mathcal{L}(u,\theta)$ is unique for all u. By using standard arguments we therefore can conclude that

$$\sup_{u\in[b_n/2\,,\,1-b_n/2]} \big\|\widehat{\theta}_n(u)-\theta_0(u)\big\|_2 \xrightarrow{P} 0\,. \tag{39}$$

We now derive the rate of convergence by using the maximal inequality. We have for each u

$$Z_n(u) := \nabla\mathcal{L}_n\big(u,\widehat{\theta}_n(u)\big) - \nabla\mathcal{L}_n\big(u,\theta_0(u)\big) = \nabla^2\mathcal{L}_n\big(u,\bar{\theta}_n(u)\big)\Big(\widehat{\theta}_n(u)-\theta_0(u)\Big) \tag{40}$$

with $\big|\bar{\theta}_n(u)-\theta_0(u)\big| \le \big|\widehat{\theta}_n(u)-\theta_0(u)\big|$. The main term on the left hand side is

$$\frac{\partial}{\partial\theta_j}\mathcal{L}_n\big(u,\theta_0(u)\big) = \frac{1}{\sqrt{n}}E_n(\phi_n) \tag{41}$$
$$+\int_{-\pi}^{\pi}\int_0^1 \frac{1}{b_n}K\Big(\frac{u-v}{b_n}\Big)\big(f_{\theta_0(v)}(\lambda)-f_{\theta_0(u)}(\lambda)\big)\frac{\partial}{\partial\theta_j}f_\theta^{-1}(\lambda)_{|\theta=\theta_0(u)}\,dv\,d\lambda$$

with $\phi_n(v,\lambda) := \frac{1}{b_n}K_t(\frac{u-v}{b_n})\ \frac{1}{4\pi}\frac{\partial}{\partial\theta_j}f_\theta^{-1}(\lambda)_{|\theta=\theta_0(u)}$. We apply the maximal inequality of Theorem 2.8 to the class $\Phi_n = \big\{\frac{1}{b_n}K(\frac{u-\cdot}{b_n})\,\frac{1}{4\pi}\frac{\partial}{\partial\theta_j}f_\theta^{-1}(\,\cdot\,)\,\big|\,u\in[b_n/2,1-b_n/2],\theta\in\Theta\big\}$. Again we can show $N(\varepsilon,\Phi_n,\rho_2) = K\big/\big(b_n^{(d+4)/2}\varepsilon^{d+2}\big)$ i.e $\int_0^1 \tilde{H}_{\Phi_n}(s)^2\,ds = O((\log b_n)^2)$. Furthermore $\tau_{\infty,V}^{(n)}$, $\tau_{V,\infty}^{(n)}$, $\tau_{V,V}^{(n)}$, and $\tau_{\infty,\infty}^{(n)}$ are of order $O(b_n^{-1})$ and $\tau_2^{(n)} = \sup_{\phi\in\Phi_n}\rho_2(\phi) \sim b_n^{-1/2}$. We now apply Theorem 2.8 with $\eta = \tau_2\,\delta$ for arbitrary δ. If $b_n n >> \log^6 n$ the conditions (22) and (23) are fulfilled and we obtain

$$P\Big(\sup_{\phi\in\Phi_n}|E_n(\phi)| > \tau_2\,\delta\,,\ B_n\Big) \le 3\,c_1\ \exp\Big\{-\frac{c_2}{4}\sqrt{\delta}\ \Big\}\,,$$

i.e. $\sup_{\phi\in\Phi_n}|\frac{1}{\sqrt{n}}E_n(\phi)| = O_p(\frac{1}{\sqrt{b_n n}})$. Since $f_\theta(\lambda)$ and $\theta_0(u)$ are twice differentiable (in θ and u respectively) with Lipschitz continuous second derivative the second summand of (41) can be uniformly bounded with a Taylor expansion by $O_p(b_n^2)$, that is we obtain

$$\sup_{u\in[b_n/2\,,\,1-b_n/2]}\big\|\nabla\mathcal{L}_n\big(u,\theta_0(u)\big)\big\|_2 = O_p(\frac{1}{\sqrt{b_n n}}+b_n^2).$$

If $\widehat{\theta}_n(u)$ lies on the boundary of Θ for some u then $\big\|\widehat{\theta}_n(u)-\theta(u)\big\|_2 \ge \kappa$ for some $\kappa > 0$ and

$$P\Big(\sup_{u\in[b_n/2\,,\,1-b_n/2]}\big\|\nabla\mathcal{L}_n\big(u,\widehat{\theta}_n(u)\big)\big\|_2 > \ \delta\,\frac{1}{\sqrt{b_n n}}\Big)$$
$$\le P\Big(\sup_{u\in[b_n/2\,,\,1-b_n/2]}\big\|\widehat{\theta}_n(u)-\theta_0(u)\big\|_2 \ge \kappa\Big) \to 0,$$

i.e. $\sup_{u\in[b_n/2\,,1-b_n/2]}\|Z_n(u)\|_2 = O_p(\frac{1}{\sqrt{b_n n}})$. In order to obtain the assertion of the theorem from (40) we now prove that the minimal eigenvalue of the matrix $\nabla^2\mathcal{L}_n\big(u,\bar{\theta}_n(u)\big)$ is uniformly bounded from below in probability. We have

$$\nabla^2\mathcal{L}_n(u,\theta) = \frac{1}{\sqrt{n}}\,E_n\Big(\frac{1}{b_n}K\big(\frac{u-\cdot}{b_n}\big)\,\frac{1}{4\pi}\nabla^2 f_\theta^{-1}\Big) + \frac{1}{n}\sum_{t=1}^{n}\frac{1}{b_n}K\big(\frac{u-t/n}{b_n}\big)\,\frac{1}{4\pi}\int_{-\pi}^{\pi}\Big(\nabla\log f_\theta(\lambda)\Big)\Big(\nabla\log f_\theta(\lambda)\Big)'\,d\lambda.$$

Since f_θ is twice differentiable in θ with Lipschitz continuous second derivative we obtain exactly as above from Theorem 2.11 for $b_n n \geq \log^4 n$ and $i,j=1,\ldots,d$

$$\sup_{u\in[b_n/2\,,1-b_n/2]}\ \sup_{\theta\in\Theta}\Big|\frac{1}{\sqrt{n}}\,E_n\Big(\frac{1}{b_n}K\big(\frac{u-\cdot}{b_n}\big)\,\frac{1}{4\pi}\frac{\partial^2}{\partial\theta_i\partial\theta_j}f_\theta^{-1}\Big)\Big| \xrightarrow{P} 0\,.$$

Therefore also

$$\sup_{u\in[b_n/2\,,1-b_n/2]}\Big\|\frac{1}{\sqrt{n}}\,E_n\Big(\frac{1}{b_n}K\big(\frac{u-\cdot}{b_n}\big)\,\frac{1}{4\pi}\big(\nabla^2 f_\theta^{-1}\big)_{|\theta=\bar{\theta}_n(u)}\Big)\Big\|_{spec} \xrightarrow{P} 0\,.$$

Since the minimal eigenvalue of $I(\theta) := \frac{1}{4\pi}\int_{-\pi}^{\pi}\big(\nabla\log f_\theta(\lambda)\big)\big(\nabla\log f_\theta(\lambda)\big)'d\lambda$ is bounded from below by $\lambda_{min}(I)>0$ uniformly in θ this implies

$$\mathbf{P}\Big(\sup_{u\in[b_n/2\,,1-b_n/2]}\big\|\nabla^2\mathcal{L}_n\big(u,\bar{\theta}_n(u)\big)^{-1}\big\|_{spec} \leq \frac{2}{\lambda_{min}(I)}\Big) \to 1\,.$$

Since

$$\big\|\widehat{\theta}_n(u)-\theta_0(u)\big\|_2 \leq \big\|\nabla^2\mathcal{L}_n\big(u,\bar{\theta}_n(u)\big)^{-1}\big\|_{spec}\,\big\|Z_n(u)\big\|_2$$

this implies the result. □

5 Proofs: CLT and exponential inequality

In this section we provide the proofs for the results of Section 2. In particular we derive the asymptotic behavior of the moments of the empirical process.

First 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 $h_n : (0,1] \to [0,\infty)$ is a data taper (with $h_n(u) = I_{(0,1]}(u)$ being the nontapered case of Section 2). This is done for two reasons:
(i) The main reason is that all proofs are greatly simplified since the data taper now automatically takes care of the range of summation ($h_n(t/n)$ is zero for all t outside the

observation domain). The consideration of arbitrary tapers h_n instead of the 'no-taper' $I_{(0,1]}$ does not introduce any extra technical complexity at all.
(ii) The use of a data-taper in the periodogram is standard for stationary time series. It leads to a better small sample performance in the presence of strong peaks in the spectrum. It may turn out that this also holds in the present situation (which requires further investigation). Furthermore, missing data can be modelled with adequate tapers.

As before the empirical spectral process is defined by $E_n(\phi) = \sqrt{n}\big(F_n(\phi) - F(\phi)\big)$ where $F_n(\phi)$ now is

$$F_n(\phi) = \sum_{t=1}^{n} \frac{h_n^2(\frac{t}{n})}{H_{2,n}} \int_{-\pi}^{\pi} \phi(\frac{t}{n},\lambda) J_n^{(h_n)}(\frac{t}{n},\lambda)\, d\lambda \tag{42}$$

with the tapered pre-periodogram

$$J_n^{(h_n)}\left(\frac{t}{n},\lambda\right) = \frac{1}{2\pi} h_n(\frac{t}{n})^{-2} \sum_{k:1\le[t+1/2\pm k/2]\le n} X^{(h_n)}_{[t+1/2+k/2],n} X^{(h_n)}_{[t+1/2-k/2],n} \exp(-i\lambda k) \tag{43}$$

and $H_{2,n} := \sum_{t=1}^{n} h_n(\frac{t}{n})^2$. $F(\phi)$ is its theoretical counterpart

$$F(\phi) = \int_0^1 \frac{h^2(u)}{\|h\|_2^2} \int_{-\pi}^{\pi} \phi(u,\lambda) f(u,\lambda)\, d\lambda\, du. \tag{44}$$

with h defined below. For stationary time series it is often assumed that the percentage of tapered data-points tends to zero as $n \to \infty$. For this reason we allow for dependence of the taper on n.

Assumption 5.1 The data taper $h_n : (0,1] \to [0,\infty)$ fulfills $\sup_n V(h_n) \le C$ and $\sup_{u,n} h_n(u) \le C$ for some $C < \infty$. There exists a function h with $\|h_n - h\|_1 = O(n^{-1+\kappa})$ with some $\kappa \in [0, 1/2)$.

Note that $\kappa = 0$ if $h_n \equiv h$ (in particular in the nontapered case).

We need the following notation. With

$$\hat{\phi}(u,j) := \int_{-\pi}^{\pi} \phi(u,\lambda) \exp(i\lambda j)\, d\lambda \tag{45}$$

we define

$$\rho_\infty(\phi) := \sum_{j=-\infty}^{\infty} \sup_u |\hat{\phi}(u,j)|, \tag{46}$$

$$\tilde{v}(\phi) := \sup_j V(\hat{\phi}(\cdot,j)) \quad \text{and} \quad v_\Sigma(\phi) := \sum_{j=-\infty}^{\infty} V(\hat{\phi}(\cdot,j)). \tag{47}$$

We mention that

$$\sup_{u,\lambda} |\phi(u,\lambda)| \leq \frac{1}{2\pi}\rho_\infty(\phi), \quad \tilde{v}(\phi) \leq v_\Sigma(\phi), \quad \tilde{v}(\phi) \leq \int V(\phi(\cdot,\lambda))\, d\lambda, \quad \rho_2(\phi) \leq \frac{1}{\sqrt{2\,\pi}}\,\rho_\infty(\phi).$$

The idea now is to prove the CLT in Theorem 2.4 by the convergence of all cumulants. The convergence of the cumulants is derived below under the assumptions $\rho_\infty(\phi) < \infty$ and $v_\Sigma(\phi) < \infty$. Unfortunately these assumptions are not fulfilled for functions of bounded variation as assumed in Theorem 2.4. Therefore its proof also uses certain approximation arguments; cf. Section 5.8. The following CLT is a by-product which follows immediately from the cumulant calculations below. It is of independent interest since the result does not follow from Theorem 2.4 (the condition $\rho_\infty(\phi) < \infty$ does not imply bounded variation in λ-direction; furthermore the conditions may be easier to check in some situations).

Therefore the proof of Theorem 2.4 also uses certain approximation arguments; cf. section 5.8.

Assumption 5.2 Suppose $\phi : [0,1] \times [-\pi,\pi] \to \mathbf{R}$ fulfills $\rho_\infty(\phi) < \infty$ and $\tilde{v}(\phi) < \infty$. If the process $X_{t,n}$ is non-Gaussian we assume in addition that $v_\Sigma(\phi) < \infty$.

Theorem 5.3 *Let $X_{t,n}$ be a locally stationary process. Suppose that Assumptions 5.1 and 5.2 hold for $\phi_1,\dots,\phi_k$. Then*

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

where $\big(E(\phi_j)\big)_{j=1,\dots,k}$ is a Gaussian random vector with mean 0 and $\operatorname{cov}(E(\phi_j),E(\phi_k)) = c_E^h(\phi_j,\phi_k)$ *with*

$$c_E^h(\phi_j,\phi_k) = 2\pi \int_0^1 \frac{h^4(u)}{\|h\|_2^4} \int_{-\pi}^{\pi} \phi_j(u,\lambda)\,[\phi_k(u,\lambda) + \phi_k(u,-\lambda)]\, f^2(u,\lambda)\, d\lambda\, du$$
$$+ \kappa_4 \int_0^1 \frac{h^4(u)}{\|h\|_2^4} \Big(\int_{-\pi}^{\pi} \phi_j(u,\lambda_1) f(u,\lambda_1) d\lambda_1\Big)\Big(\int_{-\pi}^{\pi} \phi_k(u,\lambda_2) f(u,\lambda_2) d\lambda_2\Big)\, du.$$

Proof. The result follows from the convergence of all cumulants which is proved in Lemma 5.5(ii), Lemma 5.6(ii) and Lemma 5.7(iii). □

For the following proofs we first need a result on the behaviour (decay) of the covariances of the process. The case $h_n(u) = I_{(0,1]}(u)$ gives the results for the ordinary covariances.

Proposition 5.4 *Let $X_{t,n}$ be a locally stationary process. Suppose that Assumption 5.1 holds. Then we have for all k with some K independent of k and n*

$$\sup_t |\operatorname{cov}(X_{t,n}^{(h_n)}, X_{t+k,n}^{(h_n)})| \leq \frac{K}{\ell(k)}, \tag{48}$$

$$\sup_u |c(u,k)| \leq \frac{K}{\ell(k)}, \tag{49}$$

$$\sum_{t=1}^{n} |\mathrm{cov}(X_{t+k_1,n}^{(h_n)} X_{t-k_2,n}^{(h_n)}) - h_n(\frac{t}{n})^2 c(\frac{t}{n}, k_1+k_2)| \leq K\Big(1 + \frac{min\{|k_1|, n\}}{\ell(k_1+k_2)}\Big), \tag{50}$$

$$V(c(\cdot,k)) \leq \frac{K}{\ell(k)}. \tag{51}$$

Proof. (49) follows from (7), (4) and the relation

$$\sum_{j=-\infty}^{\infty} \frac{1}{\ell(k+j)} \frac{1}{\ell(j)} \leq \frac{K}{\ell(k)} \tag{52}$$

which is easily established. Furthermore, we have with $k = k_1 + k_2$

$$\mathrm{cov}(X_{t+k_1,n}^{(h_n)}, X_{t-k_2,n}^{(h_n)}) = h_n\Big(\frac{t+k_1}{n}\Big) h_n\Big(\frac{t-k_2}{n}\Big) \sum_{j=-\infty}^{\infty} a_{t+k_1}(j+k) a_{t-k_2}(j).$$

For $k_1 = k$ and $k_2 = 0$ this gives (48) by using (3). Replacing $t + k_1$ and $t - k_2$ successively by t gives (50). (51) is obtained similarly. □

A trick which greatly simplifies the following proofs is to set $a_{t,n}(j) = 0$ for $t \notin \{1, \ldots, n\}$ and $j \in \mathbf{Z}$, $a(u,j) = 0$ for $u \notin (0,1]$ and $j \in \mathbf{Z}$, $\phi(u,\lambda) = 0$ for $u \notin (0,1]$ and $\lambda \in [-\pi,\pi]$ and $h_n(u) = 0$ for $u \notin (0,1]$. With this convention (3) - (6), (48) - (51) remain to hold for $u \in \mathbf{R}$, $t \in \mathbf{Z}$, $V(f)$ now denoting the total variation over $\mathbf{R}$, and t in (5) and (50) ranging from $-\infty$ to ∞. Furthermore the summation range of k in (12) and t in (11) can be extended from $-\infty$ to ∞. Therefore, all summation ranges are from $-\infty$ to ∞ in the following proofs unless otherwise indicated.

We also set $\tilde{a}(j) = \sup_u |a(u,j)|$, $\tilde{\phi}(j) = \max\{\sup_u |\hat{\phi}(u,j)|, \sup_u |\hat{\phi}(u,-j)|\}$ and $\tilde{c}(j) = \sup_u |c(u,j)|$.

Lemma 5.5 *(i) Let $X_{t,n}$ be a locally stationary process. Suppose Assumption 5.1 holds and $\phi : [0,1] \times [-\pi,\pi] \to \mathbf{R}$ is a function possibly depending on n. Then we have with $K > 0$*

$$|\mathbf{E}F_n(\phi) - F(\phi)| \leq K n^{-1+\kappa}\Big(\rho_\infty(\phi) + \tilde{v}(\phi)\Big).$$

(ii) If in addition ϕ is independent of n and fulfills Assumption 5.2 then

$$\mathbf{E}E_n(\phi) = O(n^{-1/2+\kappa}).$$

Proof. (i) We have

$$F_n(\phi) = \frac{1}{2\pi H_{2,n}} \sum_{t=1}^{n} \sum_{k} \hat{\phi}(\frac{t}{n}, -k) X^{(h_n)}_{[t+1/2+k/2],n} X^{(h_n)}_{[t+1/2-k/2],n}. \tag{53}$$

Since $H_{2,n} = n\|h\|_2^2 + O(n^\kappa)$ we therefore obtain from Proposition 5.4

$$\begin{aligned} \mathbf{E} F_n(\phi) &= \frac{1}{2\pi H_{2,n}} \sum_{t,\,|k|\leq n} \hat{\phi}(\frac{t}{n}, -k) \operatorname{cov}(X^{(h_n)}_{[t+1/2+k/2],n}, X^{(h_n)}_{[t+1/2-k/2],n}) \\ &= \frac{1}{2\pi H_{2,n}} \sum_{t,k} h_n^2(\frac{t}{n}) \, \hat{\phi}(\frac{t}{n}, -k) \, c(\frac{t}{n}, k) + R \end{aligned} \tag{54}$$

with

$$|R| \leq \frac{K}{n} \sum_{|k|\leq n} \tilde{\phi}(k) \, [1 + \frac{\min(|k|, n)}{\ell(k)}] + K \sum_{|k|>n} \tilde{\phi}(k) \frac{1}{\ell(k)} \leq \frac{K}{n} \rho_\infty(\phi). \tag{55}$$

Furthermore,

$$\begin{aligned} &\Big|\frac{1}{2\pi H_{2,n}} \sum_{t,k} h_n^2(\frac{t}{n}) \, \hat{\phi}(\frac{t}{n}, -k) \, c(\frac{t}{n}, k) - F(\phi)\Big| \\ &\quad\leq \Big| \frac{1}{2\pi} \sum_{t,k} \int_0^{1/n} \Big[\frac{h_n^2(\frac{t}{n})}{\frac{1}{n} H_{2,n}} \, \hat{\phi}(\frac{t}{n}, -k) \, c(\frac{t}{n}, k) \\ &\qquad - \frac{h^2(\frac{t-1}{n} + x)}{\|h\|_2^2} \, \hat{\phi}(\frac{t-1}{n} + x, -k) \, c(\frac{t-1}{n} + x, k) \Big] dx \Big| \\ &\quad\leq \frac{K}{2\pi n} \sum_{k} \big[V(\hat{\phi}(\cdot, -k)) \, \tilde{c}(k) + \tilde{\phi}(k) \, V(c(\cdot, k)) + n^\kappa \tilde{\phi}(k) \tilde{c}(k) \big] \end{aligned} \tag{56}$$

leading to the result. (ii) follows immediately. □

Lemma 5.6 *(i) Let $X_{t,n}$ be a locally stationary process. Suppose Assumption 5.1 holds and $\phi_1, \phi_2 : [0,1] \times [-\pi, \pi] \to \mathbf{R}$ are functions possibly depending on n. Then we have*

$$\operatorname{cov}\big(E_n(\phi_1), E_n(\phi_2)\big) = c_E^h(\phi_1, \phi_2) + R_n$$

with

$$|R_n| \leq \frac{K}{n} \sum_{k_1,k_2} \tilde{\phi}_1(k_1)\tilde{\phi}_2(k_2) \left[1 + \frac{min\{|k_1|, n\}}{\ell(k_1+k_2)}\right] + Kn^{-1+\kappa} \sum_{k_1,k_2} \left[\tilde{\phi}_1(k_1)\tilde{\phi}_2(k_2)\right]$$
$$+ \frac{K}{n} \sum_{k_1,k_2,k_3} \left[\tilde{\phi}_1(k_1)V(\hat{\phi}_2(\cdot,k_2)) + V(\hat{\phi}_1(\cdot,k_1))\tilde{\phi}_2(k_2)\right] \frac{\min\{|k_1|+|k_2|+|k_3|, n\}}{\ell(k_3)\ell(k_1+k_2+k_3)}$$
$$+ \frac{K}{n} \sum_{k_1,k_2} \left[\tilde{\phi}_1(k_1)V(\hat{\phi}_2(\cdot,k_2)) + V(\hat{\phi}_1(\cdot,k_1))\tilde{\phi}_2(k_2)\right] \left[\frac{1}{\ell(k_1)} + \frac{1}{\ell(k_2)}\right].$$

where the last term can be omitted if $X_{t,n}$ is Gaussian.
(ii) If ϕ_1 and ϕ_2 are independent of n and fulfill Assumption 5.2 then $R_n = o(1)$.

PROOF. (i) We have with (53)

$$\mathrm{cov}\big(E_n(\phi_1), E_n(\phi_2)\big) = n\ \mathrm{cov}\big(F_n(\phi_1), F_n(\phi_2)\big)$$
$$= \frac{n}{(2\pi)^2 H_{2,n}^2} \sum_{t_1,t_2,k_1,k_2} \hat{\phi}_1(\frac{t_1}{n}, -k_1)\, \hat{\phi}_2(\frac{t_2}{n}, -k_2) \times \tag{57}$$
$$\times \Big[\mathrm{cov}(X^{(h_n)}_{[t_1+1/2+k_1/2],n}, X^{(h_n)}_{[t_2+1/2+k_2/2],n})\mathrm{cov}(X^{(h_n)}_{[t_1+1/2-k_1/2],n}, X^{(h_n)}_{[t_2+1/2-k_2/2],n})$$
$$+ \mathrm{cov}(X^{(h_n)}_{[t_1+1/2+k_1/2],n} X^{(h_n)}_{[t_2+1/2-k_2/2],n})\mathrm{cov}(X^{(h_n)}_{[t_1+1/2-k_1/2],n} X^{(h_n)}_{[t_2+1/2+k_2/2],n})$$
$$+ \mathrm{cum}(X^{(h_n)}_{[t_1+1/2+k_1/2],n}, X^{(h_n)}_{[t_1+1/2-k_1/2],n}, X^{(h_n)}_{[t_2+1/2+k_2/2],n}, X^{(h_n)}_{[t_2+1/2-k_2/2],n})\Big].$$

Let $k_3 := t_1 - t_2 + [k_1/2 + 1/2] - [k_2/2 + 1/2]$. By using Proposition 5.4 we replace the first summand in $[\ldots]$ by $h_n(\frac{t_1}{n})^4 c(\frac{t_1}{n}, k_3) c(\frac{t_1}{n}, k_3 + k_2 - k_1)$. The remainder can be bounded by

$$\frac{K}{n} \sum_{k_1,k_2,k_3} \Big[\tilde{\phi}_1(k_1)\tilde{\phi}_2(k_2)\Big\{1 + \frac{\min\{|k_1|, n\}}{\ell(k_3)}\Big\} \frac{1}{\ell(k_3+k_2-k_1)}$$
$$+ \tilde{\phi}_1(k_1)\tilde{\phi}_2(k_2) \frac{1}{\ell(k_3)} \Big\{1 + \frac{\min\{|k_1|, n\}}{\ell(k_3+k_2-k_1)}\Big\}\Big].$$

(52) implies that this is bounded as asserted. Therefore the first term is equal to

$$\frac{n}{(2\pi)^2 H_{2,n}^2} \sum_{t_1,k_1,k_2,k_3} h_n(\frac{t_1}{n})^4\, \hat{\phi}_1(\frac{t_1}{n}, -k_1)\, \hat{\phi}_2(\frac{t_1 + k_o}{n}, -k_2)$$
$$\times c(\frac{t_1}{n}, k_3)\, c(\frac{t_1}{n}, k_3 + k_2 - k_1) + R_n \tag{58}$$

where $k_o = -k_3 + [k_1/2 + 1/2] - [k_2/2 + 1/2]$. Replacing $\hat{\phi}_2(\frac{t_1+k_o}{n}, -k_2)$ by $\hat{\phi}_2(\frac{t_1}{n}, -k_2)$ yields the error term

$$\frac{K}{n} \sum_{k_1,k_2,k_3} \tilde{\phi}(k_1)\tilde{c}(k_3)\tilde{c}(k_3+k_2-k_1) \sum_t |\hat{\phi}_2(\frac{t+k_o}{n}, -k_2) - \hat{\phi}_2(\frac{t}{n}, -k_2)|$$

which also is bounded as asserted. As in (56) we now replace the $\frac{1}{n}\Sigma_{t_1}-$sum in (58) (with $k_o=0$) by the integral over [0,1] with the same replacement error. Furthermore, we replace h_n by h. Direct calculation (or repeated application of Parseval's equality) yields

$$\begin{aligned}&\frac{1}{(2\pi)^2}\int_0^1 \frac{h^4(u)}{\|h\|_2^4}\sum_{k_1,k_2,k_3}\hat{\phi}_1(u,-k_1)\,\hat{\phi}_2(u,-k_2)\,c(u,k_3)\,c(u,k_3+k_2-k_1)\,du\\ &\qquad = 2\pi\int_0^1 \frac{h^4(u)}{\|h\|_2^4}\int_{-\pi}^{\pi}\phi_1(u,\lambda)\phi_2(u,-\lambda)f(u,\lambda)^2 d\lambda du.\end{aligned}$$

The second term in (57) is treated in the same way. The third term is with the representation $X_{t,n}=\sum_{j=-\infty}^{\infty}a_{t,n}(t-j)\,\varepsilon_j$ and the abbreviations $t_\nu^+=t_\nu^+(t_\nu,k_\nu)=[t_\nu+1/2+k_\nu/2]$, $t_\nu^-=t_\nu^-(t_\nu,k_\nu)=[t_\nu+1/2-k_\nu/2]$ equal to

$$\begin{aligned}&\frac{n\,\kappa_4}{(2\pi)^2\,H_{2,n}^2}\sum_{t_1,t_2,\,k_1,\,k_2}\hat{\phi}_1(\frac{t_1}{n},-k_1)\,\hat{\phi}_2(\frac{t_2}{n},-k_2)\\ &\times\sum_i h_n(\frac{t_1^+}{n})h_n(\frac{t_1^-}{n})h_n(\frac{t_2^+}{n})h_n(\frac{t_2^-}{n})\,a_{t_1^+,n}(t_1^+-i)\,a_{t_1^-,n}(t_1^--i)\,a_{t_2^+,n}(t_2^+-i)\,a_{t_2^-,n}(t_2^--i).\end{aligned}$$

By using Definition 2.1 and (52) we now replace this by

$$\begin{aligned}&\frac{n\,\kappa_4}{(2\pi)^2\,H_{2,n}^2}\sum_{t_1,t_2,\,k_1,\,k_2}\hat{\phi}_1(\frac{t_1}{n},-k_1)\,\hat{\phi}_2(\frac{t_2}{n},-k_2)\\ &\qquad\times\sum_i h_n(\frac{t_1}{n})^2h_n(\frac{t_2}{n})^2\,a(\frac{t_1}{n},t_1^+-i)\,a(\frac{t_1}{n},t_1^--i)\,a(\frac{t_2}{n},t_2^+-i)\,a(\frac{t_2}{n},t_2^--i)\end{aligned}\tag{59}$$

with replacement error $Kn^{-1}\sum_{k_1,k_2}\tilde{\phi}_1(k_1)\tilde{\phi}_2(k_2)$. We now replace the term $a(\frac{t_2}{n},t_2^--i)$ in the above expression by $a(\frac{t_1}{n},t_2^--i)$ leading with the substitutions $d=t_2-t_1$ and $j=i-t_1$ to a replacement error of

$$\begin{aligned}&\frac{K}{n}\sum_{k_1,\,k_2}\tilde{\phi}_1(k_1)\tilde{\phi}_2(k_2)\\ &\qquad\times\sum_{d,j}\frac{1}{\ell\big([\frac{1}{2}+\frac{k_1}{2}]-j\big)}\frac{1}{\ell\big([\frac{1}{2}-\frac{k_1}{2}]-j\big)}\frac{1}{\ell\big([d+\frac{1}{2}+\frac{k_2}{2}]-j\big)}\\ &\qquad\times\sum_{t_1}\Big|a\big(\frac{t_1+d}{n},[d+\frac{1}{2}-\frac{k_2}{2}]-j\big)-a\big(\frac{t_1}{n},[d+\frac{1}{2}-\frac{k_2}{2}]-j\big)\Big|.\end{aligned}$$

The last sum is bounded by $|d|\,V(a(\cdot,[d+\frac{1}{2}-\frac{k_2}{2}]-j))$ leading to the upper bound

$$\begin{aligned}
&\frac{K}{n}\sum_{k_1,k_2}\tilde{\phi}_1(k_1)\tilde{\phi}_2(k_2)\\
&\quad\times\sum_{d,j}\frac{\big|[d+\frac{1}{2}+\frac{k_2}{2}]-j\big|+\big|[d+\frac{1}{2}-\frac{k_2}{2}]-j\big|+\big|[\frac{1}{2}+\frac{k_1}{2}]-j\big|+\big|[\frac{1}{2}-\frac{k_1}{2}]-j\big|}{\ell\big([\frac{1}{2}+\frac{k_1}{2}]-j\big)\,\ell\big([\frac{1}{2}-\frac{k_1}{2}]-j\big)\,\ell\big([d+\frac{1}{2}+\frac{k_2}{2}]-j\big)\,\ell\big([d+\frac{1}{2}-\frac{k_2}{2}]-j\big)}\\
&\le\frac{K}{n}\sum_{k_1,k_2}\tilde{\phi}_1(k_1)\tilde{\phi}_2(k_2)
\end{aligned}$$

for the replacement error. In the same way we replace (59) by

$$\begin{aligned}
&\frac{n\,\kappa_4}{(2\pi)^2\,H_{2,n}^2}\sum_{t_1,t_2,k_1,k_2}\hat{\phi}_1(\frac{t_1}{n},-k_1)\,\hat{\phi}_2(\frac{t_1}{n},-k_2) &(60)\\
&\qquad\times\sum_i h_n(\frac{t_1}{n})^4\,a(\frac{t_1}{n},t_1^+-i)\,a(\frac{t_1}{n},t_1^--i)\,a(\frac{t_1}{n},t_2^+-i)\,a(\frac{t_1}{n},t_2^--i)\\
&\quad=\frac{n\,\kappa_4}{(2\pi)^2\,H_{2,n}^2}\sum_{t_1}h_n(\frac{t_1}{n})^4\sum_{k_1,k_2}\hat{\phi}_1(\frac{t_1}{n},-k_1)\,\hat{\phi}_2(\frac{t_1}{n},-k_2)\,c(\frac{t_1}{n},k_1)\,c(\frac{t_1}{n},k_2) &(61)
\end{aligned}$$

with replacement error $Kn^{-1}\sum_{k_1,k_2}\tilde{\phi}_1(k_1)\big[\,\tilde{\phi}_2(k_2)+V(\hat{\phi}_2(\cdot,k_2))\big]\big[\frac{1}{\ell(k_1)}+\frac{1}{\ell(k_2)}\big]$. As in (56) we now replace the $\frac{1}{n}\Sigma_{t_1}$ – sum by the integral over [0,1] and h_n by h. Application of Parseval's equality gives the final form of the fourth order cumulant term.
(ii) Considering the cases $|k|\le\sqrt{n}$ and $|k|>\sqrt{n}$ separately shows that

$$\frac{1}{n}\sum_k\min\{|k|,n\}\tilde{\phi}_i(k)=o(1)\quad\text{ and }\quad\frac{1}{n}\sum_k\min\{|k|,n\}\frac{1}{\ell(k)}=o(1).\tag{62}$$

This implies that the first term of R_n tends to zero. Since

$$\begin{aligned}
&\min\{|k_1|+|k_2|+|k_3|,n\}\\
&\quad\le 2\min\{|k_1+k_2+k_3|,n\}+2\min\{|k_1\,|,n\}+2\min\{|k_3|,n\}
\end{aligned}$$

and $|V(\hat{\phi}_i(\cdot,k))|\le K$ also the third term of R_n tends to zero. □

Lemma 5.7 *Let $X_{t,n}$ be a locally stationary process. Suppose Assumption 5.1 holds and $\phi_1,\ldots,\phi_\ell:[0,1]\times[-\pi,\pi]\to\mathbf{R}$ are functions possibly depending on n.*

(i) If $\ell\ge 2$ then

$$|\mathrm{cum}\big(E_n(\phi_1),\ldots,E_n(\phi_\ell)\big)|\le Kn^{1-\ell/2}\rho_{2,n}(\phi_1)\,\rho_{2,n}(\phi_2)\prod_{j=3}^{\ell}\rho_\infty(\phi_j)$$

with a constant K independent of n.

(ii) If $\ell \geq 2$ then

$$\left|\mathrm{cum}\big(E_n(\phi_1),\dots,E_n(\phi_\ell)\big)\right| \leq K^\ell\,(2\ell)!\prod_{j=1}^{\ell}\rho_{2,n}(\phi_j). \tag{63}$$

with a constant K independent of n and ℓ.

(iii) If $\ell \geq 3$ and $\phi_1,\dots,\phi_\ell$ are independent of n and fulfill Assumption 5.2 then

$$\left|\mathrm{cum}\big(E_n(\phi_1),\dots,E_n(\phi_\ell)\big)\right| = o(1).$$

PROOF. (i) We have with (53)

$$\begin{aligned}
&\mathrm{cum}\big(E_n(\phi_1),\dots,E_n(\phi_\ell)\big) = n^{\ell/2}\,\mathrm{cum}\big(F_n(\phi_1),\dots,F_n(\phi_\ell)\big) = \\
&\quad = \frac{n^{\ell/2}}{(2\pi)^\ell\,H_{2,n}^\ell}\sum_{t_1,\dots,t_\ell}\sum_{k_1,\dots,k_\ell}\hat{\phi}_1\Big(\frac{t_1}{n},-k_1\Big)\cdots\hat{\phi}_\ell\Big(\frac{t_\ell}{n},-k_\ell\Big)\times \\
&\qquad \times\mathrm{cum}\big(X^{(h_n)}_{[t_1+1/2+k_1/2],n}X^{(h_n)}_{[t_1+1/2-k_1/2],n},\dots,X^{(h_n)}_{[t_\ell+1/2+k_\ell/2],n}X^{(h_n)}_{[t_\ell+1/2-k_\ell/2],n}\big).
\end{aligned}$$

We now use the representation $X_{t,n} = \sum_{j=-\infty}^{\infty} a_{t,n}(t-j)\,\varepsilon_j$ and obtain with the product theorem for cumulants (c.f. Brillinger, 1981, Theorem 2.3.2) and the abbreviations $t_\nu^+ = t_\nu^+(t_\nu,k_\nu) = [t_\nu+1/2+k_\nu/2]$, $t_\nu^- = t_\nu^-(t_\nu,k_\nu) = [t_\nu+1/2-k_\nu/2]$ that this is equal to

$$\begin{aligned}
&\frac{n^{\ell/2}}{(2\pi)^\ell\,H_{2,n}^\ell}\sum_{t_1,\dots,t_\ell}\sum_{k_1,\dots,k_\ell}\hat{\phi}_1\Big(\frac{t_1}{n},-k_1\Big)\cdots\hat{\phi}_\ell\Big(\frac{t_\ell}{n},-k_\ell\Big)\times \\
&\quad\times\sum_{i_1,\dots,i_\ell,\,j_1,\dots,j_\ell}\prod_{\nu=1}^{\ell}\Big[h_n\Big(\frac{t_\nu^+}{n}\Big)\,h_n\Big(\frac{t_\nu^-}{n}\Big)\,a_{t_\nu^+,n}(t_\nu^+-i_\nu)\,a_{t_\nu^-,n}(t_\nu^--j_\nu)\Big] \\
&\quad\times\sum_{\{P_1,\dots,P_m\}\,i.p.}\prod_{j=1}^{m}\mathrm{cum}\big(\varepsilon_s|s\in P_j\big)
\end{aligned}$$

where the last sum is over all indecomposable partitions (i.p.) $\{P_1,\dots,P_m\}$ of the table

$$\begin{matrix} i_1 & j_1 \\ \cdot & \\ \cdot & \\ \cdot & \\ i_\ell & j_\ell \end{matrix} \tag{64}$$

with $|P_\nu| \geq 2$ (since $\mathbf{E}X(t)$=0). Using the upper bound $\sup_t |a_{t,n}(j)| \leq \frac{K}{\ell(j)}$ gives

$$\begin{aligned}
&|\mathrm{cum}\big(E_n(\phi_1), \ldots, E_n(\phi_\ell)\big)| \\
&\qquad \leq K n^{-\ell/2} \sum_{t_1,\ldots,t_\ell} \sum_{k_1,\ldots,k_\ell} \Big|\hat{\phi}_1(\frac{t_1}{n}, -k_1) \cdots \hat{\phi}_\ell(\frac{t_\ell}{n}, -k_\ell)\Big| \times \\
&\qquad\quad \times \sum_{i_1,\ldots,i_\ell, j_1,\ldots,j_\ell} \prod_{\nu=1}^{\ell} \Big[\frac{1}{\ell(t_\nu^+ - i_\nu)} \frac{1}{\ell(t_\nu^- - j_\nu)}\Big] \\
&\qquad\quad \times \sum_{\{P_1,\ldots,P_m\}\, i.p.} \prod_{j=1}^{m} \big|\mathrm{cum}\big(\varepsilon_s | s \in P_j\big)\big|.
\end{aligned} \tag{65}$$

The Cauchy-Schwarz inequality (with squares of ϕ_1 and ϕ_2) now leads to the upper bound

$$\begin{aligned}
&K n^{-\ell/2} \Bigg\{ \sum_{t_1,\ldots,t_\ell} \sum_{k_1,\ldots,k_\ell} \hat{\phi}_1(\frac{t_1}{n}, -k_1)^2 \; \tilde{\phi}_3(k_3) \cdots \tilde{\phi}_\ell(k_\ell) \times \\
&\qquad \times \sum_{i_1,\ldots,i_\ell, j_1,\ldots,j_\ell} \prod_{\nu=1}^{\ell} \Big[\frac{1}{\ell(t_\nu^+ - i_\nu)} \frac{1}{\ell(t_\nu^- - j_\nu)}\Big] \\
&\qquad \times \sum_{\{P_1,\ldots,P_m\}\, i.p.} \prod_{j=1}^{m} \big|\mathrm{cum}\big(\varepsilon_s | s \in P_j\big)\big| \Bigg\}^{1/2} \Bigg\{ \text{ similar term } \Bigg\}^{1/2}
\end{aligned} \tag{66}$$

which by using (52) is bounded by

$$\begin{aligned}
&K n^{-\ell/2} \Bigg\{ \sum_{t_1} \sum_{k_1,\ldots,k_\ell} \hat{\phi}_1(\frac{t_1}{n}, -k_1)^2 \; \tilde{\phi}_3(k_3) \cdots \tilde{\phi}_\ell(k_\ell) \times \\
&\qquad \times \sum_{i_1,\ldots,i_\ell, j_1,\ldots,j_\ell} \frac{1}{\ell(t_1^+ - i_1)} \frac{1}{\ell(t_1^- - j_1)} \prod_{\nu=2}^{\ell} \frac{1}{\ell(k_\nu - i_\nu + j_\nu)} \\
&\qquad \times \sum_{\{P_1,\ldots,P_m\}\, i.p.} \prod_{j=1}^{m} \big|\mathrm{cum}\big(\varepsilon_s | s \in P_j\big)\big| \Bigg\}^{1/2} \Bigg\{ \text{ similar term } \Bigg\}^{1/2}.
\end{aligned} \tag{67}$$

Note that the term $\mathrm{cum}\big(\varepsilon_s | s \in P_j\big)$ leads to the restriction that all $i_\nu, j_\nu \in P_j$ are equal. We now sum over the remaining indices from $k_2, i_1, \ldots, i_\ell, j_1, \ldots, j_\ell$ leading due to the indecomposability of the partition and the fact that $1/\ell(j) \le K$ to the upper bound

$$
\begin{aligned}
&K n^{-\ell/2} \Big\{ \sum_{t_1} \sum_{k_1, k_3, \ldots, k_\ell} \hat{\phi}_1(\frac{t_1}{n}, -k_1)^2 \, \tilde{\phi}_3(k_3) \cdots \tilde{\phi}_\ell(k_\ell) \Big\}^{1/2} \Big\{ \text{ similar term } \Big\}^{1/2} \\
&\le K n^{1-\ell/2} \rho_{2,n}(\phi_1) \, \rho_{2,n}(\phi_2) \prod_{j=3}^{\ell} \rho_\infty(\phi_j)
\end{aligned}
$$

and therefore to the result.

In (ii) the generic constant K needs to be independent of ℓ. Again we have (65) (with K replaced by K^ℓ). Remember that the term $\mathrm{cum}\big(\varepsilon_s | s \in P_j\big)$ leads to the restriction that all $i_\nu, j_\nu \in P_j$ are equal. We denote this index by $i^{(j)}$ $(j = 1, \ldots, m)$.

We start by considering the case ℓ even. For each fixed partition $\{P_1, \ldots, P_m\}$ we can renumber the indices $\{1, \ldots, \ell\}$ in such a way that for every $j \in \{1, \ldots, m\}$ there exists at least one even $\nu \in \{1, \ldots, \ell\}$ and one odd $\nu \in \{1, \ldots, \ell\}$ such that $i^{(j)} = i_\nu$ or $i^{(j)} = j_\nu$. This can be derived from the indecomposability of the partition and $|P_k| \ge 2$ for all k. The Cauchy-Schwarz inequality now yields as an upper bound for (65)

$$
\begin{aligned}
K^\ell \, n^{-\ell/2} \sum_{\{P_1, \ldots, P_m\} i.p.} \Big\{ \sum_{t_1, \ldots, t_\ell} \sum_{k_1, \ldots, k_\ell} \prod_{j \text{ even}} \hat{\phi}_j(\frac{t_j}{n}, -k_j)^2 \sum_{i^{(1)}, \ldots, i^{(m)}} \prod_{\nu=1}^{\ell} \Big[\frac{1}{\ell(t_\nu^+ - i_\nu)} \, \frac{1}{\ell(t_\nu^- - j_\nu)} \Big] \Big\}^{1/2} \\
\times \Big\{ \text{ the same term with } \ldots \prod_{j \text{ odd}} \ldots \Big\}^{1/2}. \qquad (68)
\end{aligned}
$$

where $i_\nu = i^{(j)}$ if $i_\nu \in P_j$ and $j_\nu = i^{(j)}$ if $j_\nu \in P_j$. By using relation (52) we have

$$
\sum_{t_\nu, k_\nu} \frac{1}{\ell(t_\nu^+ - i_\nu)} \, \frac{1}{\ell(t_\nu^- - j_\nu)} \le K \sum_{k_\nu} \frac{1}{\ell(k_\nu - i_\nu + j_\nu)} \le K
$$

that is the first bracket in (68) is bounded by

$$
\begin{aligned}
&K^\ell \Big\{ \sum_{t_\nu, k_\nu;\, \nu \text{ even}} \prod_{\nu \text{ even}} \hat{\phi}_\nu(\frac{t_\nu}{n}, -k_\nu)^2 \sum_{i^{(1)}, \ldots, i^{(m)}} \prod_{\nu \text{ even}} \Big[\frac{1}{\ell(t_\nu^+ - i_\nu)} \, \frac{1}{\ell(t_\nu^- - j_\nu)} \Big] \Big\}^{1/2} \\
&\le K^\ell \, n^{\ell/4} \prod_{\nu \text{ even}} \rho_{2,n}(\phi_\nu).
\end{aligned}
$$

The same applies for the second bracket in (68). Since the number of indecomposable partitions is bounded by $4^\ell \, (2\ell)!$ we obtain (63).

The case ℓ odd is a bit more involved. For each partition $\{P_1,\dots,P_m\}$ with $m<\ell$ the result follows as in the case ℓ even. For $m=\ell$ we can renumerate the indices $\{1,\dots,\ell\}$ such that each P_ν contains exactly one element of $\{i_\nu,j_\nu\}$ and $\{i_{\nu+1},j_{\nu+1}\}$ (where $i_{\ell+1}=i_1, j_{\ell+1}=j_1$). For simplicity we treat the case where $P_\nu=\{i_\nu,j_{\nu+1}\}$ ($\nu=1,\dots,\ell$) (the other cases follow analogously). We obtain with the Cauchy-Schwarz inequality as an upper bound for (65)

$$\begin{aligned}
&K^\ell\, n^{-\ell/2} \sum_{t_\ell,k_\ell} \left|\hat{\phi}_\ell(\frac{t_\ell}{n},-k_\ell)\right| \times \\
&\qquad \times \Big\{ \sum_{t_1,\dots,t_{\ell-1}} \sum_{k_1,\dots,k_{\ell-1}} \prod_{j \text{ even}} \hat{\phi}_j(\frac{t_j}{n},-k_j)^2 \sum_{i_1,\dots,i_\ell} \prod_{\nu=1}^{\ell} \Big[\frac{1}{\ell(t_\nu^+ - i_\nu)}\, \frac{1}{\ell(t_\nu^- - i_{\nu-1})}\Big] \Big\}^{1/2} \\
&\qquad \times \Big\{ \text{ the same term with } \dots \prod_{j=1,3,\dots,\ell-2} \dots \Big\}^{1/2}
\end{aligned}$$

where $i_0=i_\ell$. By using relation (52) this is bounded by

$$\begin{aligned}
&K^\ell \Big\{\prod_{j=2}^{\ell-2}\rho_{2,n}(\phi_j)\Big\}\, n^{-3/2} \sum_{t_\ell,k_\ell} \left|\hat{\phi}_\ell(\frac{t_\ell}{n},-k_\ell)\right| \times \\
&\quad \times \Big\{ \sum_{t_{\ell-1},k_{\ell-1}} \hat{\phi}_{\ell-1}(\frac{t_{\ell-1}}{n},-k_{\ell-1})^2 \frac{1}{\ell(t_\ell^- - t_{\ell-1}^+)} \Big\}^{1/2} \Big\{ \sum_{t_1,k_1} \hat{\phi}_1(\frac{t_1}{n},-k_1)^2 \frac{1}{\ell(t_\ell^+ - t_1^-)} \Big\}^{1/2}.
\end{aligned}$$

The Cauchy-Schwarz inequality now yields

$$\begin{aligned}
&K^\ell \Big\{\prod_{j=2}^{\ell-2}\rho_{2,n}(\phi_j)\Big\}\, \rho_{2,n}(\phi_\ell) \times \\
&\quad \times\; n^{-1} \Big\{ \sum_{t_1,k_1,t_{\ell-1},k_{\ell-1}} \hat{\phi}_1(\frac{t_1}{n},-k_1)^2\, \hat{\phi}_{\ell-1}(\frac{t_{\ell-1}}{n},-k_{\ell-1})^2 \sum_{t_\ell,k_\ell} \frac{1}{\ell(t_\ell^- - t_{\ell-1}^+)} \frac{1}{\ell(t_\ell^+ - t_1^-)} \Big\}^{1/2} \\
&\quad \le K^\ell \Big\{\prod_{j=1}^{\ell}\rho_{2,n}(\phi_j)\Big\}
\end{aligned}$$

which finally leads to (ii).

(iii) follows from (i) since $\rho_{2,n}(\phi)^2 \le \rho_2(\phi)^2 + \frac{1}{n}\rho_\infty(\phi)\tilde{v}(\phi)$. □

5.8 (Proof of Theorem 2.4)
(The proof in the tapered case is exactly the same - compare Remark 2.5).
We have for each ϕ_j (denoted by ϕ for simplicity) and $k \neq 0$

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

where $\phi_R(u, d\lambda)$ is the signed measure corresponding to $\phi_R(u,\lambda) := \lim_{\mu \downarrow \lambda} \phi_R(u,\mu)$ (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 for $k \neq 0$

$$\tilde{\phi}(k) \leq \sup_u |\hat{\phi}(u,k)| \leq \frac{K}{|k|} \sup_u V\big(\phi(u,\cdot)\big) \quad \text{ and } \quad V\big(\hat{\phi}(\cdot,k)\big) \leq \frac{K}{|k|} V^2(\phi) \tag{69}$$

and for $k = 0$

$$\tilde{\phi}(0) = \sup_u |\hat{\phi}(u,0)| \leq 2\pi \, \|\phi\|_{\infty,\infty} \quad \text{ and } \quad V\big(\hat{\phi}(\cdot,0)\big) \leq 2\pi \sup_\lambda V\big(\phi(\cdot,\lambda)\big). \tag{70}$$

Thus $\rho_\infty(\phi)$ is not necessarily bounded and Theorem 5.3 cannot be applied. The trick now is to smooth $\phi(u,\lambda)$ 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

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

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

$$\widehat{\phi_n}(u,k) = \hat{\phi}(u,k)\, \widehat{k_b}(k) \tag{71}$$

with $\widehat{k_b}(k) = \exp(-k^2 b^2/2)$. We obtain from Lemma 5.7(i)

$$\begin{aligned} n \operatorname{var}\big[F_n(\phi_n) - F_n(\phi)\big] &\leq \frac{K}{n} \sum_{t=1}^{n} \sum_{k=-\infty}^{\infty} \big(\widehat{\phi_n}(\frac{t}{n},k) - \hat{\phi}(\frac{t}{n},k)\big)^2 \\ &\leq K \sum_k \frac{\big[\exp(-k^2b^2/2) - 1\big]^2}{k^2}. \end{aligned} \tag{72}$$

Since $|1 - \exp(-k^2b^2/2)| \leq \min\{1, \frac{k^2b^2}{2}\}$ this 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

$$\sqrt{n}\big(\{F_n(\phi_n) - \mathbf{E}F_n(\phi_n)\} - \{F_n(\phi) - \mathbf{E}F_n(\phi)\}\big) \xrightarrow{P} 0. \tag{73}$$

We now derive a CLT for $\sqrt{n}\big(F_n(\phi_{j,n}) - \mathbf{E}F_n(\phi_{j,n})\big)_{j=1\ldots,k}$ by applying Lemma 5.6(i) and Lemma 5.7(i). We obtain from (70) and (71)

$$\rho_\infty(\phi_n) \le 2\pi\|\phi\|_{\infty,\infty} + K\|\phi\|_{\infty,V}\sum_{k=1}^{\infty}\frac{1}{|k|}\exp(-k^2b^2/2) \le Kb^{-1} \quad \text{and} \quad v_\Sigma(\phi_n) \le Kb^{-1}.$$

Therefore the remainder term R_n in Lemma 5.6(i) and the higher cumulants in Lemma 5.7(i) converge to zero provided we choose b such that $bn^{(1-\kappa)/2} \to \infty$. Furthermore

$$c_E(\phi_{j,n},\phi_{k,n}) = c_E(\phi_j,\phi_k) + O(b^{1/2}) \qquad (i,j = 1,\ldots,k).$$

This follows easily by application of the Cauchy-Schwarz inequality and $\sup_u \int(\phi_{j,n}(u,\lambda) - \phi_j(u,\lambda))^2\,d\lambda = O(b)$ (obtained with the Parseval formula as in (72)) and $\int(\phi_{j,n}(u,\lambda))^2\,d\lambda \le K$. This gives the required CLT and with (73) also the CLT for $\sqrt{n}\big(F_n(\phi_j) - \mathbf{E}F_n(\phi_j)\big)_{j=1\ldots,k}$. We obtain from (55) and (56)

$$\begin{aligned}
&\sqrt{n}\big|\mathbf{E}F_n(\phi) - F(\phi)\big| \\
&\quad \le Kn^{-1/2}\Big[\sum_{|k|\le n}\tilde{\phi}(k) + \sum_{|k|>n}|k|\tilde{\phi}(k)\frac{1}{\ell(k)}\Big] + Kn^{-1/2+\kappa}\sum_k \frac{1}{\ell(k)} \\
&\quad \le Kn^{-1/2+\kappa}\log n
\end{aligned}$$

which finally proves Theorem 2.4. □

5.9 (Proof of Theorem 2.6 and Remark 2.7)

We obtain from Lemma 5.7(ii) for the ℓ-th order cumulant in the case $\ell \ge 2$

$$\big|\mathrm{cum}_\ell\left(E_n(\phi)\right)\big| \le K^\ell\,(2\ell)!\;\rho_{2,n}(\phi)^\ell.$$

The result now follows in the same way as in the proof of Lemma 2.3 in Dahlhaus (1988).

We now prove the inequalities in Remark 2.7: We obtain from the proof of Lemma 5.5 and an application of the Cauchy-Schwarz inequality

$$\begin{aligned}
&\sqrt{n}\,|\mathbf{E}F_n(\phi) - F^+(\phi)\,| \le \sqrt{n}\;\rho_{2,n}(\phi)\,\times \\
&\quad \times\Big(\frac{1}{n}\sum_{t=1}^{n}\Big\{\sum_{|k|\le n}\big[\mathrm{cov}(X^{(h_n)}_{[t+1/2+k/2],n},X^{(h_n)}_{[t+1/2-k/2],n}) - c(\frac{t}{n},k)\big]^2 + \sum_{|k|>n}c(\frac{t}{n},k)^2\Big\}\Big)^{1/2}.
\end{aligned}$$

Application of Proposition 5.4 yields that the term in the bracket is of order $n^{-1/2}$ leading to (16). (17) and (18) follow by straightforward calculations noting that Definition 2.1 implies $\sup_{u,\lambda} |f(u,\lambda)| \leq \infty$. (20) follows from an upper bound of (55) obtained by using (70). □

5.10 (Proof of Theorem 2.8) The proof uses a chaining technique as in Alexander (1984). Let

$$B_n = \{ \max_{t=1,\dots,n} |X_{t,n}| \leq 2 \log n \}. \tag{74}$$

Lemma 5.14 gives $\lim_{n\to\infty} P(B_n) = 1$. We will replace ϕ by

$$\phi_n^*(u,\lambda) = n \int_{u-\frac{1}{n}}^{u} \phi(v,\lambda)\, dv \qquad (\text{with } \phi(v,\lambda) = 0 \text{ for } v < 0). \tag{75}$$

The reason for doing so is that otherwise we needed the exponential inequality (15) to hold with $\rho_2(\phi)$ instead of $\rho_{2,n}(\phi)$. Such an inequality does not hold. Instead we exploit the following property of ϕ_n^*:

$$\begin{aligned} \rho_{2,n}^2(\phi_n^*) &= \frac{1}{n} \sum_{t=1}^{n} \int_{-\pi}^{\pi} \phi_n^*(\frac{t}{n},\lambda)^2 d\lambda = \frac{1}{n} \sum_{t=1}^{n} \int_{-\pi}^{\pi} \big(n \int_{\frac{t-1}{n}}^{\frac{t}{n}} \phi(u,\lambda)\, du\big)^2 d\lambda \\ &\leq \sum_{t=1}^{n} \int_{-\pi}^{\pi} \int_{\frac{t-1}{n}}^{\frac{t}{n}} \phi^2(u,\lambda)\, du\, d\lambda = \rho_2^2(\phi). \end{aligned} \tag{76}$$

Denote

$$\tilde{E}_n^*(\phi) := \tilde{E}_n(\phi_n^*) = \sqrt{n}\, (F_n - \mathbf{E} F_n)(\phi_n^*). \tag{77}$$

Since the assertion and the proof of Theorem 2.6 are for n fixed we obtain from (15)

$$P(|\tilde{E}_n^*(\phi)| \geq \eta) \leq c_1 \exp\Big(- c_2 \sqrt{\frac{\eta}{\rho_2(\phi)}}\, \Big). \tag{78}$$

On B_n we have by using Lemma 5.13 and Lemma 5.14 that

$$\begin{aligned} |\tilde{E}_n^*(\phi) - \tilde{E}_n(\phi)| \; &\leq \sqrt{n}\, |F_n(\phi_n^*) - F_n(\phi)| + \sqrt{n}\, |\mathbf{E} F_n(\phi_n^*) - \mathbf{E} F_n(\phi)| \\ &\leq 4\, K_1 \Big(\|\phi\|_{V,V} \frac{(\log n)^3}{n^{1/2}} + \|\phi\|_{V,\infty} \frac{(\log n)^2}{n^{1/2}} \Big) \\ &\quad + K_2\, \|\phi\|_{V,\infty} \frac{1}{n^{1/2}} \end{aligned}$$

and therefore with (22)

$$\sup_{\phi\in\Phi} |\tilde{E}_n^*(\phi) - \tilde{E}_n(\phi)| \leq \frac{\eta}{2}. \tag{79}$$

Thus

$$P(\sup_{\phi\in\Phi}|\tilde{E}_n(\phi)| > \eta,\, B_n\,) \le P(\sup_{\phi\in\Phi}|\tilde{E}^*_n(\phi)| > \frac{\eta}{2},\, B_n\,).$$

Let $\alpha := \tilde{H}_\Phi^{-1}\big(\frac{c_2}{4}\sqrt{\frac{\eta}{\tau_2}}\big)$. We obtain for any sequence $(\delta_j)_j$ with $\alpha = \delta_0 > \delta_1 > \ldots > 0$ where $\delta_{j+1} \le \delta_j/2$ with $\eta_{j+1} := \frac{9}{c_2^2}\,\delta_{j+1}\,\tilde{H}_\Phi(\delta_{j+1})^2$

$$\frac{\eta}{4} \ge \frac{18}{c_2^2}\int_0^\alpha \tilde{H}_\Phi(s)^2 ds \ge \frac{18}{c_2^2}\sum_{j=0}^\infty(\delta_{j+1}-\delta_{j+2})\,\tilde{H}_\Phi(\delta_{j+1})^2 \ge \sum_{j=0}^\infty \eta_{j+1}. \tag{80}$$

For each number δ_j choose a finite subset A_j corresponding to the definition of the covering numbers $N(\delta_j,\Phi,\rho_2)$. In other words, the set A_j consists of the smallest possible number $N_j = N(\delta_j,\Phi,\rho_2)$ of midpoints of ρ_2-balls of radius δ_j such that the corresponding balls cover Φ. Now telescope

$$\tilde{E}^*_n(\phi) = \tilde{E}^*_n(\phi_0) + \sum_{j=0}^\infty \tilde{E}^*_n(\phi_{j+1}-\phi_j) \tag{81}$$

where the ϕ_j are the approximating functions to ϕ from A_j, i.e. $\rho_2(\phi,\phi_j) < \delta_j$. The above equality holds on B_n, because Lemma 5.13 implies that

$$\sup_{\phi\in\Phi}|\tilde{E}^*_n(\phi-\phi_j)| \le 5\,K_3\,n\,(\log n)^2 \sup_{\phi\in\Phi}\rho_2(\phi-\phi_j) \le 5\,K_3\,n\,(\log n)^2\,\delta_j \to 0 \quad \text{for all } n.$$

Thus

$$\begin{aligned} &P(\sup_{\phi\in\Phi}|\tilde{E}^*_n(\phi)| > \frac{\eta}{2},\, B_n\,) \\ &\quad\le\; P(\sup_{\phi\in\Phi}|\tilde{E}^*_n(\phi_0)| > \frac{\eta}{4}\,) + \sum_{j=0}^\infty N_j N_{j+1} \sup_{\phi\in\Phi} P(\,|\tilde{E}^*_n(\phi_{j+1}-\phi_j)| > \eta_{j+1}) \\ &\quad=\; I + II. \end{aligned}$$

Hence, using exponential inequality (78) we have by definition of α that

$$I \le c_1\,\exp\Big\{\,\tilde{H}_\Phi(\alpha) - \frac{c_2}{2}\sqrt{\frac{\eta}{\tau_2}}\,\Big\} = c_1\,\exp\Big\{-\frac{c_2}{4}\sqrt{\frac{\eta}{\tau_2}}\,\Big\}. \tag{82}$$

In order to estimate II we need a particular definition of the δ_j. We set

$$\delta_{j+1} = \sup\left\{x \,:\, x \le \delta_j/2\,;\, \tilde{H}_\Phi(x) \ge \tilde{H}_\Phi(\delta_j) + \frac{1}{\sqrt{j+1}}\,\right\}.$$

Since $\sum_{\ell=1}^{j+1} \frac{1}{\sqrt{\ell}} \geq \int_0^{j+1} \frac{1}{\sqrt{x}}\, dx = 2\sqrt{j+1} \geq 2\log(j+1)$ we obtain for term II:

$$\begin{aligned}
II &\leq \sum_{j=0}^{\infty} c_1 \ \exp\Big\{ 2\, \tilde{H}_\Phi(\delta_{j+1}) - c_2 \sqrt{\frac{\frac{9}{c_2^2}\, \delta_{j+1}\, \tilde{H}_\Phi(\delta_{j+1})^2}{\delta_{j+1}}} \Big\} \\
&\leq \sum_{j=0}^{\infty} c_1 \ \exp\Big\{ -\tilde{H}(\delta_{j+1}) \Big\} \leq \sum_{j=0}^{\infty} c_1 \ \exp\Big\{ -\tilde{H}(\alpha) - 2\log(j+1) \Big\} \\
&= c_1 \ \exp\Big\{ -\frac{c_2}{4}\sqrt{\frac{\eta}{\tau_2}} \Big\} \sum_{j=1}^{\infty} \frac{1}{j^2} \ \leq 2\, c_1 \ \exp\Big\{ -\frac{c_2}{4}\sqrt{\frac{\eta}{\tau_2}} \Big\}.
\end{aligned}$$

This implies the maximal inequality (24). To prove (25) we note that $|E_n(\phi) - \tilde{E}_n(\phi)| = \sqrt{n}\,|\mathbf{E}F_n(\phi) - F(\phi)|$, that is we obtain instead of (79) on B_n with (83), (84), (17), (20) and (22)

$$\sup_{\phi \in \Phi} |\tilde{E}_n^*(\phi) - E_n(\phi)| \leq 13\, L \ \max\{\tau_{\infty,V}\,, \tau_{V,\infty}, \tau_{V,V}\,,\ \tau_{\infty,\infty}\}\, \frac{(\log n)^3}{\sqrt{n}} \leq \frac{\eta}{2}.$$

The rest of the proof is the same, i.e. we also obtain (25). □

5.11 (Proof of Theorem 2.10) To prove weak convergence of E_n we have to show weak convergence of the finite dimensional distributions and asymptotic equicontinuity in probability of E_n (cf. van der Vaart and Wellner, 1996, Theorems 1.5.4 and 1.5.7). Convergence of the finite dimensional distributions has been shown in Theorem 2.4. Asymptotic equicontinuity means that for every $\epsilon, \eta > 0$ there exists an $\tau_2 > 0$ such that

$$\liminf_n P\Big(\sup_{\rho_2(\phi,\psi) < \tau_2} |E_n(\phi - \psi)| > \eta \Big) < \epsilon.$$

In order to see this we apply Theorem 2.8. For fixed $\eta > 0$ there exists a $\tau_2 > 0$ small enough such that (23) holds. To see this notice that $\alpha \to 0$ as $\tau_2 \to 0$, and hence, using assumption (26), it follows that the integral on the right hand side of (23) also tends to zero if $\tau_2 \to 0$. As $\eta > 0$ is fixed (22) holds for n large enough. Hence we obtain with B_n from (74) for τ_2 small enough

$$\begin{aligned}
&\liminf_n P\Big(\sup_{\rho_2(\phi,\psi) < \tau_2} |E_n(\phi - \psi)| > \eta \Big) \\
&\quad \leq \liminf_n \ P\Big(\sup_{\rho_2(\phi,\psi) < \tau_2} |E_n(\phi - \psi)| > \eta,\ B_n \Big) + \lim_n P(B_n^c) \\
&\quad \leq 3\, c_1 \ \exp\Big\{ -\frac{c_2}{4}\sqrt{\frac{\eta}{\tau_2}} \ \Big\} < \epsilon.
\end{aligned}$$

□

5.12 (Proof of Theorem 2.11) Let $\delta > 0$. The assumptions of Theorem 2.8 are fulfilled for $\eta = \delta\sqrt{n}$ and n sufficiently large. For those n we obtain from Theorem 2.8

$$P\Big(\sup_{\phi\in\Phi_n} |F_n(\phi) - F(\phi)| > \delta\Big) = P\Big(\sup_{\phi\in\Phi_n} |E_n(\phi)| \;>\; \delta\sqrt{n}\,\Big)$$
$$\leq 3\,c_1\;\exp\Big\{-\frac{c_2}{4}\sqrt{\frac{\delta\sqrt{n}}{\tau_2^{(n)}}}\;\Big\} + P(B_n^c) \to 0.$$

□

Lemma 5.13 (Properties of $\mathbf{F_n}(\phi^*_\mathbf{n})$)
Let $X_{t,n}$ be a locally stationary process and $\phi^*_n(u,\lambda) = n\int_{u-\frac{1}{n}}^{u}\phi(v,\lambda)\,dv$ (with $\phi(v,\lambda)=0$ for $v<0$). Then we have with $X_{(n)} := \max_{t=1,\ldots,n}|X_{t,n}|$

$$|F_n(\phi) - F_n(\phi^*_n)| \leq K_1\,X^2_{(n)}\,\Big(\|\phi\|_{V,V}\,\frac{\log n}{n} + \|\phi\|_{V,\infty}\frac{1}{n}\Big)\,, \tag{83}$$

$$|\mathbf{E}F_n(\phi) - \mathbf{E}F_n(\phi^*_n)| \leq K_2\,\|\phi\|_{V,\infty}\,\frac{1}{n}\,, \tag{84}$$

$$|F_n(\phi^*_n) - \mathbf{E}F_n(\phi^*_n)| \leq K_3\Big(\sqrt{n}\,X^2_{(n)} + 1\Big)\;\rho_2(\phi). \tag{85}$$

PROOF. We have with (70)

$$\begin{aligned}
|F_n(\phi) - F_n(\phi^*_n)| = \;& \left|\frac{1}{n}\sum_{t=1}^{n}\int_{-\pi}^{\pi}\Big(\phi(\frac{t}{n},\lambda) - \phi^*_n(\frac{t}{n},\lambda)\Big)\;J_n(\frac{t}{n},\lambda)\,d\lambda\right| \\
&\leq O(X^2_{(n)})\sum_{t=1}^{n}\frac{1}{n}\sum_{k=-n}^{n} n\int_{\frac{t-1}{n}}^{\frac{t}{n}}\Big|\widehat{\phi}(\frac{t}{n},-k) - \widehat{\phi}(u,-k)\Big|\,d\,u \\
&\leq O(X^2_{(n)})\,\frac{1}{n}\sum_{k=-n}^{n} V(\hat{\phi}(\cdot,k)) \\
&\leq K_1\,X^2_{(n)}\,\Big(V^2(\phi)\,\frac{\log n}{n} + \sup_{\lambda} V\big(\phi(\cdot,\lambda)\big)\frac{1}{n}\Big).
\end{aligned}$$

Furthermore we obtain with Proposition 5.4 and (70)

$$\begin{aligned}
&|\mathbf{E}F_n(\phi) - \mathbf{E}F_n(\phi^*_n)| \\
&\quad\leq \frac{1}{2\pi}\frac{1}{n}\sum_{t=1}^{n}\sum_{k=-n}^{n} n\int_{\frac{t-1}{n}}^{\frac{t}{n}}\big|\widehat{\phi}(\frac{t}{n},-k) - \widehat{\phi}(u,-k)\big|du\;\big|\mathrm{cov}(X_{[t+1/2+k/2],n}, X_{[t+1/2-k/2],n}\big| \\
&\quad\leq K_2\,\sup_{\lambda} V\big(\phi(\cdot,\lambda)\big)\,\frac{1}{n}.
\end{aligned}$$

(85) has been proved in Lemma A.3 of Dahlhaus and Polonik (2006). □

Lemma 5.14 *Let $X_{t,n}$ be a locally stationary process with $E|\varepsilon_t|^k \leq C_\varepsilon^k$ for all $k \in \mathbf{N}$ for the ε_t from Definition 2.1 and let $X_{(n)} := \max_{t=1,\dots,n} |X_{t,n}|$. Then we have*

$$\mathbf{P}(X_{(n)} > 2\,\log n\,) \to 0.$$

PROOF. From (3) we have $\sup_{t,n} \sum_{j=-\infty}^{\infty} |a_{t,n}(j)| < m_0 < \infty$. The monotone convergence theorem and Jensen's inequality then imply

$$\mathbf{E}\,|X_{t,n}|^k \leq \mathbf{E}\Big(\sum_{j=-\infty}^{\infty} |a_{t,n}(j)|\,|\varepsilon_{t-j}|\Big)^k \leq m_0^k\,\mathbf{E}\Big(\sum_{j=-\infty}^{\infty} \frac{|a_{t,n}(j)|}{m_0}\,|\varepsilon_{t-j}|^k\Big) \leq m_0^k\,C_\varepsilon^k$$

leading to

$$\begin{aligned}\mathbf{P}(X_{(n)} > 2\,\log n\,) &\leq n \max_{t=1,\dots,n} \mathbf{P}(X_{t,n} > 2\,\log n\,)\\ &\leq n\frac{\mathbf{E}\,e^{|X_{t,n}|}}{e^{2\,\log n}} \leq \frac{1}{n}\sum_{k=0}^{\infty}\frac{m_0^k\,C_\varepsilon^k}{k!} \leq \frac{1}{n}\,e^{m_0 C_\varepsilon} \to 0.\end{aligned}$$

□

6 Appendix: Proof of Proposition 2.3

6.1 (Proof of Proposition 2.3) We only give the proof for tvAR-processes ($q = 0$). The extension to tvARMA-processes then is straightforward. The proof is similar to Künsch (1995) who proved the existence of a solution of the form (2) under the assumption that the functions $\alpha_i(u)$ are continuous. Let

$$\boldsymbol{\alpha}(u) = \begin{pmatrix} -\alpha_1(u) & -\alpha_2(u) & \dots & \dots & -\alpha_p(u)\\ 1 & 0 & \dots & \dots & 0\\ 0 & \ddots & \ddots & & \vdots\\ \vdots & \ddots & \ddots & \ddots & \vdots\\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}$$

and $\boldsymbol{\alpha}(u) = \boldsymbol{\alpha}(0)$ for $u < 0$. Since $\det(\lambda E_p - \boldsymbol{\alpha}(u)) = \lambda^p(\Sigma_{j=0}^p \alpha_j(u)\lambda^{-j})$ it follows that $\delta(\boldsymbol{\alpha}(u)) \leq \frac{1}{1+\delta}$ for all u where $\delta(A) := \max\{|\lambda| : \lambda \text{ eigenvalue of } A\}$. Let

$$a_{t,n}(j) = \Big(\prod_{\ell=0}^{j-1} \boldsymbol{\alpha}(\frac{t-\ell}{n})\Big)_{11}\sigma(\frac{t-j}{n})$$

and

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

It is easy to check that $X_{t,n}$ is a solution of (8) provided the coefficients are absolutely summable.

To prove this we note (cf. Householder, 1964, p.46) that for every $\varepsilon > 0$ and $u \in [0,1]$ there exists a matrix $M(u)$ with

$$\|\boldsymbol{\alpha}(u)\|_{M(u)} \leq \delta\big(\boldsymbol{\alpha}(u)\big) + \varepsilon$$

where $\|A\|_M := \sup\{\|Ax\|_M : \|x\|_M = 1\}$ and $\|x\|_M = \|M^{-1}x\|_1 = \sum_{i=1}^{p} |(M^{-1}x)_i|$. Since the $\alpha_i(u)$ are functions of bounded variation (i.e. the difference of two monotonic functions) there exists for all $\varepsilon > 0$ a finite partition of intervals $I_1 \cup \dots \cup I_m = [0,1]$ such that $|\alpha_i(u) - \alpha_i(v)| < \varepsilon$ for all i whenever u, v are in the same I_k. Let $M_k := M(u_k)$ for an arbitrary $u_k \in I_k$. Therefore m (and the partition) can be chosen such that

$$\|\boldsymbol{\alpha}(v)\|_{M_k} \leq \rho := \big(1 + \frac{\delta}{2}\big)^{-1} < 1 \quad \text{for all } v \in I_k.$$

We now replace the first interval I_1 by $I_1 \cup (-\infty, 0)$ (remember that $\boldsymbol{\alpha}(u) = \boldsymbol{\alpha}(0)$ for $u < 0$). There exists a constant c_0 such that $\|B\|_1 := \sum_{i,j} |B_{i,j}| \leq c_0 \|B\|_{M_k}$ for all k. For t and n fixed we now define $L_k := \{\ell \geq 0 : \frac{t-\ell}{n} \in I_k\}$ and $L_{k,j} := L_k \cap \{0, \dots, j-1\}$. Then

$$\begin{aligned}
|a_{t,n}(j)| \; &= \; \Big|\Big(\prod_{\ell=0}^{j-1} \boldsymbol{\alpha}\big(\frac{t-\ell}{n}\big)\Big)_{11} \sigma\big(\frac{t-j}{n}\big)\Big| \; \leq \; \Big\|\prod_{\ell=0}^{j-1} \boldsymbol{\alpha}\big(\frac{t-\ell}{n}\big)\Big\|_1 \sigma\big(\frac{t-j}{n}\big) \\
&\leq \prod_{k=1}^{m} \; \Big\|\prod_{\ell \in L_{k,j}} \boldsymbol{\alpha}\big(\frac{t-\ell}{n}\big)\Big\|_1 \sigma\big(\frac{t-j}{n}\big) \; \leq \; c_0^m \prod_{k=1}^{m} \; \Big\|\prod_{\ell \in L_{k,j}} \boldsymbol{\alpha}\big(\frac{t-\ell}{n}\big)\Big\|_{M_k} \sigma\big(\frac{t-j}{n}\big) \\
&\leq \; c_0^m \sup_u \sigma(u) \prod_{k=1}^{m} \rho^{|L_{k,j}|} \; = \; K\rho^j \quad \text{(since m is fixed)},
\end{aligned}$$

that is we have proved (3). Since $\|\boldsymbol{\alpha}(\frac{t-k}{n}) - \boldsymbol{\alpha}(\frac{t}{n})\|_1 = \sum_{i=1}^{p} |\alpha_i(\frac{t-k}{n}) - \alpha_i(\frac{t}{n})|$ we obtain with similar arguments and

$$a(u,j) := (\boldsymbol{\alpha}(u)^j)_{11} \sigma(u)$$

$$\sum_{t=1}^{n} |a_{t,n}(j) - a(\frac{t}{n}, j)| \leq \sum_{t=1}^{n}\sum_{k=1}^{j-1} \Big\| \boldsymbol{\alpha}(\frac{t}{n})^k \Big(\boldsymbol{\alpha}(\frac{t-k}{n}) - \boldsymbol{\alpha}(\frac{t}{n})\Big) \prod_{\ell=k+1}^{j-1} \boldsymbol{\alpha}(\frac{t-\ell}{n}) \Big\|_1 \sigma(\frac{t-j}{n})$$

$$+ \sum_{t=1}^{n} \Big\| \boldsymbol{\alpha}(\frac{t}{n})^j \Big\|_1 \Big| \sigma(\frac{t-j}{n}) - \sigma(\frac{t}{n}) \Big|$$

$$\leq \sum_{t=1}^{n}\sum_{k=1}^{j-1} c_0 \rho^k \sum_{i=1}^{p} \Big| \alpha_i(\frac{t-k}{n}) - \alpha_i(\frac{t}{n}) \Big| \, c_0^m \rho^{j-1-k}$$

$$\leq K j^2 \rho^{j-1}$$

i.e. (5). (4) and (6) follow similarly. □

Acknowledgement. This work has been partially supported by the NSF (grants #0103606 and #0406431) and the Deutsche Forschungsgemeinschaft (DA 187/15-1).

References

Brillinger, D.R. (1981). *Time Series: Data Analysis and Theory.* Holden Day, San Francisco.

Dahlhaus, R. (1988). Empirical spectral processes and their applications to time series analysis. *Stoch. Proc. Appl.* **30** 69-83.

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

Dahlhaus, R. (2000). A likelihood approximation for locally stationary processes. *Ann. Statist.* **28** 1762-1794.

Dahlhaus, R. and Neumann, M.H. (2001). Locally adaptive fitting of semiparametric models to nonstationary time series. *Stoch. Proc. and Appl.* **91** 277-308.

Dahlhaus, R. and Polonik, W. (2002). Empirical spectral processes and nonparametric maximum likelihood estimation for time series. In *Empirical Process Techniques for Dependent Data*, H. Dehling, T. Mikosch, and M. Sørensen, editors, pp. 275 - 298.

Dahlhaus, R. and Polonik, W. (2006). Nonparametric quasi maximum likelihood estimation for Gaussian locally stationary processes. *Ann. Statist.* to appear.

Davis, R.A. and Lee, T., and Rodriguez-Yam, G. (2005). Structural Break Estimation for Nonstationary Time Series Models. *J. Am. Statist. Ass.* **101** 223 - 239.

Fay, G. and Soulier, P. (2001) The periodogram of an i.i.d. sequence. *Stochastic Processes and Their Applications* **92** 315 - 343.

Fryzlewicz, P., Sapatinas, T. and Subba Rao, S. (2006). A Haar-Fisz technique for locally stationary volatility estimation. *Biometrika* **93** 687-704.

Künsch, H.R. (1995). A note on causal solutions for locally stationary AR processes. Preprint. ETH Zürich.

Nason, G. P.,von Sachs, R. and Kroisandt, G. (2000). Wavelet processes and adaptive estimation of evolutionary wavelet spectra. *J Royal Statist. Soc. B* **62** 271-292.

Mikosch, T. and Norvaisa, R. (1997). Uniform convergence of the empirical spectral distribution function. *Stoch. Proc. Appl.* **70** 85 - 114.

Moulines, E., Priouret, P. and Roueff, F. (2005). On recursive estimation for locally stationary time varying autoregressive processes. *Ann. Statist.* **33** 2610-2654.

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

Priestley, M.B. (1965). Evolutionary spectra and non-stationary processes. *J Royal Statist. Soc. B* **27** 204-237.

Sakiyama, K. and Taniguchi, M. (2004). Discriminant analysis for locally stationary processes. *J. Multiv. Anal.* **90** 282-300.

Van Bellegem, S. and Dahlhaus, R. (2006). Semiparametric estimation by model selection for locally stationary processes. *J Royal Statist. Soc. B* **68** 721-746.

van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes.* Springer, New York

Whittle, P. (1953) Estimation and information in stationary time series. *Ark. Mat* **2** 423-434.