Institut für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Spectral density estimation via nonlinear wavelet methods for stationary non–Gaussian time series

Michael H. Neumann

submitted: 9th May 1994

Institute of Applied Analysis
and Stochastics
Mohrenstraße 39
D – 10117 Berlin
Germany

Preprint No. 99
Berlin 1994

1991 Mathematics Subject Classification. Primary 62M15, Secondary 62M10, 62G07.
Key words and phrases. Spectral density estimation, wavelet estimators, asymptotic normality, large deviations.

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

Fax: + 49 30 2004975
e–mail (X.400): c=de;a=d400;p=iaas-berlin;s=preprint
e–mail (Internet): preprint@iaas-berlin.d400.de

Spectral Density Estimation via Nonlinear Wavelet Methods for Stationary non-Gaussian Time Series

M. H. Neumann
Institute of Applied Analysis and Stochastics
Mohrenstr. 39
Berlin, 10117, Germany

Abstract. In the present paper we consider nonlinear wavelet estimators of the spectral density f of a zero mean stochastic process, which is stationary in the wide sense. It is known in the case of Gaussian regression that these estimators outperform traditional linear methods if the degree of smoothness of the regression function varies considerably over the interval of interest. Such methods are based on a nonlinear treatment of estimators of coefficients that arise from a Fourier series expansion according to a wavelet basis.
The main goal of this paper is to prepare the ground for the application of these methods to spectral density estimation, which is done by showing the asymptotic normality of certain empirical coefficients based on the tapered periodogram. For that we derive upper estimates for their cumulants, which yield the asymptotic normality in terms of probabilities of large deviations. Using these results we can conclude the risk equivalence to the Gaussian case for monotone estimators based on such empirical coefficients. Hence, we obtain estimators of f, which keep all interesting properties like high spatial adaptivity that are already known from wavelet estimators in the case of Gaussian regression.
It turns out that optimally tuned versions of these estimators attain the optimal uniform rate of convergence of their L_2-risk in a wide variety of Besov smoothness classes.

1991 Mathematics Subject Classification. Primary 62M15; Secondary 62M10, 62G07.
Key words and phrases. Spectral density estimation, wavelet estimators, asymptotic normality, large deviations.

1. Introduction and main results

In the present paper we consider the problem of estimating a spectral density of a real, zero mean stochastic process, which is stationary in the wide sense. An important goal in spectral analysis is the recognition of peaks of the spectral density, which are an indication for periodicities of the underlying stochastic process.

Whereas globally homogeneous smoothing methods (e.g. kernel estimators with global bandwidth, spline estimators) are an appropriate tool for estimating functions with a homogeneous degree of smoothness, one should prefer methods with a locally varying degree of smoothing in estimating objects that are quite smooth on one part of the domain but much less regular on another part.

Such a problem was recently investigated by Donoho and Johnstone (1992), who derived minimax results in Besov classes $B^m_{p,q}$ in the Gaussian white noise model

$$dX_t = f(t)\,dt + \epsilon\,dW_t, \quad t \in [0,1].$$

They rediscovered the phenomenon, originally detected by Nemirovskii, Tsybakov and Polyak (1985) and Nemirovskii (1985) for Sobolev smoothness classes, that linear estimators are unable to attain the optimal uniform convergence rate in balls of Besov spaces $B^m_{p,q}$ with $p < 2$. Moreover, they showed that thresholded wavelet estimators attain the minimax bound up to a small constant.

These estimators are based on an orthonormal system $\{\tilde{\phi}_{l,k}\}_{k\in I_l^0} \cup \{\tilde{\psi}_{j,k}\}_{j\geq l, k\in I_j}$ of basis functions, which are essentially generated by dilations and translations of two special functions ϕ and ψ. Having the Fourier series expansion $f = \sum \alpha_k \tilde{\phi}_{l,k} + \sum \alpha_{j,k}\tilde{\psi}_{j,k}$ in mind, one determines first empirical versions $\tilde{\alpha}_k$ and $\tilde{\alpha}_{j,k}$ of the Fourier coefficients. These empirical coefficients are again normally distributed with homogeneous variance ϵ^2. Then one applies level-wise nonlinear shrinkage rules δ_j to these coefficients, which finally yields an estimator

$$\hat{f} = \sum \tilde{\alpha}_k \tilde{\phi}_{l,k} + \sum \delta_j(\tilde{\alpha}_{j,k})\tilde{\psi}_{j,k}$$

of the regression function f.

We intend to transfer this locally adaptive estimation technique to spectral density estimation. For that we define a similar wavelet basis, here for $\tilde{L}_2(\Pi)$, the collection of 2π-periodic functions of $L_2(\Pi)$, $\Pi = [-\pi, \pi]$. On the basis of a, possibly tapered, periodogram we obtain empirical versions of the Fourier coefficients of f, which are then treated with the same shrinkage methods as known from Donoho and Johnstone (1992) in the case of Gaussian regression.

Our aim is to show that the resulting estimator of the spectral density behaves as good as a corresponding estimator in the Gaussian white noise model. For that we derive uniform estimates of the cumulants of the empirical wavelet coefficients, which are the basis for showing the equivalence to the Gaussian case in terms of probabilities of large deviations. These strong results allow us to conclude the risk equivalence between all monotone estimators based on the empirical coefficients and the same estimators with Gaussian random variables inserted. Therefore, thresholded wavelet

estimators keep all their appealing properties known from the Gaussian case. Finally, we discuss briefly a simple possibility to adapt the smoothing parameters involved in the procedure. It is shown that optimally thresholded wavelet estimators attain the minimax rate of convergence in a large scale of Besov smoothness classes.
We remark that Gao (1993) obtained related results for the case of a Gaussian time series. He found estimates for the L_2-risk of thresholded wavelet estimators not via an appropriate asymptotic normal approximation, but he regarded the empirical coefficients as quadratic forms of independent Gaussian random variables and used estimates for tail probabilities of them.

2. Basic tools

We assume that we observe a stretch $\{X_1, \dots, X_T\}$ of a real zero mean process $\{X_t\}_{t=1,\dots,\infty}$, which is stationary in the wide sense. We intend to estimate the spectral density

$$f(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} c_k \exp(i\omega k) = \frac{1}{2\pi} c_0 + \frac{1}{\pi} \sum_{k=1}^{\infty} c_k \cos(\omega k) \tag{2.1}$$

on the interval $\Pi = [-\pi, \pi]$, where $c_k = cov(X_t, X_{t+k})$.
The basis for our estimates will be a tapered periodogram

$$I_T(\omega) = \frac{1}{2\pi H_2^{(T)}} \left| \sum_{t=1}^{T} h_t X_t e^{it\omega} \right|^2 = \frac{1}{2\pi H_2^{(T)}} \sum_{s,t=1}^{T} h_s h_t X_s X_t \cos(\omega(t-s)), \tag{2.2}$$

where $H_k^{(T)} = \sum_{t=1}^{T} h_t^k$, $h_t = h(t/T)$. (We put $I_T(\omega) \equiv 0$ if $H_2^{(T)} = 0$.)
For the data taper we assume

(A1) $H = \int_0^1 h^2(x)\,dx > 0$, h is of bounded variation.

In particular, we obtain then $H_2^{(T)} \sim TH$. Choosing $h \equiv 1$ we obtain in (2.2) the ordinary (nontapered) periodogram I_T^0 as a special case. For the sake of greater generality we consider throughout the paper the tapered one, keeping in mind that this includes the nontapered analog as a special case.
It is well-known that $I_T(\omega)$ is asymptotically unbiased for $f(\omega)$ under quite general assumptions, however in many instances it is not a consistent estimator of $f(\omega)$. On the other hand, for $\omega_1 \neq \omega_2$, $I_T(\omega_1)$ and $I_T(\omega_2)$ are asymptotically uncorrelated. Therefore, there is some hope that one can obtain via smoothing of the periodogram estimators that are consistent under certain smoothness conditions on f.
Since the object of interest is 2π-periodic, we do not need any boundary correction of the wavelet basis. We start with two special compactly supported functions ϕ and ψ with the property that the collection

$$\left\{\tilde{\phi}_{l,k}\right\}_{k\in\mathbb{Z}} \cup \left\{\tilde{\psi}_{j,k}\right\}_{j\geq l, k\in\mathbb{Z}}$$

with $\tilde{\phi}_{l,k}(x) = 2^{l/2}\phi(2^l x - k)$ and $\tilde{\psi}_{j,k}(x) = 2^{j/2}\psi(2^j x - k)$ forms an orthonormal basis of $L_2(\mathbb{R})$. It is easy to see that

$$\{\phi_{l,k}\}_{k\in I_l} \cup \{\psi_{j,k}\}_{j\geq l, k\in I_j}$$

with $I_j = \{1, \ldots, 2^j\}$,

$$\phi_{l,k}(x) = \sum_{n\in\mathbb{Z}} (2\pi)^{-1/2} \tilde{\phi}_{l,k}\left((2\pi)^{-1}(x+n)\right)$$

and

$$\psi_{j,k}(x) = \sum_{n\in\mathbb{Z}} (2\pi)^{-1/2} \tilde{\psi}_{j,k}\left((2\pi)^{-1}(x+n)\right)$$

is an orthonormal basis of $\tilde{L}_2(\Pi)$, the L_2-space of 2π-periodic functions on Π. For $f \in \tilde{L}_2(\Pi)$ we have the representation

$$f = \sum_{k\in I_l} \alpha_k \phi_{l,k} + \sum_{j\geq l} \sum_{k\in I_j} \alpha_{j,k} \psi_{j,k},$$

where $\alpha_k = \int f(t)\phi_{l,k}(t)\, dt$ and $\alpha_{j,k} = \int f(t)\psi_{j,k}(t)\, dt$ are the usual Fourier coefficients, which are also called wavelet coefficients. Gao (1993) used another version of periodized wavelet functions.
Now we are in position to define empirical versions of these coefficients, which will be the starting point for our estimation procedures. First, we define an integral version as

$$\tilde{\alpha}_{j,k} = \int \psi_{j,k}(\omega) I_T(\omega)\, d\omega. \tag{2.3}$$

($\tilde{\alpha}_k$ is defined analogously.)
It is also possible to define discrete versions, e.g.

$$\overline{\alpha}_{j,k} = \sum_{i=1}^{T} \int_{\omega_i - 2\pi/T}^{\omega_i} \psi_{j,k}(\omega)\, d\omega\ I_T(\omega_i) \tag{2.4}$$

and

$$\overline{\overline{\alpha}}_{j,k} = \frac{2\pi}{T} \sum_{i=1}^{T} \psi_{j,k}(\omega_i)\, I_T(\omega_i), \tag{2.5}$$

where $\omega_i = 2\pi i/T - \pi$, $i = 1, \ldots, T$. ($\overline{\alpha}_k$ and $\overline{\overline{\alpha}}_k$ are defined in the same way.)
One advantage of (2.4) and (2.5) in the nontapered case $h \equiv 1$ is undoubtly that, because of $\sum_{t=1}^{T} \exp\{it\omega_j\} = 0$ for $\omega_j = 2\pi j/T - \pi$, $0 < j < T$, the equation

$$I_T^0(\omega_j) = \frac{1}{2\pi T} \left| \sum_{t=1}^{T} X_t e^{it\omega_j} \right| = \frac{1}{2\pi T} \left| \sum_{t=1}^{T} \left(X_t - T^{-1} \sum X_s \right) e^{it\omega_j} \right|, \tag{2.6}$$

holds, which means that the discrete version is robust against a deviation from our assumption $E\, X_t = 0$, since any nonzero mean of the X_t's is cancelled out. Moreover, it allows the computationally advantageous Fast Fourier Transform, see e.g. Brillinger (1975). On the other hand, we feel that $\tilde{\alpha}_{j,k}$ causes perhaps less bias problems. In the following we consider first estimators based on $\tilde{\alpha}_{j,k}$, but we show in Section 6 that the cumulants of $\overline{\alpha}_{j,k}$ and $\overline{\overline{\alpha}}_{j,k}$ are very close to those of $\tilde{\alpha}_{j,k}$, which means that the corresponding estimators have the same asymptotic behavior.

3. Cumulants of the empirical wavelet coefficients

In this section we develop approximations to the mean and the variance of $\tilde{\alpha}_{j,k}$, and we derive upper estimates for the higher order cumulants. It will turn out that the empirical wavelet coefficients are asymptotically normal if their variances are not too small. Without this restriction we obtain that their convolution with a Gaussian variable yields an asymptotically Gaussian random variable. In all of these cases it will turn out that the wavelet estimators considered in the present paper do not behave worse than in an idealized situation, where the empirical wavelet coefficients are Gaussian.

In the sequel we impose the following assumptions.

(A2) $TV(f) \le C_1$

(A3) $\sup_{1<t_1<\infty} \left\{ \sum_{t_2,\dots,t_k=1}^{\infty} |cum(X_{t_1},\dots,X_{t_k})| \right\} \le C_2^k (k!)^{1+\gamma}$ for all $k = 2,3,\dots$, where $\gamma \ge 0$.

Remark 1. If $\{X_t\}$ is α-mixing with coefficients $\alpha(s,t) \le K\exp(-b|s-t|)$ and $E|X_t|^k \le C^k (k!)^\gamma$ for all k, then

$$\sup_{1<t_1<\infty} \left\{ \sum_{t_2,\dots,t_k=1}^{\infty} |cum(X_{t_1},\dots,X_{t_k})| \right\} \le C_2^k (k!)^{3+\gamma} \quad \text{for all } k = 2,3,\dots$$

Proposition 3.1. *Assume (A1), (A2) and (A3). Then*

(i) $E\tilde{\alpha}_{j,k} = \alpha_{j,k} + O\left(2^{j/2} T^{-1} \log T\right).$

(ii) $var(\tilde{\alpha}_{j,k}) = 2\pi \left(H_4^{(T)}/(H_2^{(T)})^2\right) \int_\Pi \psi_{j,k}(\alpha)\left[\psi_{j,k}(\alpha) + \psi_{j,k}(-\alpha)\right] |f(\alpha)|^2 \, d\alpha$
$+ o(T^{-1}) + O(T^{-1}2^{-j})$
holds uniformly over $k \in I_j$ *and* $2^j \le CT^{1-\alpha}$.

(iii) *Let* $a_{j,k}(l) = \int_\Pi \psi_{j,k}(t)\cos(lt)\,dt$ *and* $M = \sum_{l=-(T-1)}^{T-1} |a_{j,k}(l)|$. *Then*

$$|cum_n(\tilde{\alpha}_{j,k})| \le var(\tilde{\alpha}_{j,k})(n-1)! \left(4\pi \|\psi_{j,k}\|_\infty \|f\|_\infty \frac{\sup\{|h(t)|^2\}}{H_2^{(T)}}\right)^{n-2}$$
$$+ \frac{2^{n-2}C_2^{2n}}{\pi^n}((2n)!)^{1+\gamma} \left(\frac{\sup\{|h(t)|^2\}}{H_2^{(T)}}\right)^n \sup_{l=0,\pm1,\dots,\pm T-1} \{|a_{j,k}(l)|^2\} T M^{n-2}.$$

Lemma 3.1. *It holds*

(i) $a_{j,k}(l) = O\left((2^{j/2} l^{-1}) \wedge 2^{-j/2}\right)$,

(ii) $M = O\left(2^{j/2} \log T\right)$.

4. Asymptotic normality of the empirical wavelet coefficients

In this section we intend to state the asymptotic normality of the empirical coefficients $\tilde{\alpha}_{j,k}$. It is known (cf. Theorems 5.2.7 in Brillinger (1975)) that a value $I_T(\omega)$ of the periodogram is asymptotically chi-squared distributed. Hence, we can only expect asymptotic normality of $\tilde{\alpha}_{j,k}$, if a certain summation effect works. Therefore, we restrict our considerations in this section to coefficients with an index (j,k) from a set

$$\mathcal{J} = \mathcal{J}(T) = \left\{(j,k) \,\middle|\, 2^j \leq CT^{1-\alpha},\ k \in I_j\right\},$$

where $C < \infty$ and $\alpha > 0$ are arbitrary, but fixed constants. Let $\sigma_{j,k}^2$ denote the variance of $\tilde{\alpha}_{j,k}$.
By (ii) of Proposition 3.1 we obtain that

$$\sup_{j,k}\{\sigma_{j,k}\} = O(T^{-1/2}). \tag{4.1}$$

Let, for some fixed $C_0 > 0$,

$$\mathcal{J}^0 = \mathcal{J}^0(T) = \{(j,k) \in \mathcal{J} \,|\, \sigma_{j,k} \geq C_0 T^{-1/2}\}.$$

On the basis of the estimates given in Proposition 3.1 we can derive by Lemma 1 in Rudzkis, Saulis and Statulevicius (1978) the following assertion.

Theorem 4.1. *Assume (A1) through (A3). Then*

$$\frac{P\left(\pm(\tilde{\alpha}_{j,k} - \alpha_{j,k})/\sigma_{j,k} \geq x\right)}{1 - \Phi(x)} \longrightarrow 1$$

holds uniformly in $(j,k) \in \mathcal{J}^0$, $-\infty < x \leq \Delta_\gamma$, *where* $\Delta_\gamma = o(\Delta^{1/(3+4\gamma)})$ *and* $\Delta = T^{1/2} 2^{-j/2} (\log T)^{-1}$.

Note that we have those empirical coefficients excluded from the assertion of Theorem 4.1, which have a standard deviation $\sigma_{j,k}$ below the level $C_0 T^{-1/2}$. Since our cumulant estimates in (iii) of Proposition 3.1 are essentially all of the same order of magnitude, it is not possible to state the large deviation property by results from Rudzkis, Saulis and Statulevicius (1978) for empirical coefficients with a too small variance.
However, we can repair this defect. Let

$$\sigma_T = \max\{\sup\{\sigma_{j,k}\}, C_0 T^{-1/2}\}.$$

and let $\theta_{j,k} \sim N(0, \sigma_T^2 - \sigma_{j,k}^2)$ be independent of $\tilde{\alpha}_{j,k}$. Then the new random variable $\tilde{\alpha}_{j,k} + \theta_{j,k}$ has the same mean and the same n-th order cumulants for $n \geq 3$ as $\tilde{\alpha}_{j,k}$, whereas its variance is equal to $\sigma_T^2 \asymp T^{-1}$. Therefore, we can derive in complete analogy to Theorem 4.1 the following result.

Theorem 4.2. *Assume (A1) through (A3). Then*

$$\frac{P\left(\pm((\tilde{\alpha}_{j,k} + \theta_{j,k}) - \alpha_{j,k})/\sigma_{j,k} \geq x\right)}{1 - \Phi(x)} \longrightarrow 1$$

holds uniformly in $(j,k) \in \mathcal{J}$, $-\infty < x \le \Delta_\gamma$.

On the basis of these two theorems we will conclude in the next section that the risk of monotone estimators based on $\tilde{\alpha}_{j,k}$ is asymptotically the same as in the case of Gaussian empirical coefficients. This will provide the link to existing results on wavelet estimators in Gaussian regression.

5. Wavelet estimators of the spectral density

First, we recall some known facts about wavelet estimators in the White Noise model. Assume

$$dY_t = f(t) + \epsilon dW_t, \quad t \in [0,1]. \tag{5.1}$$

If we consider minimax estimation problems for $f \in \mathcal{F}$, it is known (for details see Donoho and Johnstone (1992)) that linear estimators are unable to attain the optimal convergence rate if $\mathcal{F}$ is some ball in a Besov space $B^m_{p,q}$ with $p < 2$. Moreover, certain nonlinear wavelet estimators attain this minimax rate. In other words, this result indicates that wavelet estimators can behave much better than traditional linear estimators whenever a certain amount of inhomogeneity in the smoothness of f is present. Since spectral densities have certain peaks in the case of the presence of some near-periodicities, the application of these methods seems to be an interesting attempt to estimate spectral densities both at certain smooth regions as well as near peaks with a good quality.
In various papers by Donoho and Johnstone there were the following nonlinear methods used in the framework of regression:

$$\delta^{(h)}(\tilde{\alpha}_{j,k}, \lambda) = \tilde{\alpha}_{j,k} I(|\tilde{\alpha}_{j,k}| \ge \lambda), \tag{5.2}$$
$$\delta^{(s)}(\tilde{\alpha}_{j,k}, \lambda) = (|\tilde{\alpha}_{j,k}| - \lambda)_+ \, sgn(\tilde{\alpha}_{j,k}), \tag{5.3}$$

where $\tilde{\alpha}_{j,k} = \int \tilde{\psi}_{j,k}(t)\, dY_t$ are appropriate estimates of $\alpha_{j,k}$. These two nonlinear procedures on the empirical coefficients are usually called hard and soft thresholding. There exist two main results concerning the choice of these thresholds.

A) If the thresholds $\lambda = \lambda(j, \epsilon, \mathcal{F})$ are chosen optimally, then

$$\left\{\delta^{(\cdot)}(\tilde{\alpha}_{j,k}, \lambda(j,\epsilon,\mathcal{F}))\right\}$$

is rate optimal in a ball $\mathcal{F} = B^m_{p,q}(C)$ of a certain Besov class.

B) If $\lambda = \epsilon\sqrt{2 \log \#\text{coefficients}}$, i.e. λ is neither depending on j nor $\mathcal{F}$, then the corresponding estimator is optimal in a wide range of smoothness classes within a factor of $\log(1/\epsilon)$.

We turn to the problem of spectral density estimation. As approximating models for our empirical wavelet coefficients we consider

$$\xi_{j,k} = \alpha_{j,k} + \sigma_{j,k}\varepsilon_{j,k} \tag{5.4}$$

and

$$\bar{\xi}_{j,k} = \alpha_{j,k} + \sigma_T \varepsilon_{j,k} \tag{5.5}$$

where $\varepsilon_{j,k} \sim N(0,1)$. Then we have the following basic result for monotone estimators.

Theorem 5.1. *Assume (A1) through (A3). Let $\delta_j = \delta_{j,T}$ be monotone functions with*

$$|\delta_j(y)| \leq |y|. \tag{5.6}$$

Then, for $0 < p' < \infty$,

(i) $\sum_{(j,k)\in\mathcal{J}^0} E|\delta_j(\tilde{\alpha}_{j,k}) - \alpha_{j,k}|^{p'} = \sum_{(j,k)\in\mathcal{J}^0} E|\delta_j(\xi_{j,k}) - \alpha_{j,k}|^{p'}(1+o(1)) + O(T^{-p'/2})$

(ii) $\sum_{(j,k)\in\mathcal{J}} E|\delta_j(\tilde{\alpha}_{j,k}) - \alpha_{j,k}|^{p'} \leq 2 \sum_{(j,k)\in\mathcal{J}} E|\delta_j(\overline{\xi}_{j,k}) - \alpha_{j,k}|^{p'}(1+o(1)) + O(T^{-p'/2})$

holds uniformly over $f \in \mathcal{F}$.

Proceeding from this theorem we can obtain risk properties of thresholded wavelet estimators. Since the estimators (5.2) and (5.3) obey the assumption (5.6), we can immediately derive due to Theorem 5.1 the risk equivalence of our spectral density estimators to analogous estimators in the much simpler models (5.4) and (5.5). Let $\delta^{(\cdot)}$ denote either the hard-threshold rule $\delta^{(h)}$ defined by (5.2) or the soft-threshold rule $\delta^{(s)}$ given by (5.3). Then we can state the following assertion.

Corollary 5.1. *Assume (A1) through (A3). Then, for nonrandom thresholds λ_j,*

(i) $\sum_{(j,k)\in\mathcal{J}^0} E\left(\delta^{(\cdot)}(\tilde{\alpha}_{j,k},\lambda_j) - \alpha_{j,k}\right)^2 = \sum_{(j,k)\in\mathcal{J}^0} E\left(\delta^{(\cdot)}(\xi_{j,k},\lambda_j) - \alpha_{j,k}\right)^2 + O(T^{-1}),$

(ii) $\sum_{(j,k)\in\mathcal{J}} E\left(\delta^{(\cdot)}(\tilde{\alpha}_{j,k},\lambda_j) - \alpha_{j,k}\right)^2 \leq 2 \sum_{(j,k)\in\mathcal{J}} E\left(\delta^{(\cdot)}(\overline{\xi}_{j,k},\lambda_j) - \alpha_{j,k}\right)^2 + O(T^{-1}).$

Let us now assume that the spectral density f lies in a set of the following type:

$$\mathcal{F} = \mathcal{F}(C_1, C_3, C_4) = \left\{ f = \sum_k \alpha_k \phi_{l,k} + \sum_{j,k} \alpha_{j,k}\psi_{j,k} \,\middle|\, TV(f) \leq C_1,\ |\alpha_k| \leq C_3,\ \|\alpha\| \leq C_4 \right\}, \tag{5.7}$$

where

$$\|\alpha\| = \left(\sum_{j\geq l} \left(2^{js} \left(\sum_k |\alpha_{j,k}|^p \right)^{1/p} \right)^q \right)^{1/q} \tag{5.8}$$

with $s = m + 1/2 - 1/p$. It is known that this norm is essentially equivalent to the norm in the Besov space $B^m_{p,q}$, if the basis functions ϕ and ψ are sufficiently regular, cf. Donoho and Johnstone (1992). Moreover, we see by the relation $B^m_{p,1} \subseteq W^m_p \subseteq B^m_{p,\infty}$, that also smoothness classes from the scale of Sobolev spaces W^m_p are covered by our results.

Let $\lambda_j^0 = \lambda_j^0(T, \mathcal{F})$ be the optimal (nonrandom) thresholds and $\lambda_j = \sigma_{\max}\sqrt{2\log K}$, where $K = \#\mathcal{J}$ and $\sigma_{\max} = \max_{(j,k)\in\mathcal{J}}\{\sigma_{j,k}\}$. Further, let

$$\widehat{f}^0 = \sum_{k\in I_l} \tilde{\alpha}_k \phi_{l,k} + \sum_{(j,k)\in\mathcal{J}} \delta^{(\cdot)}(\tilde{\alpha}_{j,k}, \lambda_j^0)\psi_{j,k} \tag{5.9}$$

and

$$\widehat{f} = \sum_{k\in I_l} \tilde{\alpha}_k \phi_{l,k} + \sum_{(j,k)\in\mathcal{J}} \delta^{(\cdot)}(\tilde{\alpha}_{j,k}, \lambda_j)\psi_{j,k}. \tag{5.10}$$

The following theorem shows that we can attain the same rates in our spectral density framework as known from the case of Gaussian regression.

Theorem 5.2. *Assume (A1) through (A3). Then*

$$\begin{aligned} &(i) \quad \sup_{f\in\mathcal{F}}\left\{E\|\widehat{f}^0 - f\|^2\right\} &&= O\left(T^{-2m/(2m+1)}\right), \\ &(ii) \quad \sup_{f\in\mathcal{F}}\left\{E\|\widehat{f} - f\|^2\right\} &&= O\left(T^{-2m/(2m+1)}\log T\right). \end{aligned}$$

(The proof of this theorem follows immediately from Theorem 5.1 and known results in the Gaussian white noise model (5.1), cf. Donoho and Johnstone (1992).)

It is known from Bentkus (1986) that $T^{-2m/(2m+1)}$ is just the optimal rate of convergence in Hölder smoothness classes with degree of smoothness m, if $\{X_t\}$ is a stationary Gaussian process. Therefore, $\widehat{f}^0$ attains the minimax rate of convergence for each smoothness class $\mathcal{F}$, which contains some ball from the corresponding Hölder space. Moreover, we see that $\widehat{f}$ is nearly minimax up to a factor of order $\log T$ in such smoothness classes.

Now we turn to the practical choice of the thresholds λ_j. In general, the adaptive choice of the smoothing parameter(s) in spectral density estimation seems to be more difficult than in regression or density estimation.

Wahba and Wold (1975) proposed some cross-validation criterion to choose the degree of smoothing of a periodic spline estimator automatically. By showing that the expectation of the cross-validation function is asymptotically close to the mean square error they gave some indication that this criterion yields an asymptotically optimal smoothing parameter. Wahba (1980) proposed for spline estimators of the log periodogram to adapt the smoothing parameter by minimizing an asymptotically unbiased risk estimate. Hurvich (1985) proposed in the framework of quite general classes of estimators some refinement of the usual leave-one-out technique used in cross-validation methods. For kernel estimators Beltrão and Bloomfield (1987) proposed some cross-validated likelihood approach to determine the bandwidth, and gave also some indications that this method can work. However, since it is based on Whittle's approximation of the likelihood, its applicability seems to be restricted to Gaussian time series only. Franke and Härdle (1992) proposed to use a bootstrap estimate of the mean square percentage error and showed that the bootstrap distribution approaches the distribution of the estimation errors in probability. This gives certainly some hope that the risk of the corresponding estimator is asymptotically close to the risk of the estimator with optimal bandwidth. However, the construction

of the bootstrap estimate requires, besides the bandwidth of interest, two additional bandwidths that must tend to zero with some suboptimal rates. It is an open problem to find a reasonable data-driven rule to determine these bandwidths, which would provide then a completely defined rule.
Here we propose a much simpler, admittedly slightly suboptimal, approach. Assume that any consistent estimator $\widetilde{f}$ of f is available. By the trivial inequality

$$var(\widetilde{\alpha}_{j,k}) \leq 2\pi \left(H_4^{(T)}/(H_2^{(T)})^2\right) \int_\Pi |\psi_{j,k}(\alpha)\,[\psi_{j,k}(\alpha)+\psi_{j,k}(-\alpha)]|\;d\alpha\;|f|_\infty^2 + O(T^{-1}2^{-j})$$

we can find

$$\widehat{\sigma}_T^2 \;=\; 2\pi \left(H_4^{(T)}/(H_2^{(T)})^2\right) \int_\Pi |\psi_{j,k}(\alpha)\,[\psi_{j,k}(\alpha)+\psi_{j,k}(-\alpha)]|\;d\alpha\;|\widetilde{f}|_\infty^2$$

as an asymptotic majorant of $var(\widetilde{\alpha}_{j,k})$ for all levels $j = j(T)$ with $2^j \gg 1$. Then we can use, according to (ii) of Theorem 5.2, $\widehat{\lambda} = \widehat{\sigma}_T\sqrt{2\log(\#\mathcal{J})}$ as a universal threshold, which is completely data-driven. We mention that Gao (1993) used thresholds of order $T^{-1/2}\log T$, which can be applied to Gaussian time series without the restriction to levels j with $2^j \ll T$. The use of these thresholds results in an L_2-risk of the estimator that is within a factor of $(\log T)^2$ of the risk with optimally thresholded coefficients.
We expect that one can show under appropriate conditions that the corresponding estimator is, within a $\log T$-term as in (ii) of Theorem 5.2, rate-optimal in a wide range of smoothness classes.

6. Cumulants of the discrete empirical coefficients

We show in this section that the cumulants of $\overline{\alpha}_{j,k}$ and $\overline{\overline{\alpha}}_{j,k}$ are asymptotically close to those of $\widetilde{\alpha}_{j,k}$, which will imply that analogous results as in Section 5 hold true for thresholded wavelet estimators based on the discrete versions of the empirical wavelet coefficients.
First, we state a technical lemma, which will be the basis for estimates of the cumulants. Let

$$\overline{a}_{j,k}(l) \;=\; \sum_i \int_{\omega_i - 2\pi/T}^{\omega_i} \psi_{j,k}(t)\,dt\;\cos(\omega_i l)$$

and

$$\overline{\overline{a}}_{j,k}(l) \;=\; \frac{2\pi}{T}\sum_i \psi_{j,k}(\omega_i)\;\cos(\omega_i l).$$

Lemma 6.1. *There exists a constant $K < \infty$ such that*

$$\begin{aligned}
&(i) && |\overline{a}_{j,k}(l) - a_{j,k}(l)| && \leq KT^{-1}l2^{-j/2},\\
&(ii) && |\overline{\overline{a}}_{j,k}(l) - \overline{a}_{j,k}(l)| && \leq KT^{-1}2^{j/2},\\
&(iii) && |\overline{a}_{j,k}(l)| + |\overline{\overline{a}}_{j,k}(l)| && \leq K\left(2^{j/2}\left(\frac{1}{l}+\frac{1}{T-l}\right)\wedge 2^{-j/2}\right)
\end{aligned}$$

hold uniformly over $(j,k) \in \mathcal{J}$.

On the basis of this lemma we can state the following assertions concerning the cumulants of $\overline{\alpha}_{j,k}$ and $\overline{\overline{\alpha}}_{j,k}$.

Proposition 6.1. *Assume (A1) through (A3). Then*

$$
\begin{aligned}
&(i) \quad E\overline{\alpha}_{j,k} - E\widetilde{\alpha}_{j,k} = O(2^{j/2}T^{-1}),\\
&(ii) \quad var(\overline{\alpha}_{j,k}) = var(\widetilde{\alpha}_{j,k}) + O(T^{-1}2^{-j}) + o(T^{-1}),\\
&(iii) \quad |cum_n(\overline{\alpha}_{j,k})| \leq var(\widetilde{\alpha}_{j,k})(n-1)!\left(4\pi\|\psi_{j,k}\|_\infty\|f\|_\infty\frac{\sup\{|h(t)|^2\}}{H_2^{(T)}}\right)^{n-2}\\
&\qquad + \frac{2^{n-2}C_2^{2n}}{\pi^n}((2n)!)^{1+\gamma}\left(\frac{\sup\{|h(t)|^2\}}{H_2^{(T)}}\right)^n \sup_{l=0,\pm 1,\dots,\pm T-1}\{|\overline{a}_{j,k}(l)|^2\}T\overline{M}^{n-2}.
\end{aligned}
$$

hold uniformly over $f \in \mathcal{F}$ and $(j,k) \in \mathcal{J}$.
(For the cumulants of $\overline{\overline{\alpha}}_{j,k}$ we have analogous estimates.)

In view of the estimates of the cumulants of $\overline{\alpha}_{j,k}$ and $\overline{\overline{\alpha}}_{j,k}$ given in this proposition we can infer that the results of Section 5 remain true for estimators based on the discrete empirical coefficients.

7. Concluding remarks

1) There exist several possibilities to weaken the cumulant assumption (A3). One could obtain analogous results as in Theorem 5.1 under the following modifications of (A3).
   (A3') There exist random variables $\widetilde{X}_{t_i}$ such that
   (i) $P(X_{t_i} \neq \widetilde{X}_{t_i}) \leq Cn^{-\gamma}$ for γ large enough
   (ii) $\sup_{1<t_1<\infty}\left\{\sum_{t_2,\dots,t_k=1}^{\infty}\left|cum(\widetilde{X}_{t_1},\dots,\widetilde{X}_{t_k}\right|\right\} \leq C_2^k(k!)^{1+\gamma}$ for all $k = 2, 3, \dots$
   (iii) $E|X_t|^\gamma \leq C(\gamma)$ for some finite but sufficiently large γ.
   (A3") The constant C_2 in (A3) can be allowed to increase with T with a sufficiently slow rate.
2) Sometimes a mixing condition as in Remark 1 is imposed instead of our assumption (A3). However, by the relation

$$|cov(X_t, X_{t+k})| \leq 8\alpha^{1/r}(t, t+k)\|X_t\|_p\|X_{t+k}\|_q$$

for $p, q, r \geq 1$, $1/p + 1/q + 1/r = 1$ (cf. Davydov (1970)) we infer that

$$\left|f^{(\theta)}(\omega)\right| \leq \frac{1}{\pi}\sum k^\theta |c_k| = O(1)$$

holds for all $\theta < \infty$, which means that the underlying spectral density is arbitrarily smooth. In this case, also traditional kernel estimators based on $I_T(\omega)$ would do a good job in estimating f and there is no need to apply nonlinear wavelet estimators.
In contrast, under (A3) we include essentially more irregular spectral densities, which is demonstrated by the following somewhat artificial example.
Let $\{X_t^{(i)}\}$, $i = 1, \dots, m$, be independent time series satisfying (A3) with constants $C_2^{(i)}$, respectively. Let $\gamma \sim U\{1, \dots, m\}$ be independent of these processes and let

$$Y_{(t-1)m+i} = X_t^{(i+\gamma \bmod m)}.$$

Then $\{Y_t\}$ is a stationary process satisfying (A3) with constant $C_2 = m^{-1}\sum_i C_2^{(i)}$. For the covariances we have

$$cov(Y_t, Y_{t+jm+k}) = \begin{cases} m^{-1}\sum_i cov(X_0^{(i)}, X_j^{(i)}), & \text{for } k = 0 \\ 0, & \text{for } 1 \le k < m \end{cases},$$

which implies by $f_{\{Y_t\}}(\omega) = m^{-1}\sum_i f_{\{X_t^{(i)}\}}(m\omega)$ that

$$\frac{d^\theta}{d\omega^\theta} f_{\{Y_t\}}(\omega) = m^{\theta-1}\sum_i \frac{d^\theta}{d\omega^\theta} f_{\{X_t^{(i)}\}}(m\omega).$$

Letting m tend to infinity we cover more and more irregular spectral densities.

3) It is pointed out in several papers, e.g. Dahlhaus (1983, 1990), that linear estimators can be improved by the use of an appropriate data taper. In particular, in Dahlhaus (1990) it is shown that such a data taper is necessary to get estimators with a high resolution property.
The results presented here allow the use of the tapered periodogram as an alternative to the nontapered one, too. It would be a challenging problem to investigate the effect of an taper to our nonlinear estimation rules.
4) As may be seen in Theorem 5.1, the risk equivalence to the Gaussian case can also be shown for other loss functions as the quadratic one. If we are especially interested in recognizing peaks of the true spectral density, we could take this special goal into account by an appropriate choice of the loss function. The L_p-loss for large values of p is more sensitive than the L_2-loss to relatively large estimation errors at certain small intervals, and therefore more sensitive to a misjudgement of a peak.

8. Proofs

For simplicity of notation we agree that C denotes any finite constant, which may take different values at different places.

Proof of Remark 1. Using 2) of Theorem 3 in [20] with $\delta = 1$ we obtain, for $t_1 \le t_2 \le \ldots \le t_k$,

$$|\Gamma(X_{t_1}, \ldots, X_{t_k})| \le C^k (k-1)!((2k)!)^{\gamma/2} \prod_{j=2}^{k} \alpha^{1/(2(k-1))}(t_{j-1}, t_j), \qquad (8.1)$$

which implies

$$\begin{aligned} \sum_{t_2,\ldots,t_k=1}^{\infty} |\Gamma(X_{t_1}, \ldots, X_{t_k})| &\le k! \sum_{1\le t_2\le t_3\le\ldots\le t_k<\infty} |\Gamma(X_{t_1}, \ldots, X_{t_k})| \\ &\le C^k (k!)^2((2k)!)^{\gamma/2} \sum_{t_2\ge t_1} \alpha^{1/(2(k-1))}(t_1, t_2) * \cdots * \sum_{t_k \ge t_{k-1}} \alpha^{1/(2(k-1))}(t_{k-1}, t_k) \\ &\le C^k (k!)^{2+\gamma}(k-1)^{k-1} \\ &\le C^k (k!)^{3+\gamma}. \end{aligned}$$

□

Proof of Proposition 3.1. (i) In the nontapered case we obtain this assertion immediately by formula (1.7) in [14]. In the tapered case Lemma 4 in [6] provides under the weaker condition $f \in L_2$ an error estimate of order $TV(\psi_{j,k})o(T^{-1/2})$, which is $o(T^{-1/2}2^{j/2})$. For reader's convenience we adopt the notation introduced in [6] as much as possible.
Using Theorem 5.2.3 in [5] we obtain

$$E\, I_T(\omega) = \int_\Pi \Phi_2^{(T)}(\alpha) f(\omega-\alpha)\, d\alpha,$$

where

$$\Phi_k^{(T)}(\alpha) = \begin{cases} \frac{H_1^{(T)}(\alpha_1)\cdots H_1^{(T)}(\alpha_{k-1})H_1^{(T)}(-\sum_{j=1}^{k-1}\alpha_j)}{(2\pi)^{k-1}H_k^{(T)}(0)}, & \text{if } H_k^{(T)}(0) \neq 0 \\ 0, & \text{if } H_k^{(T)}(0) = 0 \end{cases},$$

$H_k^{(T)}(\alpha) = \sum_t h^k(t/T)\exp\{-i\alpha t\}$, which implies

$$\begin{aligned} E\tilde{\alpha}_{j,k} &= \int_\Pi \int_\Pi \Phi_2^{(T)}(\alpha) f(\omega-\alpha)\psi_{j,k}(\omega)\, d\alpha\, d\omega \\ &= \int_\Pi \Phi_2^{(T)}(\alpha) \int_\Pi f(\omega-\alpha)\psi_{j,k}(\omega)\, d\omega\, d\alpha. \end{aligned} \tag{8.2}$$

From (4) in [6] we conclude

$$\int_\Pi \Phi_2^{(T)}(\alpha)\, d\alpha = \frac{1}{2\pi H_2^{(T)}(0)} \int_\Pi H_1^{(T)}(\alpha) H_1^{(T)}(-\alpha)\, d\alpha = 1,$$

which implies that

$$\begin{aligned} E\tilde{\alpha}_{j,k} - \alpha_{j,k} &= \int_\Pi \Phi_2^{(T)}(\alpha) \int_\Pi (f(\omega-\alpha) - f(\omega))\, \psi_{j,k}(\omega)\, d\omega\, d\alpha \\ &= \frac{1}{2\pi H_2^{(T)}} \int_\Pi H_1^{(T)}(\alpha) H_1^{(T)}(-\alpha) \int_\Pi (f(\omega-\alpha) - f(\omega))\, \psi_{j,k}(\omega)\, d\omega\, d\alpha. \end{aligned} \tag{8.3}$$

Using (6) in [6] we get

$$\left|H_1^{(T)}(\alpha)\right| \leq C|\alpha|^{-1}$$

and, by Lemma 1 in that paper,

$$\int_\Pi \left|H_1^{(T)}(\alpha)\right|\, d\alpha \leq C \log T.$$

Because of

$$\left|\int_\Pi (f(\omega-\alpha) - f(\omega))\, \psi_{j,k}(\omega)\, d\omega\right| = O\left(2^{j/2}|\alpha|\right),$$

we obtain finally

$$|E\tilde{\alpha}_{j,k} - \alpha_{j,k}| \leq C 2^{j/2} T^{-1} \int_\Pi |H_1^{(T)}(\alpha)||H_1^{(T)}(-\alpha)||\alpha|\, d\alpha = O\left(2^{j/2}T^{-1}\log T\right).$$

(ii) The proof is, of course, very similar to that of Lemma 6 in [6]. The only, but important modification for our purposes is, that the weight functions $\psi_{j,k}$ are not uniformly bounded, and therefore we carry out the proof again.
According to Lemma 6 in [6] we have

$$var(\tilde{\alpha}_{j,k}) = \left(H_4^{(T)}/(H_2^{(T)})^2\right)\int_{\Pi^3} G(u)\Phi_4^{(T)}(u)\,\lambda^3(du),$$

where

$$\begin{aligned} G_1(u) &= 2\pi\int_{\Pi^2}\psi_{j,k}(\alpha_1)\psi_{j,k}(-\alpha_2)f_4(\alpha_1-u_1,-\alpha_1-u_2,\alpha_2-u_3)\lambda^2(d\alpha),\\ G_2(u) &= 2\pi\int_{\Pi}\psi_{j,k}(\alpha)\psi_{j,k}(\alpha-u_1-u_3)f(\alpha-u_1)f(-\alpha-u_2)\,d\alpha,\\ G_3(u) &= 2\pi\int_{\Pi}\psi_{j,k}(\alpha)\psi_{j,k}(-\alpha-u_2-u_3)f(\alpha-u_1)f(-\alpha-u_2)\,d\alpha. \end{aligned}$$

Now, we have

$$|G_1(u)| \le 2\pi|f_4|_\infty\left(\int_{\Pi^2}\psi_{j,k}(\alpha)\,d\alpha\right)^2 = O(2^{-j}),$$

which implies

$$\int_{\Pi^3} G_1(u)\Phi_4^{(T)}(u)\,\lambda^3(du) = O(2^{-j}).$$

Further, we have

$$\begin{aligned} &|G_2(u) - G_2(0)|\\ &\le 2\pi\left|\int_{\Pi}\psi_{j,k}(\alpha)\left(\psi_{j,k}(\alpha-u_1-u_3)-\psi_{j,k}(\alpha)\right)f(\alpha-u_1)f(-\alpha-u_2)\,d\alpha\right|\\ &\quad + 2\pi\int_{\Pi}|\psi_{j,k}(\alpha)|^2|f(\alpha-u_1)-f(\alpha)||f(-\alpha-u_2)|\,d\alpha\\ &\quad + 2\pi\int_{\Pi}|\psi_{j,k}(\alpha)|^2|f(\alpha)||f(-\alpha-u_2)-f(-\alpha)|\,d\alpha\\ &= O(|u|2^j). \end{aligned}$$

This yields, by the estimate contained in the proof of Lemma 3 in [6],

$$\begin{aligned} &\left|\int_{\Pi^3} G_2(u)\Phi_4^{(T)}(u)\,\lambda^3(du) - G_2(0)\int_{\Pi^3}\Phi_4^{(T)}(u)\lambda^3(du)\right|\\ &\le \int_{\Pi^3\setminus\{|u|<\delta\}}|G_2(u)-G_2(0)||\Phi_4^{(T)}(u)|\lambda^3(du)\\ &\quad + \int_{\{|u|<\delta\}}|G_2(u)-G_2(0)||\Phi_4^{(T)}(u)|\lambda^3(du)\\ &\le O\left(\frac{\log^3 T}{\delta T}\|\psi_{j,k}\|_2^2\|f\|_\infty^2\right) + O(\delta 2^j). \end{aligned}$$

Choosing $\delta = \delta(T) \asymp T^{-1/2}2^{-j/2}(\log T)^{-3/2}$ we obtain a residual term of order $O(2^{j/2}T^{-1/2}(\log T)^{3/2})$, which is $o(1)$ under $2^j \le CT^{1-\alpha}$.
An analogous upper estimate can be obtained for the difference containing G_3.
Since $H_4^{(T)}/(H_2^{(T)})^2 = O(T^{-1})$, we obtain (ii).

(iii) This proof is similar to that of Lemma 2 in [18]. Because of the use of a data taper and because of the more general cumulant assumption we have to modify it slightly. For the reader's convenience we sketch the main steps. The higher order cumulants of $\tilde{\alpha}_{j,k}$ can be estimated with the aid of Leonov-Shiryaev's formula, [15]. For that we have to write $\tilde{\alpha}_{j,k}$ as a polynomial in certain random variables, where each of them enters at most with the power one into each of the terms of this polynomial. Therefore, we define $X_{T+t} = X_t$ for $t = 1, \ldots, T$, such that $\eta = \tilde{\alpha}_{j,k}$ can be written in the form

$$\eta = \sum_{\vartheta \in \Theta} a(\vartheta) X_1^{\vartheta_1} \cdots X_{2T}^{\vartheta_{2T}}, \tag{8.4}$$

where $\vartheta = (\vartheta_1, \ldots, \vartheta_{2T})$, $\Theta = \{\vartheta \,|\vartheta_t \in \{0,1\}, \sum \vartheta_t = 2\}$.
Using Leonov-Shiryaev's formula we obtain

$$cum_n(\eta) = \sum_{\vartheta^{(1)},\ldots,\vartheta^{(n)} \in \Theta} a(\vartheta^{(1)}) \cdots a(\vartheta^{(n)}) \sum_{D(\vartheta)=D_1+\ldots+D_q}^{*} c(t_{D_1}) \cdots c(t_{D_q}), \tag{8.5}$$

where the sum $\sum^*$ is taken over all indecomposable partitions of $D(\vartheta) = \left\{(i,j) \,\middle|\, \vartheta_j^{(i)} = 1\right\}$. Here, $c(t_{D_p})$ is defined as $cum\left(X_{m_{p,1}}, \ldots, X_{m_{p,r_p}}\right)$, where $D_p = \left\{(i_{m_{p,1}}, m_{p,1}), \ldots, (i_{m_{p,r_p}}, m_{p,r_p})\right\}$.
For our further calculations it is not necessary to work with the general, but quite involved formula (8.5); it tells us only, which products of cumulants must be included in the calculation of $cum_n(\eta)$. We rewrite (8.5) as

$$cum_n(\eta) = (2\pi H_2^{(T)})^{-n} \sum_{t_1,\ldots,t_{2n}=1}^{T} a_{j,k}(t_1 - t_2) \cdots a_{j,k}(t_{2n-1} - t_{2n}) \sum_{D=D_1+\ldots+D_q}^{**} c(t_{D_1}) \cdots c(t_{D_q}), \tag{8.6}$$

where $D = \{t_1, \ldots, t_{2n}\}$ and $\sum^{**}$ is the analog of $\sum^*$ from (8.5). Analogously to (8.5) we define $c(t_{D_p}) = c(t_{m_{p,1}}, \ldots, t_{m_{r_p,1}}) = cum\left(h_{t_{m_{p,1}}} X_{t_{m_{p,1}}}, \ldots, h_{t_{m_{p,r_p}}} X_{t_{m_{p,r_p}}}\right)$ if $D_p = \left\{t_{m_{p,1}}, \ldots, t_{m_{p,r_p}}\right\}$.
Since $E\,X_t = 0$ for all t, we can restrict our considerations to indecomposable decompositions with $r_p \geq 2$. We divide the right-hand side of (8.6) into two parts, R_1 and R_2, the first one containing all indecomposable partitions $D = D_1 + \ldots + D_n$, which have exactly n sets, and the second one containing the remaining partitions $D = D_1 + \ldots + D_q$ for any $q < n$.
Because of $r_p \geq 2$, R_1 contains all pure products of second order cumulants and is just the n-th cumulant of η, if the time series would be Gaussian with zero mean and the same spectral density as $\{X_t\}$. Following the lines of the proof of Lemma 3.1 in [4], we conclude

$$|R_1| = var(\tilde{\alpha}_{j,k})(n-1)! \left(4\pi \|\psi_{j,k}\|_\infty \|f\|_\infty \frac{\sup\{|h(t)|^2\}}{H_2^{(T)}}\right)^{n-2}. \tag{8.7}$$

Now we estimate $|R_2|$. Let $D(\vartheta) = D_1 + \ldots + D_q$ be any indecomposable partition, which occurs in the sum $\sum^*$. Since it is indecomposable, there exist numbers $p_2, \dots, p_q \in \{1, \dots, n\}$ and associated indices t_{p_i}, $\tilde{t}_{p_i}$ such that

$$(p_i, t_{p_i}) \in D_i \quad \text{and} \quad (p_i, \tilde{t}_{p_i}) \in \bigcup_{j=1}^{i-1} D_j \quad \text{for all } i = 2, \dots, q. \tag{8.8}$$

Let us now fix a partition $D = D_1 + \ldots + D_q$, which occurs in $\sum^{**}$. According to (8.8) there exist pairs $(s_p, \widehat{s}_p)$ with $\{s_p, \widehat{s}_p\} = \{2i-1, 2i\}$ for some $i \in \{1, \dots, n\}$ with

$$s_p \in D_p, \quad \widehat{s}_p \in \bigcup_{j=1}^{p-1} D_j. \tag{8.9}$$

Since the sets D_p are not ordered, we set w.l.o.g. $m_{1,r_1} = \widehat{s}_2$ and $m_{p,r_p} = s_p$, $p = 2, \dots, q$.
Now we have

$$\begin{aligned} &\left| \sum_{t_1,\dots,t_{2n}=1}^{T} a_{j,k}(t_1 - t_2) \cdots a_{j,k}(t_{2n-1} - t_{2n}) c(t_{D_1}) \cdots c(t_{D_q}) \right| \\ &\le \sup_l |a_{j,k}(l)|^{n-q-1} * \\ &* \sum_{t_{m_{1,1}},\dots,t_{m_{1,r_1-1}},t_{\widehat{s}_2}=1}^{T} \left| c\left(t_{m_{1,1}}, \dots, t_{m_{1,r_1-1}}, t_{\widehat{s}_2}\right) \right| * \\ &* \sum_{t_{s_2}=1}^{T} \left| a_{j,k}(t_{s_2} - t_{\widehat{s}_2}) \right| \sum_{t_{m_{2,1}},\dots,t_{m_{2,r_2-1}}=1}^{T} \left| c\left(t_{m_{2,1}}, \dots, t_{m_{2,r_2-1}}, t_{s_2}\right) \right| * \\ &\dots \\ &* \sum_{t_{s_q}=1}^{T} \left| a_{j,k}(t_{s_q} - t_{\widehat{s}_q}) \right| \sum_{t_{m_{q,1}},\dots,t_{m_{q,r_q-1}}=1}^{T} \left| c\left(t_{m_{q,1}}, \dots, t_{m_{q,r_q-1}}, t_{s_q}\right) \right| . \end{aligned} \tag{8.10}$$

Now we stepwise introduce upper estimates for the sums on the right-hand side of (8.10). The last sum is, uniformly in $\widehat{s}_q$, less or equal than $M\widetilde{C}_2^{r_q}(r_q!)^{1+\gamma}$, where $\widetilde{C}_2 = C_2 \sup\{|h(t)|\}$. Then we estimate the last but one sum above by $M\widetilde{C}_2^{r_{q-1}}(r_{q-1}!)^{1+\gamma}$, and so on. Finally we obtain that

$$\begin{aligned} &\left| \sum_{t_1,\dots,t_{2n}=1}^{T} a_{j,k}(t_1 - t_2) \cdots a_{j,k}(t_{2n-1} - t_{2n}) c(t_{D_1}) \cdots c(t_{D_q}) \right| \\ &\le \sup_l \left\{ |a_{j,k}(l)|^{n-q-1} \right\} T M^{q-1} \widetilde{C}_2^{2n} (r_1! \cdots r_q!)^{1+\gamma} \\ &\le \sup_l \left\{ |a_{j,k}(l)|^2 \right\} T M^{n-2} \widetilde{C}_2^{2n} (r_1! \cdots r_q!)^{1+\gamma} . \end{aligned} \tag{8.11}$$

From the proof of Lemma 2 in [18] we get

$$\sum_{D=D_1+\dots+D_q}^{**} (r_1! \cdots r_q!)^{1+\gamma} \le ((2n)!)^{\gamma} \sum^{**} r_1! \cdots r_q! \le 2^{2n-2} ((2n)!)^{1+\gamma}, \tag{8.12}$$

which yields

$$|R_2| \leq \frac{2^{n-2}C_2^{2n}}{\pi^n}((2n)!)^{1+\gamma}\left(\frac{\sup\{|h(t)|^2\}}{H_2^{(T)}}\right)^n \sup_{l=0,\pm 1,\dots,\pm T-1}\{|a_{j,k}(l)|^2\}TM^{n-2}. \tag{8.13}$$

□

Proof of Lemma 3.1. For $l \neq 0$ we obtain via integration by parts

$$a_{j,k}(l) = \int \psi_{j,k}(t)\cos(lt)\,dt = -\frac{1}{l}\int \sin(lt)\,d\psi_{j,k}(t),$$

which implies

$$|a_{j,k}(l)| \leq l^{-1}TV(\psi_{j,k}) = O\left(l^{-1}2^{j/2}\right). \tag{8.14}$$

On the other hand, we have

$$|a_{j,k}(l)| \leq \int |\psi_{j,k}(t)|\,dt = O\left(2^{-j/2}\right), \tag{8.15}$$

which proves (i).
(ii) follows immediately from $\sum_{m=1}^{T} m^{-1} \leq 1 + \log T$. □

Proof of Theorem 4.1. Using (iii) of Proposition 3.1 as well as the assertion of Lemma 3.1 we obtain due to Lemma 1 in [19] that

$$\frac{P\left(\pm(\tilde{\alpha}_{j,k} - E\tilde{\alpha}_{j,k})/\sigma_{j,k} \geq x\right)}{1-\Phi(x)} = e^{L(x)}\left(1 + g(x)\frac{x+1}{\Delta^{1/(3+4\gamma)}}\right) \tag{8.16}$$

for $0 \leq x \leq \Delta^{1/(3+4\gamma)}$, where Δ is an appropriate constant of order $T^{1/2}2^{-j/2}(\log T)^{-1}$. Here, g is some bounded function, and by (3) and (4) in [19] we infer that $|L(x)| \leq Cx^3/\Delta$. Therefore, the right-hand side of (8.16) converges uniformly to 1, if both $0 \leq x \leq \Delta^{1/(3+4\gamma)}$ and $x = o(\Delta^{1/3})$ are satisfied. (For $x < 0$ this convergence is an immediate consequence.)
Since $b = (\alpha_{j,k} - E\tilde{\alpha}_{j,k})/\sigma_{j,k} = O(\Delta^{-1})$, we conclude that

$$\frac{P\left(\pm(\tilde{\alpha}_{j,k} - \alpha_{j,k})/\sigma_{j,k} \geq x\right)}{1-\Phi(x)} = (1 + o(1))\frac{1-\Phi(x+b)}{1-\Phi(x)} \tag{8.17}$$

holds uniformly over $-\infty \leq x \leq \Delta_\gamma$. Now it remains to estimate

$$\eta = \left|\frac{1-\Phi(x+b)}{1-\Phi(x)} - 1\right| = \frac{|\Phi(x+b)-\Phi(x)|}{1-\Phi(x)}.$$

Let w.l.o.g. $b \geq 0$. (Otherwise we would estimate $\left|\frac{1-\Phi(x)}{1-\Phi(x+b)} - 1\right|$ in the same manner as η below.)
Fix any $c > 1$. It is clear that η tends to zero if $x \leq c$. For $c < x \leq \Delta_\gamma$ we obtain by the formula at the bottom of p. 525 in [4] that

$$\eta \leq \frac{b\frac{1}{\sqrt{2\pi}}e^{-x^2/2}}{\left(1-\frac{1}{x^2}\right)\frac{1}{\sqrt{2\pi}}\frac{1}{x}e^{-x^2/2}} \leq \frac{bx}{1-\frac{1}{c^2}} \longrightarrow 0 \quad \text{as} \quad T \to \infty,$$

which completes the proof. □

Proof of Theorem 5.1. First we choose a constant $\gamma_{j,k}$ such that

$$\begin{aligned} \delta_j(y) &\geq \alpha_{j,k}, \text{ if } y > \gamma_{j,k} \\ \delta_j(y) &\leq \alpha_{j,k}, \text{ if } y < \gamma_{j,k}. \end{aligned}$$

(W.l.o.g. we assume $\delta_j(\gamma_{j,k}) \geq \alpha_{j,k}$.)
Now we split up

$$\begin{aligned} & E\,|\delta_j(\tilde{\alpha}_{j,k}) - \alpha_{j,k}|^{p'} \\ &= E\,I\,(\gamma_{j,k} \leq \tilde{\alpha}_{j,k} \leq \alpha_{j,k} + \sigma_{j,k}\Delta_\gamma)\,|\delta_j(\tilde{\alpha}_{j,k}) - \alpha_{j,k}|^{p'} \\ &\quad + E\,I\,(\alpha_{j,k} - \sigma_{j,k}\Delta_\gamma \leq \tilde{\alpha}_{j,k} < \gamma_{j,k})\,|\delta_j(\tilde{\alpha}_{j,k}) - \alpha_{j,k}|^{p'} \\ &\quad + E\,I\,(|\tilde{\alpha}_{j,k} - \alpha_{j,k}| > \sigma_{j,k}\Delta_\gamma)\,|\delta_j(\tilde{\alpha}_{j,k}) - \alpha_{j,k}|^{p'} \\ &= R_1 + R_2 + R_3. \end{aligned} \tag{8.18}$$

According to the assertion of Theorem 4.1 there exist $C_T^{(l)}$, $C_T^{(u)}$, both tending to 1 as $t \to \infty$, such that

$$C_T^{(l)}(1 - \Phi(x)) \leq P\,(\pm(\tilde{\alpha}_{j,k} - \alpha_{j,k})/\sigma_{j,k} \geq x) \leq C_T^{(u)}(1 - \Phi(x)) \quad \forall x \leq \Delta_\gamma, \tag{8.19}$$

which implies

$$P\,(\Delta_\gamma \geq \pm(\tilde{\alpha}_{j,k} - \alpha_{j,k})/\sigma_{j,k} \geq x) \leq C_T^{(u)}(1 - \Phi(x)) \quad \forall x \in \mathbb{R}. \tag{8.20}$$

Since $\delta_j(y) - \alpha_{j,k}$ is monotone nondecreasing for $y \geq \gamma_{j,k}$, we obtain

$$R_1 \leq C_T^{(u)} E\,I\,(\gamma_{j,k} \leq \xi_{j,k})\,|\delta_j(\xi_{j,k}) - \alpha_{j,k}|^{p'}. \tag{8.21}$$

Analogously we get

$$R_2 \leq C_T^{(u)} E\,I\,(\gamma_{j,k} > \xi_{j,k})\,|\delta_j(\xi_{j,k}) - \alpha_{j,k}|^{p'}. \tag{8.22}$$

Using the formula on the bottom of p. 525 in [4], we have

$$\begin{aligned} P\,(|\tilde{\alpha}_{j,k} - \alpha_{j,k}| > \sigma_{j,k}\Delta_\gamma) &\leq C_T^{(u)}(1 - \Phi(\Delta_\gamma)) \\ &\leq C_T^{(u)} \frac{1}{\sqrt{2\pi}\Delta_\gamma} e^{-\Delta_\gamma^2/2} = O\left(T^{-\mu}\right) \end{aligned}$$

for arbitrary $\mu < \infty$, which implies by the Cauchy-Schwarz inequality

$$R_3 \leq \sqrt{P(|\tilde{\alpha}_{j,k} - \alpha_{j,k}| > \sigma_{j,k}\Delta_\gamma)}\sqrt{E|\delta_j(\xi_{j,k}) - \alpha_{j,k}|^{2p'}} = O\left(T^{-p'/2-1}\right). \tag{8.23}$$

By (8.18) and (8.21) through (8.23) we conclude

$$E\,|\delta_j(\tilde{\alpha}_{j,k}) - \alpha_{j,k}|^{p'} \leq E\,|\delta_j(\xi_{j,k}) - \alpha_{j,k}|^{p'} + O\left(T^{-p'/2-1}\right).$$

The reverse inequality can be proved analogously.

(ii) Let $\theta_{j,k} \sim N(0, \sigma_T^2 - \sigma_{j,k}^2)$ be independent of $\widetilde{\alpha}_{j,k}$. Then

$$\begin{aligned} E|\delta_j(\widetilde{\alpha}_{j,k}) - \alpha_{j,k}|^{p'} &= E\, I(\widetilde{\alpha}_{j,k} \geq \gamma_{j,k})|\delta_j(\widetilde{\alpha}_{j,k}) - \alpha_{j,k}|^{p'} + E\, I(\widetilde{\alpha}_{j,k} < \gamma_{j,k})|\delta_j(\widetilde{\alpha}_{j,k}) - \alpha_{j,k}|^{p'} \\ &= 2E\, I(\widetilde{\alpha}_{j,k} \geq \gamma_{j,k},\, \theta_{j,k} \geq 0)|\delta_j(\widetilde{\alpha}_{j,k}) - \alpha_{j,k}|^{p'} + 2E\, I(\widetilde{\alpha}_{j,k} < \gamma_{j,k},\, \theta_{j,k} \leq 0)|\delta_j(\widetilde{\alpha}_{j,k}) - \alpha_{j,k}|^{p'} \\ &\leq 2E\, |\delta_j(\widetilde{\alpha}_{j,k} + \theta_{j,k}) - \alpha_{j,k}|^{p'}, \end{aligned}$$

which yields (ii) due to Theorem 4.2. □

Proof of Lemma 6.1. (i) It holds

$$\begin{aligned} |a_{j,k}(l) - \overline{a}_{j,k}(l)| &= \left| \sum_i \int_{\omega_i - 2\pi/T}^{\omega_i} \psi_{j,k}(\omega) \left[\cos(\omega l) - \cos(\omega_i l)\right] d\omega \right| \\ &= O\left(T^{-1} l \int |\psi_{j,k}(\omega)|\, d\omega\right) = O\left(T^{-1} l 2^{-j/2}\right). \end{aligned}$$

(ii) Here we have the following estimate

$$\begin{aligned} |\overline{a}_{j,k}(l) - \overline{\overline{a}}_{j,k}(l)| &= \sum_i \int_{\omega_i - 2\pi/T}^{\omega_i} (\psi_{j,k}(\omega) - \psi_{j,k}(\omega_i))\, d\omega\ \cos(\omega_i l) \\ &\leq \sum_i \int_{\omega_i - 2\pi/T}^{\omega_i} |\psi_{j,k}(\omega) - \psi_{j,k}(\omega_i)|\, d\omega \\ &\leq 2\pi T^{-1} TV(\psi_{j,k}) \\ &= O\left(T^{-1} 2^{j/2}\right). \end{aligned}$$

(iii) Using the formula

$$\sum_{s=1}^{N} \exp\{is\alpha\} = \frac{\sin \frac{N\alpha}{2}}{\sin \frac{\alpha}{2}} \exp\left\{i \frac{N+1}{2} \alpha\right\}$$

from p. 33 in [3] we get for $0 < l < T$

$$\begin{aligned} d_N(l) &= \sum_{s=1}^{N} \cos(\omega_s l) \\ &= \sum_{s=1}^{N} \cos(s * 2\pi l/T) \\ &= \frac{\sin \frac{N 2\pi l/T}{2}}{\sin \frac{2\pi l/T}{2}} \cos\left(\frac{N+1}{2} 2\pi l/T\right), \end{aligned}$$

which implies

$$|d_N(l)| \leq K \left(\frac{T}{l} + \frac{T}{T-l}\right).$$

(For $l = 0, T$ we have $d_N(l) = N$.)

With the definitions $r(T) = \psi_{j,k}(\omega_T)$ and $r(m) = \psi_{j,k}(\omega_m) - \psi_{j,k}(\omega_{m+1})$ for $0 \le m \le T-1$ we have, for $0 < l < T$,

$$\begin{aligned}
\bar{a}_{j,k}(l) &= \frac{2\pi}{T}\sum_{i=1}^{T}\psi_{j,k}(\omega_i)\cos(\omega_i l) \\
&= \frac{2\pi}{T}\sum_{i}\left(\sum_{j\ge i} r(j)\right)\cos(\omega_i l) \\
&= \frac{2\pi}{T}\sum_{j} r(j)\sum_{i=1}^{j}\cos(\omega_i l) \\
&= \frac{2\pi}{T}\sum_{j=1}^{T-1} r(j) d_j(l).
\end{aligned}$$

Since $\sum_{j=1}^{T-1}|f(j)| \le TV(\psi_{j,k}) \le K2^{j/2}$, we obtain

$$|\bar{a}_{j,k}(l)| \le 2\pi 2^{j/2} K\left(\frac{1}{l} + \frac{1}{T-l}\right). \tag{8.24}$$

The proof for $\bar{\bar{a}}_{j,k}(l)$ is analogous, we replace only the $r(m)$'s by $\tilde{r}(T) = \int_{\omega_T-2\pi/T}^{\omega_T}\psi_{j,k}(t)\,dt$ and $\tilde{r}(m) = \int_{\omega_m-2\pi/T}^{\omega_m}\psi_{j,k}(t)\,dt - \int_{\omega_{m+1}-2\pi/T}^{\omega_{m+1}}\psi_{j,k}(t)\,dt$ for $0 \le m \le T-1$.
Since $\sum_{m=1}^{T-1}|\tilde{r}(m)| = O(T^{-1}TV(\psi_{j,k}))$, we obtain for $0 < l < T$

$$|\bar{\bar{a}}_{j,k}(l)| \le K2^{j/2}\left(\frac{1}{l} + \frac{1}{T-l}\right). \tag{8.25}$$

□

Proof of Proposition 6.1. (i) We have

$$E\,I_T(\omega) = \int_{\Pi}\Phi_2^{(T)}(u) f(\omega+u)\,du,$$

where $\sup_T\left\{\int|\Phi_2^{(T)}(u)|\,du\right\} < \infty$, cf. Theorem 5.2.3 in [5] and (i) on the bottom of p. 166 in [6]. Therefore, we obtain

$$TV(E\,I_T(\omega)) \le \int_{\Pi}|\Phi_2^{(T)}(u)|\,du * TV(f) \le C,$$

which implies

$$\begin{aligned}
|E\tilde{\alpha}_{j,k} - E\bar{\alpha}_{j,k}| &\le \sum_{t=1}^{T}\int_{\omega_t-2\pi/T}^{\omega_t}|\psi_{j,k}(\omega)||E\,I_T(\omega) - E\,I_T(\omega_t)|\,d\omega \\
&\le C2^{j/2}T^{-1}TV(E\,I_T(\omega)|_{supp(\psi_{j,k})}).
\end{aligned}$$

(ii) Using (8.6) we get analogously to (22) in [18], that

$$var(\overline{\alpha}_{j,k}) = \frac{1}{(2\pi H_2^{(T)})^2} \sum_{t_1,t_2,t_3,t_4=1}^{T} \overline{a}_{j,k}(t_1-t_2)\overline{a}_{j,k}(t_3-t_4) * \tag{8.26}$$
$$* \{c(t_1,t_3)c(t_2,t_4) + c(t_1,t_4)c(t_2,t_3) + c(t_1,t_2,t_3,t_4)\},$$

whereas $var(\tilde{\alpha}_{j,k})$ can be written in the same way with $a_{j,k}(.)$ instead of $\overline{a}_{j,k}(.)$. This implies

$$|var(\overline{\alpha}_{j,k}) - var(\tilde{\alpha}_{j,k})|$$
$$\leq KT^{-2} \sum_{t_1,t_2,t_3,t_4=1}^{T} |\overline{a}_{j,k}(t_1-t_2)||\overline{a}_{j,k}(t_3-t_4) - a_{j,k}(t_3-t_4)| * \{\ldots\} \tag{8.27}$$
$$+ KT^{-2} \sum_{t_1,t_2,t_3,t_4=1}^{T} |\overline{a}_{j,k}(t_1-t_2) - a_{j,k}(t_1-t_2)||a_{j,k}(t_3-t_4)| * \{\ldots\}.$$

W.l.o.g. we estimate only the first term on the right-hand side of (8.27). Because of $|a_{j,k}(l)| + |\overline{a}_{j,k}(l)| \leq K2^{-j/2}$ we obtain by (A3)

$$\sum_{t_1,t_2,t_3,t_4=1}^{T} |\overline{a}_{j,k}(t_1-t_2)||\overline{a}_{j,k}(t_3-t_4) - a_{j,k}(t_3-t_4)||c(t_1,t_2,t_3,t_4)| = O(T2^{-j}). \tag{8.28}$$

We have by (i) of Lemma 3.1 and (iii) of Lemma 6.1 that

$$\sum_{t_4} |\overline{a}_{j,k}(t_3-t_4) - a_{j,k}(t_3-t_4)| \sum_{t_1} |c(t_1,t_3)| \sum_{t_2} |c(t_2,t_4)||\overline{a}_{j,k}(t_1-t_2)| = O(\log T) \tag{8.29}$$

holds uniformly in t_3, which implies

$$\sum_{t_3 \leq T^\delta \text{ or } t_3 \geq T-T^\delta} \sum_{t_1,t_2,t_4} |\overline{a}_{j,k}(t_1-t_2)||\overline{a}_{j,k}(t_3-t_4) - a_{j,k}(t_3-t_4)||c(t_1,t_3)||c(t_2,t_4)| = O(T^\delta \text{lo} \tag{8$$

Let now $T^\delta < t_3 < T - T^\delta$. Then we conclude by (i) of Lemma 3.1 and (i), (iii) of Lemma 6.1 that

$$|\overline{a}_{j,k}(t_3-t_4) - a_{j,k}(t_3-t_4)| \leq T^{-1/2} + 2^{j/2}T^{-\delta}$$

holds uniformly in t_4. This implies

$$\sum\nolimits_{T^\delta<t_3<T-T^\delta} \sum_{t_1} |c(t_1,t_3)| \sum_{t_2} |\overline{a}_{j,k}(t_1-t_2)| \sum_{t_4} |c(t_2,t_4)||\overline{a}_{j,k}(t_3-t_4) - a_{j,k}(t_3-t_4)|$$
$$= \sum_{T^\delta<t_3<T-T^\delta} O\left(2^{j/2}\log T\left(T^{-1/2} + 2^{j/2}T^{-\delta}\right)\right)$$
$$= O\left(\left(2^{j/2}T^{1/2} + 2^jT^{1-\delta}\right)\log T\right). \tag{8.31}$$

Choosing $\delta = \delta(T)$ such that $1 - \alpha < \delta < 1$ we get by (8.30) and (8.31)

$$\sum_{t_1,t_2,t_3,t_4} |\overline{a}_{j,k}(t_1-t_2)||\overline{a}_{j,k}(t_3-t_4) - a_{j,k}(t_3-t_4)||c(t_1,t_3)||c(t_2,t_4)| = o(T). \tag{8.32}$$

The term containing $c(t_1, t_4)$ and $c(t_2, t_3)$ instead of $c(t_1, t_3)$ and $c(t_2, t_4)$, respectively, can be estimated analogously, which yields in view of (8.28) and (8.32) assertion (ii).

(iii) The proof is analogous to that of (iii) of Proposition 3.1.

□

Acknowledgment. I thank Professor R. Dahlhaus for helpful discussions. Further, I thank R. von Sachs for a careful reading of the manuscript and pointing out some inaccuracies.

References

1. Beltrão, K. I. and Bloomfield, P. (1987). Determining the bandwidth of a kernel spectrum estimate. *J. Time Ser. Anal.* **8**, 21-38.
2. Bentkus, R. (1986). Rate of uniform convergence of statistical estimators of spectral density in spaces of differentiable functions. *Lithuanian Math. J* **26**, 209-219.
3. Bentkus, R. and Rutkauskas, V. (1973). On the asymptotic of the first two moments of second-order spectral estimates. *Lithuanian Math. J.* **13**, 19-30.
4. Bentkus, R. and Rudzkis, R. (1976). Large deviations for the spectral estimate of a stationary Gaussian sequence. *Lithuanian Math. J.* **16**, 519-529.
5. Brillinger, D. R. (1975). *Time Series: Data Analysis and Theory.* Holt, Rinehart and Winston: New York.
6. Dahlhaus, R. (1983). Spectral analysis with tapered data. *J. Time Ser. Anal.* **4**, 163-175.
7. Dahlhaus, R. (1990). Nonparametric high resolution spectral estimation. *Probab. Th. Rel. Fields* **85**, 147-180.
8. Davydov, Y. A. (1970). The invariance principle for stationary processes. *Theory Probab. Appl.* **14**, 487-498.
9. Donoho, D. L. and Johnstone, I. M. (1992). Minimax estimation via wavelet shrinkage. Technical Report No. 402, Department of Statistics, Stanford University.
10. Franke, J. and Härdle, W. (1992). On bootstrapping kernel spectral estimates. *Ann. Statist.* **20**, 121-145.
11. Gao, H.-Ye (1993). Wavelet estimation of spectral densities in time series analysis. Ph.D. dissertation. University of California, Berkeley.
12. Gasser, T. and Müller, H. G. (1979). Kernel estimation of regression functions. In *Smoothing techniques for curve estimation. Lecture Notes in Math.* **757**, 23-68.
13. Hurvich, C. M. (1985). Data-driven choice of a spectrum estimate: Extending the applicability of cross-validation methods. *J. Amer. Statist. Assoc.* **80**, 933-940.
14. Ibragimov, I. A. (1963). On estimation of the spectral function of a stationary Gaussian process. Theory Probab. Appl. **8**, 366-401.
15. Leonov, V. P. and Shiryaev, A. N. (1959). On a method of calculation of semi-invariants. *Theory Probab. Appl.* 4, 319-329.
16. Nemirovskii, A. S. (1985). Nonparametric estimation of smooth regression functions. *Izv. Akad. Nauk SSR Techn. Kibernet.* **3**, 50-60 (in Russian).
17. Nemirovskii, A. S., Polyak, B. T. and Tsybakov, A. B. (1985). Rate of convergence of nonparametric estimates of maximum-likelihood type. *Problems of Information Transmission* **21**, 258-272.
18. Rudzkis, R. (1978). Large deviations for estimates of spectrum of stationary series. *Lithuanian Math. J.* **18**, 214-226.
19. Rudzkis, R., Saulis, L. and Statulevicius, V. (1978). A general lemma on probabilities of large deviations. *Lithuanian Math. J.* **18**, 226-238.

20. Statulevicius, V. and Jakimavicius, D. (1988). Estimates of semiinvariants and centered moments of stochastic processes with mixing I. *Lithuanian Math. J.* **28**, 67-80.
21. Wahba, G. (1980). Automatic smoothing of the log periodogram. *J. Amer. Statist. Assoc.* **75**, 122-132.
22. Wahba, G. and Wold, S. (1975). Periodic splines for spectral density estimation: The use of cross-validation for determining the degree of smoothing. *Comm. Statist.* 4, (2), 125-144.

Recent publications of the Institut für Angewandte Analysis und Stochastik

Preprints 1993

70. Herbert Gajewski, Willi Jäger, Alexander Koshelev: About loss of regularity and "blow up" of solutions for quasilinear parabolic systems.

71. Friedrich Grund: Numerical solution of hierarchically structured systems of algebraic–differential equations.

72. Henri Schurz: Mean square stability for discrete linear stochastic systems.

73. Roger Tribe: A travelling wave solution to the Kolmogorov equation with noise.

74. Roger Tribe: The long term behavior of a Stochastic PDE.

75. Annegret Glitzky, Konrad Gröger, Rolf Hünlich: Rothe's method for equations modelling transport of dopants in semiconductors.

76. Wolfgang Dahmen, Bernd Kleemann, Siegfried Prößdorf, Reinhold Schneider: A multiscale method for the double layer potential equation on a polyhedron.

77. Hans–Günter Bothe: Attractors of non invertible maps.

78. Gregori Milstein, Michael Nussbaum: Autoregression approximation of a nonparametric diffusion model.

Preprints 1994

79. Anton Bovier, Véronique Gayrard, Pierre Picco: Gibbs states of the Hopfield model in the regime of perfect memory.

80. Roland Duduchava, Siegfried Prößdorf: On the approximation of singular integral equations by equations with smooth kernels.

81. Klaus Fleischmann, Jean–François Le Gall: A new approach to the single point catalytic super–Brownian motion.

82. Anton Bovier, Jean–Michel Ghez: Remarks on the spectral properties of tight binding and Kronig–Penney models with substitution sequences.

83. Klaus Matthes, Rainer Siegmund–Schultze, Anton Wakolbinger: Recurrence of ancestral lines and offspring trees in time stationary branching populations.

84. Karmeshu, Henri Schurz: Moment evolution of the outflow–rate from nonlinear conceptual reservoirs.

85. Wolfdietrich Müller, Klaus R. Schneider: Feedback stabilization of nonlinear discrete–time systems.

86. Gennadii A. Leonov: A method of constructing of dynamical systems with bounded nonperiodic trajectories.

87. Gennadii A. Leonov: Pendulum with positive and negative dry friction. Continuum of homoclinic orbits.

88. Reiner Lauterbach, Jan A. Sanders: Bifurcation analysis for spherically symmetric systems using invariant theory.

89. Milan Kučera: Stability of bifurcating periodic solutions of differential inequalities in $\mathbb{R}^3$.

90. Peter Knabner, Cornelius J. van Duijn, Sabine Hengst: An analysis of crystal dissolution fronts in flows through porous media Part I: Homogeneous charge distribution.

91. Werner Horn, Philippe Laurençot, Jürgen Sprekels: Global solutions to a Penrose–Fife phase–field model under flux boundary conditions for the inverse temperature.

92. Oleg V. Lepskii, Vladimir G. Spokoiny: Local adaptivity to inhomogeneous smoothness. 1. Resolution level.

93. Wolfgang Wagner: A functional law of large numbers for Boltzmann type stochastic particle systems.

94. Hermann Haaf: Existence of periodic travelling waves to reaction–diffusion equations with excitable–oscillatory kinetics.

95. Anton Bovier, Véronique Gayrard, Pierre Picco: Large deviation principles for the Hopfield model and the Kac–Hopfield model.

96. Wolfgang Wagner: Approximation of the Boltzmann equation by discrete velocity models.

97. Anton Bovier, Véronique Gayrard, Pierre Picco: Gibbs states of the Hopfield model with extensively many patterns.

98. Lev D. Pustyl'nikov, Jörg Schmeling: On some estimations of Weyl sums.