QUASILIKELIHOOD-BASED HIGHER-ORDER SPECTRAL ESTIMATION OF RANDOM FIELDS WITH POSSIBLE LONG-RANGE DEPENDENCE

V. V. ANH, N. N. LEONENKO, AND L. M. SAKHNO

Dedicated to Professor Chris Heyde on the occasion of his 65th birthday

Abstract. This paper provides a quasilikelihood/minimum contrast-type method for parameter estimation of random &elds in the frequency domain based on higher-order information. The estimation technique uses the spectral density of the general k-th order and allows for possible long-range dependence in the random &elds. To avoid bias due to edge effects, data tapering is incorporated in the method. The suggested minimum contrast functional is linear with respect to the periodogram of k-th order, hence kernel estimation for the spectral densities is not needed. Furthermore, discretisation is not required in the estimation of continuously observed random &elds. The consistency and asymptotic normality of the resulting estimators are established. Illustrative application of the method to some problems in mathematical &nance and signal detection will be indicated.

1. Introduction

Parameter estimation of random processes and &elds in the frequency domain has a long history and is an elaborated area of statistical inference. Many estimation methods have been available for different models of stationary processes and &elds with short or long-range dependence, but the majority of these methods rely on the information provided by the spectral density, that is, the second-order information only. We mention here some of the contributions which are most relevant to the approach of the present paper. These are the results on minimum distance estimation techniques and, in particular, on the Whittle estimators of Ibragimov (1963), Hannan (1970,1973), Dunsmuir and Hannan (1976), Guyon (1982), Rosenblatt (1985), Fox and Taqqu (1986), Bentkus and Maliukevicius (1988), Giraitis and Surgailis (1990), Heyde and Gay (1993), Giraitis and Taqqu 1999), Gao *et al.* (2001,2002). The related quasi-likelihood approach has been elucidated in Heyde (1997).

This paper is concerned with parameter estimation of random &elds in the frequency domain based on higher-order information. Statistical techniques relying on higher-order moments and cumulants and higher-order spectra are of increasing demand in many &elds of applications. The beginning of higher-order statistics can be traced back to Kolmogorov's work and those contributions in the 60s, such as Brillinger (1965) and Brillinger and Rosenblatt (1967a, 1967b), but it is only during the past two decades

Date: 25 May 2003.

Key words and phrases. Quasilikelihood estimation, minimum contrast estimator, higher-order spectral densities, long-range dependence.

Partially supported by the Australian Research Council grants A89601825 and A10024117, and the NATO grant PST.CLG.976361. The authors are grateful to a referee for many suggestions for improvement of the paper.

that the area has been rapidly expanding. The bibliography on higher-order statistics presented by Swami, Giannakis and Zhou (1997) listed 1759 papers, and these papers are only those related to statistical signal processing and engineering applications.

We mention here some contexts where higher-order statistics play an essential role. Firstly, various problems related to non-Gaussianity include detection and classi&cation of non-Gaussian signals, separation of additive mixtures of independent non-Gaussian signal and Gaussian noise, suppression of additive Gaussian noise, and tests for non-Gaussianity. Next are those problems related to nonlinearity such as spectral analysis of nonlinear processes, identi&cation of non-linear systems and tests for non-linearity, and modelling and analysis of chaotic systems. Another area treated with higher-order statistics is non-stationarity and, in particular, multiplicative noise models, useful for describing certain kinds of non-stationary behaviour. Higher-order information can also be used to obtain improved estimates, and, of course, to estimate those parameters which are not covered by second-order information. We again refer to Swami, Giannakis and Zhou (1997), where many further areas of application of higher-order statistics are indicated. Some recent references and examples can also be found in Anh, Leonenko and Sakhno (2002b).

It should be noted that, although a variety of techniques and algorithms based on higher-order statistics has been available for a wide range of applications, many practical problems remain challenging, and new problems continue to arise. At the same time, there exist an increasing need for the development of rigorous theories underpinning many ad-hoc treatments in practice, as well as the need for the development of new statistical methods based on higher-order information. We provide an approach here for such a method, which is based on an extension of the idea of quasilikelihood in such a way that the estimating function is generated from information on higher-order spectral densities.

This estimation technique is based on the spectral density of the general k-th order. The classes of processes and &elds for which this technique is applicable are described by the conditions on the spectral densities of the k-th order and the weight functions incorporated into the minimum contrast functional. This approach does not exclude the possibility of the process or &eld being long-range dependent. It should also be noted that our functional is linear with respect to the periodogram of k-th order, therefore we do not need to consider kernel estimation for the spectral densities. This approach will be presented in a uni&ed manner, suitable for estimation of random processes and &elds, continuous- and discrete-time settings, and with the use of data transformation by means of tapering. Furthermore, the method does not require discretisation in the estimation of continuous-time random processes and &elds based on continuously observed data. Section 2 contains the main results on the consistency and asymptotic normality of contrast estimators. A discussion and some examples are also provided to illustrate the method. The proofs are given in Section 3.

2. Results and discussion

We begin with the following assumption.

I. Let $Y(t)$, $t \in \mathbb{R}^n$, be a real-valued, measurable, strictly stationary random &eld with zero mean and spectral densities of order $k = 2, 3, ...$, that is, functions $f_k(\lambda_1, ..., \lambda_{k-1}) \in L_1\left(\mathbb{R}^{(k-1)n}\right)$ exist such that the cumulant function of the k-th order of the &eld $Y(t)$ is given by

$$c_k(t_1,...,t_{k-1}) = \int_{\mathbb{R}^{(k-1)n}} f_k(\lambda_1,...,\lambda_{k-1})\, e^{i\sum_{j=1}^{k-1}\lambda_j t_j} d\lambda_1...d\lambda_{k-1}. \tag{2.1}$$

We will assume that the spectral densities depend on an unknown parameter vector $\theta \in \Theta \subset \mathbb{R}^m$:

$$\begin{aligned} f_k(\lambda_1,...,\lambda_{k-1}) &= f_k(\lambda_1,...,\lambda_{k-1};\theta) \\ &= \operatorname{Re} f_k(\lambda_1,...,\lambda_{k-1};\theta) + i \operatorname{Im} f_k(\lambda_1,...,\lambda_{k-1};\theta) \\ &= f_k^{(1)}(\lambda_1,...,\lambda_{k-1};\theta) + i f_k^{(2)}(\lambda_1,...,\lambda_{k-1};\theta); \end{aligned}$$

the parametric set Θ is a compact set and the true value of the parameter vector $\theta_0 \in$ *int* Θ, the interior of Θ. Suppose further that $f_k(\lambda_1,...,\lambda_{k-1};\theta_1) \not\equiv f_k(\lambda_1,...,\lambda_{k-1};\theta_2)$ for $\theta_1 \neq \theta_2$ almost everywhere in $\mathbb{R}^{n(k-1)}$ with respect to the Lebesgue measure.

Suppose that we are given observations of the random &eld $Y(t)$ over the cube $[0,T]^n$, and we are interested in the estimation of θ which is based on the k-th order empirical information. For this purpose, we consider a generalization to the multidimensional case of the quasilikelihood/minimum contrast estimation procedure elaborated in Anh, Leonenko and Sakhno (2002b), which is based on the minimization of a certain empirical spectral functional of k-th order.

An essential problem we need to address in the case $n \geq 2$ is the bias problem due to edge effects, which features signi&cantly in the asymptotic properties of the corresponding estimators. We will use an approach based on tapered data (see Tukey 1967, Dahlhaus 1983, Dahlhaus and Künsch 1987, Guyon 1982, 1995 among many others), namely, we will base our analysis on the tapered values

$$\{h_T(t)Y(t), t \in [0,T]^n\},$$

where $h_T(t) = h(t/T)$, $t \in \mathbb{R}^n$. We will suppose that the taper function factorizes as

$$h(t) = \prod_{i=1}^{n} h_1(t_i), t_i \in \mathbb{R}^1,$$

and the measurable function $h_1(t) : \mathbb{R}^1 \to \mathbb{R}^1$ is of bounded variation, vanishes outside the interval $[0,1]$ and should be smooth with $h(0) = h(1) = 0$.

We de&ne the &nite Fourier transform of tapered data $\{h_T(t)Y(t), t \in [0,T]^n\}$, and the tapered periodogram of the k-th order, respectively, as

$$d_T^h(\lambda) = \int_{[0,T]^n} h_T(t) Y(t)\, e^{-i(\lambda,t)} dt,\ \lambda \in \mathbb{R}^n, \tag{2.2}$$

$$I_{k,T}^h(\lambda_1,...,\lambda_k) = \left((2\pi)^n H_{k,T}(0)\right)^{-1} \prod_{i=1}^{k} d_T^h(\lambda_i), \quad \lambda_i \in \mathbb{R}^n,\ i = 1,...,k, \tag{2.3}$$

where $\sum_{i=1}^{k} \lambda_i = 0$, but no proper subset of λ_i has sum 0. We have denoted here

$$H_{k,T}(\lambda) = \int_{[0,T]^n} (h_T(t))^k e^{-i(\lambda,t)} dt,\ \lambda \in \mathbb{R}^n,$$

and have supposed that $H_{k,T}(0) \neq 0$.

To simplify our exposition, we will use the following notation: We will write the spectral density of k-th order as $f_k(\lambda)$ or $f_k(\lambda;\theta)$, where $\lambda = (\lambda_1,...,\lambda_k)$ with $\lambda_k = -\sum_{j=1}^{k-1} \lambda_j$

and $\lambda_j \in \mathbb{R}^n$, $j = 1, ...k$; we denote $\lambda' = (\lambda_1, ..., \lambda_{k-1})$; the same notations will be applied to functions of k arguments $\lambda_1, ..., \lambda_k \in \mathbb{R}^n$ with $\lambda_k = -\sum_{j=1}^{k-1} \lambda_j$. Sometimes in the following we will also write the spectral density of k-th order as a function of k variables $f_k(\lambda_1, ..., \lambda_k)$, where $\lambda_k = -\sum_{j=1}^{k-1} \lambda_j$. Furthermore, where it does not cause any confusion, we will write $\int_{\mathbb{R}^{(k-1)n}} g(\lambda)\, d\lambda'$, where, again, $\lambda = (\lambda_1, ..., \lambda_k)$, $\lambda_k = -\sum_{j=1}^{k-1} \lambda_j$, $\lambda_j \in \mathbb{R}^n$, $j = 1, ...k$, and $\lambda' = (\lambda_1, ..., \lambda_{k-1})$. Sometimes we will also write such an integral in the form $\int_{\mathbb{R}^{(k-1)n}} g(\lambda)\delta\left(\sum_{i=1}^{k} \lambda_j\right) d\lambda'$, with $\delta(\cdot)$ being the Kronecker delta function.

If ν is a set of natural numbers, we will write $|\nu|$ to denote the number of elements in ν, and $\tilde{\nu}$ to denote the subset of ν which contains all the elements of ν except the last one. We will also deal with integrals of the form $\int_{\mathbb{R}^{(k-p)n}} g(u) \prod_{l=1}^{p} \delta\left(\sum_{j\in\nu_l} u_j\right) du'$, where $(\nu_1, ..., \nu_p)$ is a partition of the set $\{1, ..., k\}$. In such a case, integration is understood with respect to $(k-p)n$-dimensional vector u', obtained from the vector $u = (u_1, .., u_k)$ in view of p linear restrictions on k variables u_j.

Remark 1. *From the practical point of view, tapering can improve statistical analysis as it lessens the importance of observations close to the edges and, therefore, lessens any non-stationary effect which could be present on the boundary of the observation domain. An additional advantage of data tapering is reducing the leakage effect, hence providing better estimators for spectral densities with peaks. Furthermore, tapering provides a means to control bias for spectral estimators of the form* $\int_{\mathbb{R}^{n(k-1)}} I_{k,T}^{h}(\lambda)\varphi(\lambda)\, d\lambda'$ *with an appropriate choice of a taper* h *and a weight function* φ*.*

Although our exposition concentrates on continuous-time random &elds, the results stated below hold true for the discrete case. When the discrete case differs essentially from the continuous case, we will provide necessary details.

II. Let the real-valued functions $w_k^{(i)}(\lambda)$, $i = 1, 2$, $w_{k,0}(\lambda)$, $\lambda \in \mathbb{R}^{(k-1)n}$, and the spectral density of k-th order satisfy the following conditions:

(i) $w_k^{(i)}(\lambda)$, $i = 1, 2$, and $w_{k,0}(\lambda)$ satisfy the same conditions of symmetry as the k-th order spectral density;

(ii) $w_{k,0}(\lambda)$ is nonnegative and $w_{k,0}(\lambda) \equiv 0$ if any proper subset of λ_i has sum 0, that is, $w_{k,0}(\lambda) \equiv 0$ on all hyperplanes of the form $\sum_{i\in\nu} \lambda_i = 0$, where $\nu = \{i_1, ..., i_l\} \subset \{1, ..., k\}$ and $1 \le l < k$;

(iii) $w_k^{(i)}(\lambda)\, w_{k,0}(\lambda)\, f_k^{(i)}(\lambda; \theta) \in L_1\left(\mathbb{R}^{(k-1)n}\right)$, $i = 1, 2$, for all $\theta \in \Theta$;

(iv) $w_k^{(i)}(\lambda)\, f_k^{(i)}(\lambda; \theta) \ge 0$, $i = 1, 2$, $(\lambda; \theta) \in \mathbb{R}^{(k-1)n} \times \Theta$.

In what follows we will suppose $k \ge 3$; the case $k = 2$ will be outlined in Remark 3 below. Consider the following factorization of the real and imaginary parts of the spectral density $f_k(\lambda; \theta)$, $\lambda \in \mathbb{R}^{n(k-1)}$, $k \ge 3$:

$$\begin{aligned} f_k^{(i)}(\lambda; \theta)\, w_k^{(i)}(\lambda) &= \sigma_k^{(i)}(\theta)\, \psi_k^{(i)}(\lambda; \theta),\ i = 1, 2,\ \theta \in \Theta, \\ \lambda &= (\lambda_1, ..., \lambda_k),\ \lambda_k = -\sum_{j=1}^{k-1} \lambda_j,\ \lambda_j \in \mathbb{R}^n,\ j = 1, ..., k, \end{aligned} \tag{2.4}$$

with

$$\sigma_k^{(i)}(\theta) = \int_{\mathbb{R}^{n(k-1)}} f_k^{(i)}(\lambda; \theta)\, w_k^{(i)}(\lambda)\, w_{k,0}(\lambda)\, d\lambda',\ i = 1, 2 \tag{2.5}$$

and

$$\psi_k^{(i)}(\lambda;\theta) = \frac{f_k^{(i)}(\lambda;\theta)\, w_k^{(i)}(\lambda)}{\sigma_k^{(i)}(\theta)} \tag{2.6}$$

so that

$$\int_{\mathbb{R}^{n(k-1)}} \psi_k^{(i)}(\lambda;\theta)\, w_{k,0}(\lambda)\, d\lambda' = 1. \tag{2.7}$$

We additionally suppose

III. The derivatives $\nabla_\theta \psi_k^{(i)}(\lambda;\theta)$, $i = 1, 2$, exist and

$$\nabla_\theta \int_{\mathbb{R}^{(k-1)n}} \psi_k^{(i)}(\lambda;\theta)\, w_{k,0}(\lambda)\, d\lambda' = \int_{\mathbb{R}^{(k-1)n}} \nabla_\theta \psi_k^{(i)}(\lambda;\theta)\, w_{k,0}(\lambda)\, d\lambda' = 0,\ i = 1, 2. \tag{2.8}$$

In the following, where it does not cause any confusion, we will omit the index k in the functions $w_k^{(i)}$ and $w_{k,0}$.

We next introduce the definition of quasilikelihood/minimum contrast estimators.

Definition 1. *A nonrandom real-valued function* $K(\theta_0;\theta) \geq 0$ *is called a* contrast function *if it has a unique minimum at* $\theta = \theta_0$. *A random field* $U_T(\theta)$, $\theta \in \Theta$, *related to the observation* $\{Y(t), t \in [0,T]^n\}$ *is called the* contrast field *for a contrast function* $K(\theta_0;\theta)$ *if it satisfies the following inequality:*

$$\liminf_{T\to\infty} \left[U_T(\theta) - U_T(\theta_0)\right] \geq K(\theta_0;\theta) \tag{2.9}$$

in P_0*-probability, where* $P_0 = P_{\theta_0}$, *a member of the family of distributions* $\{P_\theta, \theta \in \Theta\}$. *The* mimimum contrast estimator $\hat{\theta}_T$ *is defined as a minimum point of the functional* $U_T(\theta)$, *that is,*

$$\hat{\theta}_T = \arg\min_{\theta\in\Theta} U_T(\theta). \tag{2.10}$$

For the random field $Y(t)$ defined above, we consider the contrast field based on the tapered periodogram of k-th order:

$$\begin{aligned} U_T^h(\theta) = U_{k,T}^h(\theta) &= -\left(p \int_{\mathbb{R}^{n(k-1)}} \operatorname{Re} I_{k,T}^h(\lambda)\, w^{(1)}(\lambda)\, w_0(\lambda) \log \psi_k^{(1)}(\lambda;\theta)\, d\lambda' \right.\\ &\left. + q \int_{\mathbb{R}^{n(k-1)}} \operatorname{Im} I_{k,T}^h(\lambda)\, w^{(2)}(\lambda)\, w_0(\lambda) \log \psi_k^{(2)}(\lambda;\theta)\, d\lambda' \right), \end{aligned} \tag{2.11}$$

with nonnegative numbers p and q satisfying $p + q = 1$. Denote

$$\begin{aligned} \mathcal{K}(\theta_0;\theta) = \mathcal{K}_k(\theta_0;\theta) &= p \int_{\mathbb{R}^{n(k-1)}} f_k^{(1)}(\lambda;\theta_0) \log \frac{\psi_k^{(1)}(\lambda;\theta_0)}{\psi_k^{(1)}(\lambda;\theta)} w^{(1)}(\lambda)\, w_0(\lambda)\, d\lambda' \\ & + q \int_{\mathbb{R}^{n(k-1)}} f_k^{(2)}(\lambda;\theta_0) \log \frac{\psi_k^{(2)}(\lambda;\theta_0)}{\psi_k^{(2)}(\lambda;\theta)} w^{(2)}(\lambda)\, w_0(\lambda)\, d\lambda' \end{aligned} \tag{2.12}$$

and

$$\begin{aligned} U(\theta) = U_k(\theta) &= -\left(p \int_{\mathbb{R}^{k-1}} f_k^{(1)}(\lambda;\theta_0)\, w^{(1)}(\lambda)\, w_0(\lambda) \log \psi_k^{(1)}(\lambda;\theta)\, d\lambda' \right.\\ &\left. + q \int_{\mathbb{R}^{k-1}} f_k^{(2)}(\lambda;\theta_0)\, w^{(2)}(\lambda)\, w_0(\lambda) \log \psi_k^{(2)}(\lambda;\theta)\, d\lambda' \right). \end{aligned} \tag{2.13}$$

We will need the following assumptions. Denote $\varphi_i(\lambda) = w^{(i)}(\lambda) \log \psi_k^{(i)}(\lambda)$.
IV. For all $\theta \in \Theta$, the functions $G_k(u; \varphi_i, w_0)$, $i = 1, 2$, are bounded and continuous at the point $u = 0$, where $G_k(u; \varphi, w)$ is de&ned by the formula

$$
\begin{aligned}
G_k(u) &= G_k(u_1, ..., u_k; \varphi, w) \\
&= \sum_{\nu=(\nu_1,...,\nu_p)} \int_{\mathbb{R}^{(k-p)n}} \prod_{l=1}^{p} f_{|\nu_l|}(\lambda_j + u_j, j \in \tilde{\nu}_l) \\
&\quad \times \varphi(\lambda) w(\lambda) \prod_{l=1}^{p-1} \delta\left(\sum_{j \in \nu_l} (\lambda_j + u_j) \right) \delta\left(\sum_{i=1}^{k} \lambda_i \right) d\lambda'.
\end{aligned} \tag{2.14}
$$

To formulate some further assumptions, we introduce for $m = 2, 3, ...$ the functions

$$G_{km}(u) = G_{km}(u; \varphi_1, ..., \varphi_m, \psi)$$

$$= \sum_{\nu=(\nu_1,...,\nu_p)} \int_{\mathbb{R}^{((k-1)m-p+1)n}} \prod_{i=1}^{m} \varphi_i\left(\lambda_{(i-1)k+1}, ..., \lambda_{ik}\right)$$

$$\times \prod_{i=1}^{m} \psi\left(\lambda_{(i-1)k+1}, ..., \lambda_{ik}\right) f_{|\nu_1|}(u_j + \lambda_j, j \in \tilde{\nu}_1) \times ... \times f_{|\nu_p|}(u_j + \lambda_j, j \in \tilde{\nu}_p)$$

$$\times \prod_{l=1}^{p-1} \delta\left(\sum_{j \in \nu_l} (u_j + \lambda_j) \right) \prod_{i=1}^{m} \delta\left(\lambda_{(i-1)k+1} + ... + \lambda_{ik}\right) d\lambda', \tag{2.15}$$

where the summation is taken over all indecomposable partitions $\nu = (\nu_1, ..., \nu_p)$ of the table

$$
\begin{array}{lll}
1 & ... & k \\
k+1 & ... & 2k \\
... & ... & ... \\
m(k-1)+1 & ... & mk
\end{array}
$$

that is, over those partitions ν of the elements of the above table into disjoint sets in which there exist no sets $\nu_{i_1}, ..., \nu_{i_n}$ $(n < p)$ such that for some rows $r_{j_1}, ..., r_{j_m}$ $(m < k)$ of the table, the following equality holds: $r_{j_1} \cup ... \cup r_{j_m} = \nu_{i_1} \cup ... \cup \nu_{i_n}$.
V. The functions $G_{2k}(u; \varphi_i, \varphi_i, w_0)$, $i = 1, 2$,
(i) are bounded;
(ii) are continuous at the point $u = 0$.
VI. There exist nonnegative functions $v_1(\lambda)$ and $v_2(\lambda)$ such that
(i) the functions

$$a_k^{(i)}(\lambda; \theta) = v_i(\lambda) \log \psi_k^{(i)}(\lambda; \theta), \ i = 1, 2,$$

are uniformly continuous in $\mathbb{R}^{(k-1)n} \times \Theta$;
(ii) the functions $G_k\left(u; \frac{w^{(i)}}{v_i}, w_0\right)$, $i = 1, 2$, are bounded and continuous at $u = 0$ and the functions $G_{2k}\left(u; \frac{w^{(i)}}{v_i}, \frac{w^{(i)}}{v_i}, w_0\right)$, $i = 1, 2$, are bounded.
VII. $w_k^{(1)}(\lambda) \operatorname{Re} I_k^T(\lambda) \geq 0$, $w_k^{(2)}(\lambda) \operatorname{Im} I_k^T(\lambda) \geq 0$.

Theorem 1. *Let the random field* $Y(t)$, $t \in \mathbb{R}^n$, *satisfy the assumptions I, II, IV, V(i), VI and VII. Then the function* $\mathcal{K}(\theta_0;\theta)$ *given by (2.12) is the contrast function for the contrast field* $U_T^h(\theta)$ *given by (2.11). The resulting minimum contrast estimator* $\hat{\theta}_T$ *is then a consistent estimator of the parameter vector* θ, *that is,* $\hat{\theta}_T \to \theta_0$ *in* P_0*-probability as* $T \to \infty$, *and the estimators*

$$\hat{\sigma}_{k,T}^{(1)} = \int_{\mathbb{R}^{(k-1)n}} \operatorname{Re} I_{k,T}^h(\lambda)\, w^{(1)}(\lambda)\, w_0(\lambda)\, d\lambda' \tag{2.16}$$

and

$$\hat{\sigma}_{k,T}^{(2)} = \int_{\mathbb{R}^{(k-1)n}} \operatorname{Im} I_{k,T}^h(\lambda)\, w^{(2)}(\lambda)\, w_0(\lambda)\, d\lambda' \tag{2.17}$$

are consistent estimators of $\sigma_k^{(1)}(\theta)$ *and* $\sigma_k^{(2)}(\theta)$, *respectively.*

Remark 2. *It should be noted that Theorem 1 actually holds with the use of the untapered periodogram* ($h(t)=1$) *in the functional (2.11) for the multidimensional case as well as for the one-dimensional case. However, in order to state the result on asymptotic normality of the estimator (2.10), tapering is essential (or another adjustment of the periodogram is needed such as constructing the k-th order periodogram by means of unbiased estimators of the moments of second and higher orders).*

We will need some further conditions to state the result on asymptotic normality of the estimator $\hat{\theta}_T$.

VIII. The functions $\psi_k^{(i)}(\lambda;\theta)$, $i=1,2$, are twice differentiable in the neighborhood of the point θ_0 and the functions

$$\varphi_l^{ij}(\lambda;\theta) = w^{(l)}(\lambda)\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log\psi_k^{(l)}(\lambda;\theta), \quad i,j=1,...,m,\ l=1,2,\ \theta\in\Theta, \tag{2.18}$$

$$\begin{aligned} g_k^{(i)}(\lambda;\theta) &= w^{(1)}(\lambda)\frac{\partial}{\partial\theta_i}\log\psi_k^{(1)}(\lambda;\theta), \quad i=1,...,m,\ \theta\in\Theta, \\ g_k^{(i+m)}(\lambda;\theta) &= w^{(2)}(\lambda)\frac{\partial}{\partial\theta_i}\log\psi_k^{(2)}(\lambda;\theta), \quad i=1,...,m,\ \theta\in\Theta \end{aligned} \tag{2.19}$$

are such that

(i) the functions $G_k(u;\varphi_l^{ij},w_0)$, $i,j=1,...,m,\ l=1,2$, are bounded and continuous at $u=0$ for all $\theta\in\Theta$;

(ii) the functions $G_{2k}(u;\varphi_l^{ij},\varphi_l^{ij},w_0)$, $i,j=1,...,m,\ l=1,2$, are bounded for all $\theta\in\Theta$;

(iii) the functions $G_{kl}(u;g_k^{(m_1)},\ ...\ ,g_k^{(m_l)},w_0)$, are bounded for all $\theta\in\Theta$, $l=2,3,...$ and all choices of $(m_1,\ ...\ ,m_l)$, $1\le m_i\le 2m$, $i=1,...,l$.

The most essential assumption needed for the extension of the result on asymptotic normality of Anh, Leonenko and Sakhno (2002b) to the multidimensional case is the following assumption.

IX. The following convergence holds:

$$\begin{aligned} T^{n/2}\Bigg(E\int_{\mathbb{R}^{n(k-1)}} & I_{k,T}^h(\lambda)\, g_k^{(i)}(\lambda;\theta)\, w_0(\lambda)\, d\lambda' \\ & - \int_{\mathbb{R}^{n(k-1)}} f_k(\lambda)\, g_k^{(i)}(\lambda;\theta)\, w_0(\lambda)\, d\lambda'\Bigg) \to 0 \quad \text{as } T\to\infty\ \ \forall\theta\in\Theta \end{aligned} \tag{2.20}$$

with the functions $g_k^{(i)}(\lambda;\theta)$ defined by (2.19).

X. The matrices $S_k(\theta) = \left\{ s_{ij}^{(k)}(\theta) \right\}_{i,j=1,\dots,m}$ and $A_k(\theta) = \left\{ a_{ij}^{(k)}(\theta) \right\}_{i,j=1,\dots,m}$ are positive de&nite, where

$$
\begin{aligned}
(2.21)\quad s_{ij}^{(k)}(\theta) &= p \int_{\mathbb{R}^{(k-1)n}} f_k^{(1)}(\lambda;\theta)\, \varphi_1^{ij}(\lambda;\theta)\, w_0(\lambda)\, d\lambda' \\
&\quad + q \int_{\mathbb{R}^{(k-1)n}} f_k^{(2)}(\lambda;\theta)\, \varphi_2^{ij}(\lambda;\theta)\, w_0(\lambda)\, d\lambda' \\
&= \sigma_k^{(1)}(\theta)\, p \int_{\mathbb{R}^{(k-1)n}} \left(\frac{\partial^2}{\partial\theta_i \partial\theta_j} \psi_k^{(1)} - \frac{1}{\psi_k^{(1)}} \frac{\partial}{\partial\theta_i} \psi_k^{(1)} \frac{\partial}{\partial\theta_j} \psi_k^{(1)} \right) d\lambda' \\
&\quad + \sigma_k^{(2)}(\theta)\, q \int_{\mathbb{R}^{(k-1)n}} \left(\frac{\partial^2}{\partial\theta_i \partial\theta_j} \psi_k^{(2)} - \frac{1}{\psi_k^{(2)}} \frac{\partial}{\partial\theta_i} \psi_k^{(2)} \frac{\partial}{\partial\theta_j} \psi_k^{(2)} \right) d\lambda',
\end{aligned}
$$

$$
\begin{aligned}
a_{ij}^{(k)}(\theta) &= \frac{1}{2} \Big\{ p^2 \operatorname{Re} G_{2k}\left(0; g_k^{(i)}, g_k^{(j)}, w_0\right) + q^2 \operatorname{Re} G_{2k}\left(0; g_k^{(i+m)}, g_k^{(j+m)}, w_0\right) \\
(2.22)\qquad &\quad + pq \operatorname{Im} G_{2k}\left(0; g_k^{(i+m)}, g_k^{(j)}, w_0\right) - pq \operatorname{Im} G_{2k}\left(0; g_k^{(i)}, g_k^{(j+m)}, w_0\right) \Big\}.
\end{aligned}
$$

Theorem 2. *Let the assumptions I - X be satis&ed. Then, as $T \to \infty$,*

$$
(2.23)\qquad T^{n/2}\left(\widehat{\theta}_T - \theta_0\right) \xrightarrow{\mathcal{D}} N_m\left(0, e(h) S_k^{-1}(\theta_0)\, A_k(\theta_0)\, S_k^{-1}(\theta_0)\right),
$$

where $N_m(\cdot,\cdot)$ is the m-dimensional normal distribution, the matrices $S_k(\theta)$ and $A_k(\theta)$ are given by the formulae (2.21) and (2.22) respectively and the tapering factor is of the form

$$
(2.24)\qquad e(h) = \left(\int (h_1(t))^{2k} dt \left(\int (h_1(t))^k dt \right)^{-2} \right)^n .
$$

Let us now discuss some particular features of the estimation method developed above and some aspects of its application.

Remark 3. *For the case of second-order spectral density, we suppose the existence of a real-valued, nonnegative function $w(\lambda)$, $\lambda \in \mathbb{R}^n$, symmetric about zero, such that $w(\lambda) f(\lambda;\theta) \in L_1(\mathbb{R}^n)$ $\forall \theta \in \Theta$, and we can introduce the factorization*

$$
f_2(\lambda;\theta) = \sigma^2(\theta)\, \psi(\lambda;\theta), \quad \lambda \in \mathbb{R}^n, \theta \in \Theta,
$$

where

$$
\sigma^2(\theta) = \int_{\mathbb{R}^n} f_2(\lambda;\theta)\, w(\lambda)\, d\lambda
$$

and

$$
\int_{\mathbb{R}^n} \psi(\lambda;\theta)\, w(\lambda)\, d\lambda = 1.
$$

The contrast &eld in this case is of the form

$$
(2.25)\qquad U_{2,T}(\theta) = -\int_{\mathbb{R}^n} I_{2,T}(\lambda)\, w(\lambda) \log \psi(\lambda;\theta)\, d\lambda,
$$

where $I_{2,T}(\lambda)$ is a tapered periodogram of the second order given by the formula (2.3) with $k = 2$ (for the case $n = 1$ we can also use the untapered periodogram). Equivalently, we can use in (2.25) an unbiased periodogram of the second order, constructed with the use of *an unbiased estimator of the correlation function (see, for example, Ivanov and Leonenko 1989). The conditions for consistency and asymptotic normality, I-X, can be rewritten for the case $k = 2$ (see Anh, Leonenko and Sakhno 2002a, where these conditions were presented for the case of processes ($n = 1$), and Anh, Leonenko and Sakhno 2003, where the case of Gaussian fields was considered).*

Remark 4. *Let us consider the reasons for introduction of the weight functions $w(\lambda)$ and $w_k^{(i)}(\lambda)$, $k \geq 3$, $i = 1, 2$, used in the contrast fields (2.25) and (2.11), respectively. Firstly, these functions are supposed to compensate for possible singularities of the spectral densities, that is, to control the behaviour of the integrands in (2.25) and (2.11) (and related integrals) at the points of singularities. In other words, introducing w, we will have scope for a trade-off between smoothness of weight functions and that of spectral densities. By means of weight functions, the high frequencies can be weighted down (or cut off).*
On the other hand, for the continuous-time case, when dealing with spectral densities with their frequencies defined on $\mathbb{R}^n$, introduction of a weight function will guarantee the convergence of corresponding integrals over infinite domains. This idea, and the contrast field of the type (2.25), were used in Leonenko and Moldavs'ka (1999) where, for the estimation of random fields with spectral density $f_2(\lambda;\theta) \in L_2(\mathbb{R}^n)$, the weight function of the form $w(\lambda) = \frac{1}{1+|\lambda|^2}$ was chosen (see also Ibragimov 1967 for the case of processes). However, in some cases, it may happen that a weight function is not required. For example, when considering the discrete case and "well behaved" spectral densities, the above conditions for consistency and asymptotic normality may be satisfied without any weight function. It should be noted that the idea of introducing a weight function into a contrast process for the case of continuous-time stochastic processes (observed continuously) was also used in Gao et al. *(2002), where a continuous version of the Gauss-Whittle contrast function (with the weight function $w(\lambda) = \frac{1}{1+\lambda^2}$) was considered.*

Remark 5. *Continuing our observation concerning weight functions, let us consider again the functional (2.25). We note the following fact: if there exists a smooth function $v : \mathbb{R}^n \to \mathbb{R}^n$ with Fourier transform $\hat{v}$ such that $w(\lambda) = |\hat{v}(\lambda)|^2$, then the product $f_2(\lambda)w(\lambda)$ can be viewed as the spectral density of the random field obtained from the original one by linear filtering with transfer function $\hat{v}(\lambda)$ (i.e., impulse response function $v(\lambda)$). Here we find a parallel situation with Heyde and Gay (1993) where the asymptotics for the smoothed periodogram were derived based on filtering of the original field, which may not have a square integrable spectral density, to produce a related one for which the spectral density is square integrable, and some standard results on the asymptotics of corresponding covariances can be obtained. The set of assumptions to be satisfied by the spectral density and smoothing function was tailored to implement this idea.*

Remark 6. *The proofs of our main results rely on large sample properties of the empirical spectral functionals of the form $\int_{\mathbb{R}^{n(k-1)}} I_{k,T}^h(\lambda)\,\varphi(\lambda)\,d\lambda'$ (see Lemmas 1 and 2 below), which can be deduced from the representation of the cumulants of these functionals in the form of singular integrals. This technique was elaborated in Bentkus (1972a, 1972b) and Bentkus and Rutkauskas (1973), for the case of untapered data. However, in our approach we formulate more general conditions on the second- and higher-order spectra as well as on the weight function φ (see Lemmas 1 and 2 below). These conditions are formulated in such a way that the spectral densities and weight function are treated simultaneously. Here again the conditions (for $k = 2$) are parallel to those of the second-order* *spectral density and smoothing function in Heyde and Gay (1993) (see conditions (A) and Theorem 1 therein), although our set of conditions has been deduced from a different technique.*

Remark 7. *It should be noted that the contrast function $\mathcal{K}(\theta_0;\theta)$, given by (2.12), takes origin from the Kullback-Leibler divergence defined as*

$$D_{KL}(p,q) = \int_\Gamma p(x)\log\left(\frac{p(x)}{q(x)}\right)d\mu(x),$$

where p,q belong to the set P of all probability densities given on (Γ,μ) with μ being a σ-finite measure on Γ, that is,

$$P = \left\{p:\Gamma\to\mathbb{R}^1;\ p(x)\geq 0,\int_\Gamma p(x)\,d\mu(x)=1\right\}.$$

Remark 8. *Let us consider a modification of the technique elaborated above with the intention to apply it to some particular models of financial processes. It has been known that financial data often display a characteristic that the data appear uncorrelated but the absolute values or the squares of the data exhibit long-range dependence (LRD). Heyde (1999) proposed a risky asset model with strong dependence through fractal-activity time. Under this model the price P_t at time t of a risky asset is given by*

$$P_t = P_0\exp\{\mu t+\sigma W(T_t)\},$$

where $\mu,\sigma^2>0$ and $\{T_t\}$ is a positive increasing stochastic process with stationary differences independent of the Brownian motion $W(t)$ and the differences $\tau_t = T_t - T_{t-1}$ of the process $\{T_t\}$ are long-range dependent. Then, under some assumptions,

$$X_t = \log P_t - \log P_{t-1} = \mu+\sigma(T_t-T_{t-1})^{1/2}W(1) \tag{2.26}$$

in distribution. Also, for $k=1,2,...$ and centered variables,

$$\begin{aligned} cov(X_t,X_{t+k}) &= 0,\\ cov(|X_t|,|X_{t+k}|) &= \sigma^2 cov\left(\tau_t^{1/2},\tau_{t+k}^{1/2}\right),\\ cov(X_t^2,X_{t+k}^2) &= 3\sigma^4 cov(\tau_t,\tau_{t+k}), \end{aligned} \tag{2.27}$$

the last covariance being defined if $E\tau_t^2<\infty$. Thus, the LRD of $\{|X_t|\}$ and $\{X_t^2\}$ follows from that of $\{\tau_t^{1/2}\}$ and $\{\tau_t\}$, respectively. The parameters of the long-range dependent process $\{\tau_t\}$ cannot be estimated from the second-order information of $\{X_t\}$, but they can be estimated, for example, from the second-order information of $\{X_t^2\}$, which is the fourth-order information of the process $\{X_t\}$. From (2.26) we can write down the following relationship between the spectral density of fourth order, $f_4^X(\lambda_1,\lambda_2,\lambda_3)$, of the process X_t, $t=1,2,\ldots$, which gives the aggregated returns for the log price process over intervals of unit length, $X_t = \log P_t - \log P_{t-1} = \int_{t-1}^t d\log P_s$, and the spectral density of second order, $f_2^\tau(\lambda)$, of the process τ_t:

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f_4^X(\lambda_1,\lambda_2-\lambda_1,\lambda_3)d\lambda_1 d\lambda_3 = 3\sigma^2 f_2^\tau(\lambda_2).$$

If we are interested in the estimation of an unknown parameter vector θ of the process τ_t, the above relation suggests considering the following contrast function:

$$U_{4,T}^*(\theta) = -\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} I_{4,T}^X(\lambda_1,\lambda_2,\lambda_3)w^*(\lambda_1+\lambda_2)\log\psi_\tau(\lambda_1+\lambda_2;\theta)d\lambda_1 d\lambda_2 d\lambda_3, \tag{2.28}$$

where $I^X_{4,T}$ is the periodogram of fourth order constructed from observations of the process X_t, $t = 1, \ldots, T$, and the function ψ_τ comes from the following factorization of the second-order spectral density of the process τ_t:

$$f_2^\tau(\lambda, \theta) = \sigma^2(\theta)\psi_\tau(\lambda, \theta),$$

$$\sigma^2(\theta) = \int_{-\pi}^{\pi} f_2^\tau(\lambda, \theta)w^*(\lambda)d\lambda, \qquad \psi_\tau(\lambda, \theta) = \frac{f_2^\tau(\lambda, \theta)}{\sigma^2(\theta)}.$$

In view of the equality in distribution (2.26), the conditions needed for the convergence

$$\begin{aligned} U^*_{4,T}(\theta) \to U^*_4(\theta) &= -\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f_4^X(\lambda_1, \lambda_2, \lambda_3)w^*(\lambda_1+\lambda_2)\log\psi_\tau(\lambda_1+\lambda_2;\theta)d\lambda_1 d\lambda_2 d\lambda_3 \\ &= -3\sigma^2\int_{-\pi}^{\pi} f_2^\tau(\lambda;\theta_0)w^*(\lambda)\log\psi_\tau(\lambda;\theta)d\lambda \end{aligned} \tag{2.29}$$

in probability and uniformly in $\theta \in \Theta$ can be formulated in terms of the conditions on the spectral densities of the process τ_t and the function w^*. Under this set of conditions, we will have

$$\begin{aligned} U^*_{4,T}(\theta) - U^*_{4,T}(\theta_0) &\to U^*_4(\theta) - U^*_4(\theta_0) \\ &= 3\sigma^2\int_{-\pi}^{\pi} f_2^\tau(\lambda;\theta_0)w^*(\lambda)\log\frac{\psi_\tau(\lambda;\theta_0)}{\psi_\tau(\lambda;\theta)}d\lambda \\ &= K^*(\theta_0;\theta) \geq 0, \end{aligned}$$

and, hence, the minimum contrast estimator based on the functional (2.28) will be consistent. Therefore, for the process τ_t (which is not observed directly), one can obtain a consistent estimator of its parameter vector θ, based on the contrast process (2.28), which is constructed with the use of fourth-order empirical information on the observable process X_t. Conditions for asymptotic normality can also be formulated in terms of spectral densities of the process τ_t.

In an analogous way to the above, we can construct a procedure for the estimation of the stochastic volatility model proposed in Anh, Heyde and Leonenko (2002), namely, the model for the evolution of an asset price,

$$dx(t) = (\mu + \beta\sigma^2(t))dt + \sigma(t)dW(t), \quad t \geq 0,$$

with $x(t)$ being a log-price process, $W(t)$ Brownian motion, and the volatility process $\sigma(t)$ following the stationary model

$$\sigma^2(t) = \int_{-\infty}^{t} G(t-s)dL(s) \geq 0, \tag{2.30}$$

where L is the Lévy process such that the distribution of $L(1)$ is nonnegative, and G is a memory function such that $\int_0^t G^2(s)ds < \infty$. Here again, for aggregated returns $y_n = \int_{(n-1)\Delta}^{n\Delta} dx(t)$, we will have $cov(y_t, y_{t+k}) = 0$, but $cov(y_t^2, y_{t+k}^2) \neq 0$, and, moreover, $cov(y_t^2, y_{t+k}^2)$ can be expressed in terms of the covariances of the process $\sigma^2(t)$ given by (2.30). Parametric families of spectral densities for the processes of the form (2.30) and corresponding choices of G and L have been provided in Anh, Heyde and Leonenko (2002). The functional analogous to (2.28) can be used for the estimation of parameters of the model (2.30) for the (unobservable) process $\sigma^2(t)$.

Remark 9. *Several further examples in which the use of higher-order information is needed for statistical inference and where use of the estimation technique developed in this section can be suggested are presented in Anh, Leonenko and Sakhno (2002b) (for the case* $n = 1$*). Some of these examples can be extended to the case of random fields* ($n \geq 2$)*, namely, in the estimation of stationary non-Gaussian linear fields and also in signal detection models of the form*

$$Y(t) = X_\theta(t) + N_\zeta(t), \quad t \in \mathbb{R}^n,$$

which is an additive mixture of independent strictly stationary non-Gaussian field $X_\theta(t)$ *and stationary Gaussian field* $N_\zeta(t)$*. Such models appear quite often in applications, for example, in image analysis. Assume that the spectral densities* $f_k^X(\lambda_1, ..., \lambda_{k-1}, \theta)$, $k \geq 2$ *of the field* $X_\theta(t)$ *exist. The spectral densities* $f_k^N(\lambda_1, ..., \lambda_{k-1}, \zeta)$ *of the Gaussian field* $N_\zeta(t)$ *exist and* $f_k^N \equiv 0$ *for* $k \geq 3$*. Therefore, the spectral densities of the field* $Y(t)$ *are of the form*

$$f_2^Y(\lambda, \theta, \zeta) = f_2^X(\lambda, \theta) + f_2^N(\lambda, \zeta),$$

$$f_k^Y(\lambda_1, ..., \lambda_{k-1}, \theta) = f_k^X(\lambda_1, ..., \lambda_{k-1}, \theta), \quad k \geq 3.$$

Observe that the spectral densities $f_k^Y, k \geq 3$ *do not depend on the parameter* ζ*. If we are interested in estimation of the parameter vector* θ *of* $X_\theta(t)$*, it seems reasonable to use the analytical information on* θ *contained in the higher-order spectral densities* $f_k^Y(\lambda_1, ..., \lambda_{k-1}, \theta)$*, and apply Theorems 1 and 2 to construct the minimum contrast estimator* $\widehat{\theta}_T$*, which is consistent and asymptotically normal. For example, we may choose* $k = 3$ *in Theorems 1 and 2. We may also use Theorems 1 and 2 for hypothesis testing.*

3. Proofs

The proofs of the results of the present paper are analogous to those of Anh, Leonenko and Sakhno (2002b). Hence we only present an outline of the proofs here. We commence with the basic ideas and facts needed for the proofs.

Firstly, we note that the following formula for the cumulants of the finite Fourier transform $d_T^h(\lambda)$, $\lambda \in \mathbb{R}^n$, can be deduced:

$$\begin{aligned} cum\left(d_T^h(\alpha_1), ..., d_T^h(\alpha_k)\right) &= \int_{[0,T]^{kn}} h_T(t_1)...h_T(t_k) e^{-i\Sigma_1^k(\alpha_j, t_j)} \\ &\quad \times cum\left(Y(t_1), ..., Y(t_k)\right) dt_1...dt_k \\ &= \int_{\mathbb{R}^{n(k-1)}} f_k(\gamma_1, ..., \gamma_{k-1}) \int_{[0,T]^{kn}} h_T(t_1)...h_T(t_k) \\ &\quad \times \exp\{i(t_1, \gamma_1 - \alpha_1) + ... + i(t_{k-1}, \gamma_{k-1} - \alpha_{k-1})\} \\ &\quad \times \exp\left\{i\left(t_k, -\sum_{j=1}^{k-1}\gamma_j - \alpha_k\right)\right\} dt_1...dt_k\, d\gamma_1...d\gamma_{k-1} \\ &= \int_{\mathbb{R}^{n(k-1)}} f_k(\gamma_1, ..., \gamma_{k-1}) H_{1,T}(\gamma_1 - \alpha_1)...H_{1,T}(\gamma_{k-1} - \alpha_{k-1}) \\ &\quad \times H_{1,T}\left(-\sum_1^{k-1}\gamma_j - \alpha_k\right) d\gamma_1...d\gamma_{k-1}, \end{aligned} \tag{3.1}$$

where

$$H_{1,T}(\lambda) = \int_{[0,T]^n} h_T(t) e^{-i(t,\lambda)} dt. \tag{3.2}$$

If $\sum_{j=1}^{k} \lambda_j = 0$, and $H_{k,T}(0) \neq 0$, then

$$\left((2\pi)^{n(k-1)} H_{k,T}(0)\right)^{-1} \prod_{j=1}^{k} H_{1,T}(\lambda_j) = \Phi_{k,T}^{h}(\lambda_1, ..., \lambda_{k-1}) \tag{3.3}$$

is a multidimensional kernel over $\mathbb{R}^{n(k-1)}$, which is an approximate identity for convolution (see the next Remark), and the following equality holds:

$$\lim_{T\to\infty} \int_{\mathbb{R}^{n(k-1)}} \Phi_{k,T}^{h}(u_1, ..., u_{k-1}) G(u_1, ..., u_{k-1}) du_1 ... du_{k-1} = G(0, ..., 0), \tag{3.4}$$

provided that the function $G(u_1, ..., u_{k-1})$ is bounded and continuous at the point $(u_1, ..., u_{k-1}) = (0, ..., 0)$.

Note that, for the discrete case, the integrals over $\mathbb{R}^{n(k-1)}$ are to be substituted by integrals over $(-\pi, \pi]^{n(k-1)}$ and the integrals over $[0,T]^n$ are to be substituted by multidimensional sums over the index set $D_T = \{(t_1, ..., t_n), t_i = 1, 2, ..., T, i = 1, ..., n\}$. The result will be a family of multidimensional kernels over $(-\pi, \pi]^{n(k-1)}$.

Remark 10. *Recall, for example, from Rudin (1991) that a family $\{\Psi_T(\lambda) = \Psi_T(\lambda_1, ..., \lambda_{k-1}), T \in N\}$ of Lebesgue integrable functions on $\Pi^{k-1} = (-\pi, \pi]^{k-1}$, with values in $\mathbb{C}$, is called an approximate identity for convolution if*
(i) $\sup_T \int_{\Pi^{k-1}} |\Psi_T(\lambda)| d\lambda < \infty$;
(ii) $\lim_{T\to\infty} \int_{\Pi^{k-1}} \Psi_T(\lambda) d\lambda = 1$;
(iii) $\lim_{T\to\infty} \int_{\Pi^{k-1} \setminus \{\lambda, \|\lambda\| < \delta\}} |\Psi_T(\lambda)| d\lambda = 0 \ \forall \delta > 0$, *where* $\|\lambda\| = \sup_{i=1,...k-1} |\lambda_i|$.
Then, for all bounded, continuous, complex-valued functions f on Π^{k-1}, $\lim_{T\to\infty} \Psi_T * f = f$. The kernels of the form

$$\Psi_T^h(\lambda_1, ..., \lambda_{k-1}) = \left((2\pi)^{(k-1)} H_{k,T}(0)\right)^{-1} \prod_{j=1}^{k-1} H_{1,T}(\lambda_j) H_{1,T}\left(-\sum_{j=1}^{k-1} \lambda_j\right), \ (\lambda_1, ..., \lambda_{k-1}) \in \Pi^{k-1},$$

which are, in essence, tapered multidimensional generalizations of the Féjer kernel, were considered in Dahlhaus (1983) and it was shown that the above conditions (i)-(iii) hold for these kernels, with h being continuously differentiable (see Lemma 3 of Dahlhaus 1983).

Analogously, this assertion can be stated also in the continuous-time setting for the kernels $\Phi_T^h(\lambda_1, ..., \lambda_{k-1})$, $(\lambda_1, ..., \lambda_{k-1}) \in \mathbb{R}^{k-1}$. Furthermore, as we have chosen a taper which factorizes, these assertions extend directly to the kernels given by (3.3) and their discrete version.

The kernels which appear (coming from (3.1)) in the untapered case, namely,

$$\begin{aligned} \Phi_{k,T}(u_1, ..., u_{k-1}) &= \left((2\pi)^{(k-1)} T\right)^{-1} \int_{[0,T]^k} \exp\{i \sum_{j=1}^{k} t_j u_j\} dt_1 \dots dt_k \\ &= \left((2\pi)^{(k-1)} T\right)^{-1} \prod_{j=1}^{k} \frac{\sin(T u_j/2)}{u_j/2}, \quad u_k = \sum_{j=1}^{k-1} u_j, \end{aligned}$$

$$\Psi_{k,T}(u_1,...,u_{k-1}) = \left((2\pi)^{(k-1)} T\right)^{-1} \sum_{t_1,...,t_k=1,...,T} \exp\{i\sum_{j=1}^{k} t_j u_j\}$$
$$= \left((2\pi)^{(k-1)} T\right)^{-1} \prod_{j=1}^{k} \frac{\sin(Tu_j/2)}{\sin(u_j/2)}, \quad u_k = \sum_{j=1}^{k-1} u_j,$$

were treated in Bentkus (1972a, 1972b) and were shown to satisfy the above conditions (i)-(iii)

For the case when $\sum_{j=1}^{k} \alpha_j = 0$, formula (3.1) leads to the following result:

(3.5)
$$\left((2\pi)^{n(k-1)} H_{k,T}(0)\right)^{-1} cum\left(d_T^h(\alpha_1),...,d_T^h(\alpha_k)\right)$$
$$= \int_{\mathbb{R}^{n(k-1)}} \Phi_{k,T}^h\left(\gamma_1-\alpha_1,...,\gamma_{k-1}-\alpha_{k-1}\right) f_k\left(\gamma_1,...,\gamma_{k-1}\right) d\gamma_1...d\gamma_{k-1}$$
$$= \int_{\mathbb{R}^{n(k-1)}} \Phi_{k,T}^h\left(u_1,...,u_{k-1}\right) f_k\left(u_1+\alpha_1,...,u_{k-1}+\alpha_{k-1}\right) du_1...du_{k-1}.$$

The proofs of Theorems 1 and 2 make use of some general results concerning large-sample properties of the empirical spectral functionals of k-th order,

$$J_{k,T}^h(\varphi) = J_{k,T}^h(\varphi; w_0) = \int_{\mathbb{R}^{n(k-1)}} I_{k,T}^h(\lambda)\varphi(\lambda) w_0(\lambda) d\lambda',$$

where $I_{k,T}^h(\lambda)$ is the periodogram based on tapered data, the function $w_0(\lambda)$ satisfies the condition II(ii), and $\varphi(\lambda)$ is a weight function.

The cumulants of the functional $J_{k,T}^h(\varphi)$ can be represented in the form of singular integrals (with the use of the formulae (3.1)-(3.3) and (3.5)), and, due to the property (3.4) of the kernels $\Phi_{k,T}^h(\lambda_1,...,\lambda_{k-1})$, the asymptotic behaviour of the functionals $J_{k,T}^h(\varphi)$ can be evaluated.

We summarize these results in the following two lemmas.

Lemma 1. *Let the random field $Y(t)$, $t \in \mathbb{R}^n$, satisfy assumption I.*
1)

$$EJ_{k,T}^h(\varphi) = \int_{\mathbb{R}^{n(k-1)}} \Phi_{k,T}^h(u) G_k(u;\varphi,w_0) du',$$

where $G_k(u;\varphi,w_0)$ is given by (2.14). Furthermore, if $G_k(u)$ is bounded and continuous at $u=0$, then

$$EJ_{k,T}^h(\varphi) \to \int_{\mathbb{R}^{n(k-1)}} f_k(\lambda)\varphi(\lambda) w_0(\lambda) d\lambda' \text{ as } T\to\infty.$$

2)

$$cov\left(J_{k,T}^h(\varphi_1), J_{k,T}^h(\varphi_2)\right) = (2\pi)^n H_{2k,T}(0)\left(H_{k,T}(0)\right)^{-2} \times \int_{\mathbb{R}^{n(2k-1)}} \Phi_{2k,T}^h(u) G_{2k}(u;\varphi_1,\varphi_2,w_0) du',$$

where the function $G_{2k}(u;\varphi_1,\varphi_2,w_0)$ is given by (2.15) with $m=2$; if the function $G_{2k}(u)$ is bounded and continuous at $u=0$, then as $T\to\infty$,

(3.6)
$$cov\left(T^{n/2}J_{k,T}^h(\varphi_1), T^{n/2}J_{k,T}^h(\varphi_2)\right) = (2\pi)^n e(h) G_{2k}(0;\varphi_1,\varphi_2,w_0),$$

where

$$e(h)=\left(\int (h_1(t))^{2k}dt\left(\int (h_1(t))^{k}dt\right)^{-2}\right)^{n}. \tag{3.7}$$

3)

$$\begin{aligned} cum\left(J_{k,T}^{h}\left(\varphi_1\right),...,J_{k,T}^{h}\left(\varphi_m\right)\right) &= (2\pi)^{n(m-1)}H_{km,T}\left(0\right)\left(H_{k,T}\left(0\right)\right)^{-m} \\ &\quad\times\int_{\mathbb{R}^{n(km-1)}}\Phi_{km,T}^{h}\left(u\right)G_{km}\left(u;\varphi_1,...,\varphi_m,w_0\right)du', \end{aligned}$$

where the function $G_{km}\left(u;\varphi_1,...,\varphi_m,w_0\right)$ *is given by the formula (2.15); if* $G_{km}\left(u\right)$ *is bounded, then as* $T\to\infty$

$$cum\left(J_{k,T}^{h}\left(\varphi_1\right),...,J_{k,T}^{h}\left(\varphi_m\right)\right)=O\left(\frac{1}{T^{n(m-1)}}\right).$$

Remark 11. *As a consequence of Lemma 1, we have that, under assumption I, and using results 1) and 2) (with* $\varphi_1=\varphi_2=\varphi$*) of Lemma 1,*

$$J_{k,T}^{h}\left(\varphi\right)\to J_k\left(\varphi\right)=\int_{\mathbb{R}^{n(k-1)}}f_k\left(\lambda\right)\varphi_i\left(\lambda\right)w_0\left(\lambda\right)d\lambda'$$

in probability.

Let us now &x the weight functions $\varphi_1,...,\varphi_m$ and set

$$\begin{aligned} J_{k,T}^{h} &= \left\{J_{k,T}^{h}\left(\varphi_i\right)\right\}_{i=1,...,m} \\ &= \left\{\int_{\mathbb{R}^{n(k-1)}}I_{k,T}^{h}\left(\lambda\right)\varphi_i\left(\lambda\right)w_0\left(\lambda\right)d\lambda'\right\}_{i=1,...,m} \end{aligned}$$

and

$$\begin{aligned} J_k &= \left\{J_k\left(\varphi_i\right)\right\}_{i=1,...,m} \\ &= \left\{\int_{\mathbb{R}^{n(k-1)}}f_k\left(\lambda\right)\varphi_i\left(\lambda\right)w_0\left(\lambda\right)d\lambda'\right\}_{i=1,...,m}. \end{aligned}$$

Let $\xi=\{\xi_i\}_{i=1,...,m}$ be a complex-valued Gaussian random vector with mean zero and second-order moments

$$w_{ij}=E\xi_i\bar{\xi}_j=(2\pi)^n e\left(h\right)G_{2k}\left(0;\varphi_i,\varphi_j,w_0\right),\ i,j=1,...,m, \tag{3.8}$$

where the tapering factor is given by (3.7).

Lemma 2. *Let the assumptions of Lemma 1 hold and the functions* $G_{kl}\left(u;\varphi_{m_1},...,\varphi_{m_l};w_0\right)$*, de&ned by (2.15), be bounded for all* $l=2,3,...$ *and all choices* $(m_1,...,m_l)$ *with* $1\le m_i\le m$*,* $i=1,...,l$*. Then, as* $T\to\infty$*,*

$$T^{n/2}\left(J_{k,T}^{h}-EJ_{k,T}^{h}\right)\xrightarrow{D}\xi, \tag{3.9}$$

and, moreover, if

$$T^{n/2}\left(EJ_{k,T}^{h}\left(\varphi_i\right)-J_k\left(\varphi_i\right)\right)\to 0\quad as\ \ T\to\infty,\quad i=1,...,m, \tag{3.10}$$

then, as $T\to\infty$*,*

$$T^{n/2}\left(J_{k,T}^{h}-J_k\right)\xrightarrow{D}\xi. \tag{3.11}$$

The proofs of Lemmas 1 and 2 parallel those of the corresponding results for the one-dimensional case in Anh, Leonenko and Sakhno (2002b). We need only keep track of data tapers which contribute, in particular, an additional tapering factor in the expression for the covariance function (3.6) and, therefore, to the covariance matrix (3.8).

Remark 12. *For the discrete-time case and $k=2$, sufficient conditions for (3.10) presented in Guyon (1995) are the following: $n=1,2,3$ and $\varphi_i \in C(\Pi^n)$, and the taper h and the spectral density f_2 belong to $C^2(\Pi^n)$. Following the argument for bias evaluation of the empirical spectral functional presented in Guyon (1995), but treating the spectral density and weight function together, less restrictive conditions for (3.10) can be obtained. We leave this problem for further investigation.*

Proof of Theorem 1.

In view of Lemma 1 and Remark 11, we can conclude that, under the assumptions of the theorem, the following convergence holds as $T\to\infty$:

$$U_T^h(\theta)\to U(\theta) \quad \text{in } P_0\text{-probability,} \tag{3.12}$$

which implies

$$U_T^h(\theta)-U_T^h(\theta_0)\to U(\theta)-U(\theta_0)=K(\theta_0;\theta)$$

in P_0-probability. By Jensen's inequality and the relations (2.4) to (2.7),

$$\begin{aligned}
-K(\theta_0;\theta) &= p\int_{\mathbb{R}^{(k-1)n}} f_k^{(1)}(\lambda;\theta_0)\log\frac{\psi_k^{(1)}(\lambda;\theta)}{\psi_k^{(1)}(\lambda;\theta_0)}w^{(1)}(\lambda)\,w_0(\lambda)\,d\lambda' \\
&\quad +q\int_{\mathbb{R}^{(k-1)n}} f_k^{(2)}(\lambda;\theta_0)\log\frac{\psi_k^{(2)}(\lambda;\theta)}{\psi_k^{(2)}(\lambda;\theta_0)}w^{(2)}(\lambda)\,w_0(\lambda)\,d\lambda' \\
&= p\sigma_k^{(1)}(\theta_0)\int_{\mathbb{R}^{(k-1)n}} \psi_k^{(1)}(\lambda;\theta_0)\log\frac{\psi_k^{(1)}(\lambda;\theta)}{\psi_k^{(1)}(\lambda;\theta_0)}w_0(\lambda)\,d\lambda' \\
&\quad +q\sigma_k^{(2)}(\theta_0)\int_{\mathbb{R}^{(k-1)n}} \psi_k^{(2)}(\lambda;\theta_0)\log\frac{\psi_k^{(2)}(\lambda;\theta)}{\psi_k^{(2)}(\lambda;\theta_0)}w_0(\lambda)\,d\lambda' \\
&\le p\sigma_k^{(1)}(\theta_0)\log\int_{\mathbb{R}^{(k-1)n}} \psi_k^{(1)}(\lambda;\theta_0)\,w_0(\lambda)\,d\lambda' \\
&\quad +q\sigma_k^{(2)}(\theta_0)\log\int_{\mathbb{R}^{(k-1)n}} \psi_k^{(2)}(\lambda;\theta_0)\,w_0(\lambda)\,d\lambda'=0,
\end{aligned}$$

that is, $K(\theta_0;\theta)\ge 0$. The condition that $f_k(\lambda;\theta_1)\not\equiv f_k(\lambda;\theta_2)$ for $\theta_1\ne\theta_2$ almost everywhere in $\mathbb{R}^{n(k-1)}$ with respect to the Lebesgue measure then assures that $K(\theta_0;\theta)>0$ if $\theta\ne\theta_0$.

For the consistency of the estimator $\widehat{\theta}_T$ it remains to show that the convergence in (3.12) holds uniformly in $\theta\in\Theta$. Under condition VI(i), let us denote $\eta(\varepsilon)=\min\{\eta_1(\varepsilon),\eta_2(\varepsilon)\}$, with $\eta_i(\varepsilon)$ being the modulus of continuity of the function $a_k^{(i)}(\lambda;\theta)$, $i=1,2$. Then,

$$\begin{aligned}
&\sup\left\{\left|U_T^h(\theta_1)-U_T^h(\theta_2)\right|,\ \theta_1,\theta_2\in\Theta,|\theta_1-\theta_2|\le\eta(\varepsilon)\right\}\\
\le\ &\varepsilon\left[p\int_{\mathbb{R}^{(k-1)n}}\operatorname{Re}I_{k,T}^h(\lambda)\frac{w^{(1)}(\lambda)}{v_1(\lambda)}w_0(\lambda)\,d\lambda'\right.\\
&\left.+q\int_{\mathbb{R}^{(k-1)n}}\operatorname{Im}I_{k,T}^h(\lambda)\frac{w^{(2)}(\lambda)}{v_2(\lambda)}w_0(\lambda)\,d\lambda'\right].
\end{aligned}\tag{3.13}$$

Further, in view of VI(ii), Lemma 1 and Remark 11, we obtain that

$$\int_{\mathbb{R}^{(k-1)n}}I_{k,T}^h(\lambda)\frac{w^{(i)}(\lambda)}{v_i(\lambda)}w_0(\lambda)\,d\lambda'\to\int_{\mathbb{R}^{(k-1)n}}f_k(\lambda;\theta_0)\frac{w^{(i)}(\lambda)}{v_i(\lambda)}w_0(\lambda)\,d\lambda'\tag{3.14}$$

in probability. From (3.14), we can conclude that

$$\int_{\mathbb{R}^{(k-1)n}}I_{k,T}^h(\lambda)\frac{w^{(i)}(\lambda)}{v_i(\lambda)}w_0(\lambda)\,d\lambda'=O_p(1),$$

which implies that the expression in the square brackets of (3.13) is $O_p(1)$. This completes the proof of Theorem 1.

Proof of Theorem 2.

From Taylor's formula we have the relation

$$\nabla_\theta U_T^h\left(\widehat{\theta}_T\right)=\nabla_\theta U_T^h(\theta_0)+\nabla_\theta\nabla_\theta' U_T^h(\theta_T^*)\left(\widehat{\theta}_T-\theta_0\right),\tag{3.15}$$

where $|\theta_T^*-\theta_0|<\left|\widehat{\theta}_T-\theta_0\right|$,

$$\begin{aligned}
\nabla_\theta U_T^h(\theta)\ =\ &-(p\int_{\mathbb{R}^{(k-1)n}}\operatorname{Re}I_{k,T}^h(\lambda)w^{(1)}(\lambda)w_0(\lambda)\nabla_\theta\log\psi_k^{(1)}(\lambda;\theta)\,d\lambda'\\
&+q\int_{\mathbb{R}^{(k-1)n}}\operatorname{Im}I_{k,T}^h(\lambda)w^{(2)}(\lambda)w_0(\lambda)\nabla_\theta\log\psi_k^{(2)}(\lambda;\theta)\,d\lambda');
\end{aligned}\tag{3.16}$$

$$\begin{aligned}
\nabla_\theta\nabla_\theta' U_T^h(\theta)\ =\ &-(p\int_{\mathbb{R}^{(k-1)n}}\operatorname{Re}I_{k,T}^h(\lambda)w^{(1)}(\lambda)w_0(\lambda)\nabla_\theta\nabla_\theta'\log\psi_k^{(1)}(\lambda;\theta)\,d\lambda'\\
&+q\int_{\mathbb{R}^{(k-1)n}}\operatorname{Im}I_{k,T}^h v(\lambda)w^{(2)}(\lambda)w_0(\lambda)\nabla_\theta\nabla_\theta'\log\psi_k^{(2)}(\lambda;\theta)\,d\lambda').
\end{aligned}\tag{3.17}$$

It follows from the definition of the minimum contrast estimator that, for sufficiently large T,

$$\nabla_\theta U_T^h(\theta_0)=-\nabla_\theta\nabla_\theta' U_T^h(\theta_T^*)\left(\widehat{\theta}_T-\theta_0\right).\tag{3.18}$$

In view of assumptions VIII(i) and (ii), and by Lemma 1, we have

$$\begin{aligned}
&\int_{\mathbb{R}^{(k-1)n}}I_{k,T}^h(\lambda)w^{(l)}(\lambda)w_0(\lambda)\nabla_\theta\nabla_\theta'\log\psi_k^{(l)}(\lambda;\theta)\,d\lambda'\\
&\xrightarrow{P_0}\int_{\mathbb{R}^{(k-1)n}}f_k(\lambda;\theta_0)w^{(l)}(\lambda)w_0(\lambda)\nabla_\theta\nabla_\theta'\log\psi_k^{(l)}(\lambda;\theta)\,d\lambda',\ l=1,2,
\end{aligned}$$

which implies

$$\nabla_\theta\nabla_\theta' U_T^h(\theta_T^*)\to S_k(\theta_0)\tag{3.19}$$

in P_0-probability, where the matrix $S_k(\theta_0)$ is given by (2.21).

On the other hand, we have

$$\begin{aligned} T^{n/2}\nabla_\theta U_T^h(\theta_0) &= T^{n/2}\left(p\,\mathrm{Re}\,J_{k,T}^h\left(g_k^{(i)}\right)+q\,\mathrm{Im}\,J_{k,T}^h\left(g_k^{(m+i)}\right)\right)_{i=1,\dots,m} \\ &= T^{n/2}(p(\mathrm{Re}\,J_{k,T}^h(g_k^{(i)})-\mathrm{Re}\,J_k(g_k^{(i)})) \\ &\quad +q(\mathrm{Im}\,J_{k,T}^h(g_k^{(m+i)})-\mathrm{Im}\,J_k(g_k^{(m+i)})))_{i=1,\dots,m}, \end{aligned}$$

where the last equality is due to (2.8). Under the conditions of the theorem, and using Lemma 2, we can conclude that the following convergence holds:

$$T^{n/2}\nabla_\theta U_T^h(\theta_0) \xrightarrow{D} N_m(0, A_k(\theta_0)) \text{ as } T\to\infty. \tag{3.20}$$

Now by Slutsky's arguments, the convergence (2.23) is a consequence of (3.18), (3.19) and (3.20). This completes the proof of the theorem.

References

[1] V. V. Anh, C. C. Heyde and N. N. Leonenko, Dynamic models of long-memory processes driven by Lévy noise, *J. Appl. Prob.* 39 (2002) 730-747.

[2] V.V. Anh, N.N. Leonenko and L.M. Sakhno, Minimum contrast estimation of random processes and fields based on information of the second and third order (2002a) (submitted).

[3] V.V. Anh, N.N. Leonenko and L.M. Sakhno, Statistical inference using higher-order information (2002b) (submitted).

[4] V.V. Anh, N.N. Leonenko and L.M. Sakhno, On a class of minimum contrast estimators for fractional stochastic processes and fields, *J. Statist. Plann. Infer.* (2003), in press.

[5] R. Bentkus, On the error of the estimate of the spectral function of a stationary process, *Liet. Mat. Rink.* 12 (1972a) 55-71 (in Russian).

[6] R. Bentkus, Asymptotic normality of an estimate of the spectral function, *Liet. Mat. Rink.* 12 (1972b) 5–18 (in Russian).

[7] R. Bentkus and R. R. Maliukevicius, Statistical estimation of a multidimentional parameter of a spectral density I, *Lithuanian Math. J.*, 28 (1988a) 115-126.

[8] R. Bentkus and R. R. Maliukevicius, Statistical estimation of a multidimentional parameter of a spectral density II, *Lithuanian Math. J.*, 28 (1988b), 415-431.

[9] R. Bentkus and R. Rutkauskas, On the asymptotics of the first two moments of second order spectral estimators, *Liet. Mat. Rink.* 13 (1973) 29–45 (in Russian).

[10] D.R. Brillinger, Introduction to polyspectra, *Ann. Math. Statist.*, 36 (1965) 1351-1374.

[11] D.R. Brillinger and M. Rosenblatt, Asymptotic theory of k-th order spectra. In B.Harris (Ed.), *Spectral Analysis of Time Series*, Wiley, New York, 1967a, 153-188.

[12] D.R. Brillinger and M. Rosenblatt, Computation and interpretation of k-th order spectra. In B.Harris (Ed.), *Spectral Analysis of Time Series*, Wiley, New York, 1967b, 189-232.

[13] R. Dahlhaus, Spectral analysis with tapered data, *J. Time. Ser. Anal.*, 4 (1983) 163-175.

[14] R. Dahlhaus and H. Künsch, Edge effects and efficient parameter estimation for stationary random fields, *Biometrika*, 74 (1987) 877-882.

[15] W.Dusmuir and E.J.Hannan, Vector linear time series models, *Adv. Appl. Prob.* 8 (1976) 339-360.

[16] R. Fox and M. S. Taqqu, Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series, *Ann. Statist.* 14 (1986) 517–532.

[17] J. Gao, V.V. Anh, C.C. Heyde and Q. Tieng. Parameter estimation of stochastic processes with long-range dependence and intermittency, *J. Time Series Anal.* 22 (2001) 517-535.

[18] J. Gao, V.V. Anh and C.C. Heyde, Statistical estimation of nonstationary Gaussian processes with long-range dependence and intermittency, *Stoch. Proc. Appl.* 99 (2002) 295-321.

[19] L. Giraitis and D. Surgailis, A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotic normality of Whittle estimate, *Prob. Theor. Rel. Fields*, 86 (1990) 87-104.

[20] L. Giraitis and M.Taqqu, Whittle estimator for finite-variance non-Gaussian time series with long memory, *Ann. Statist.* 27 (1999) 178-203.

[21] X. Guyon, Parameter estimation for a stationary process on a d-dimensional lattice, *Biometrica*, 69 (1982) 95-105.

[22] X. Guyon, *Random Fields on a Network: Modelling, Statistics and Applications,* Springer-Verlag, New York, 1995.

[23] E.J. Hannan, *Multiple Time Series,* Springer-Verlag, New York, 1970.

[24] E.J. Hannan, The asymptotic theory of linear time series models, *J. Appl. Prob.* 10 (1973) 130-145.

[25] C. C. Heyde, *Quasi-Likelihood And Its Application: A General Approach to Optimal Parameter Estimation*, Springer-Verlag, New York, 1997.

[26] C. Heyde, A risky asset model with strong dependence through fractal activity time, *J. Appl. Prob.* 36 (1999) 1234-1239.

[27] C. Heyde and R. Gay, Smoothed periodogram asymptotics and estimation for processes and fields with possible long-range dependence, *Stoch. Proc. Appl.*, 45 (1993) 169–182.

[28] I.A. Ibragimov, On estimation of the spectral function of a stationary Gaussian process, *Theor. Prob. Appl.* 8 (1963) 366–401.

[29] I.A. Ibragimov, On maximum likelihood estimation of parameters of the spectral density of stationary time series, *Theor. Prob. Appl.*, 12 (1967) 115–119.

[30] A.V. Ivanov and N.N. Leonenko, *Statistical Analysis of Random Fields*, Kluwer, 1989.

[31] N. N. Leonenko and E. M. Moldavs'ka, Minimum contrast estimators of a parameter of the spectral density of continuous time random fields, *Theor. Prob. Math. Statist.* 58 (1999) 101–112.

[32] M.R. Rosenblatt, *Stationary Sequences and Random Fields*, Birkhäuser, Boston, 1985.

[33] W. Rudin *Functional Analysis,* McGraw-Hill, 1991.

[34] A. Swami, G.B. Giannakis and G. Zhou, Bibliography on higher-order statistics, *Signal Processing*, 60 (1997) 65-126.

[35] J.W. Tukey, An introduction to the calculations of numerical spectrum analysis. In B.Harris (Ed.), *Spectral Analysis of Time Series*, Wiley, New York, 1967.

(V. V. Anh) Program in Statistics and Operations Research, Queensland University of Technology, GPO Box 2434, Brisbane QLD 4001, Australia

E-mail address, V. V. Anh: `v.anh@qut.edu.au`

(N. N. Leonenko) School of Mathematics, Cardiff University, Senghennydd Road, Cardiff CF24 4YH, UK

E-mail address, N. N. Leonenko: `LeonenkoN@Cardiff.ac.uk`

(L. M. Sakhno) Department of Mathematics, Kyiv National Taras Shevchenko University, Volodymyrska 64, 01033, Kyiv, Ukraine

E-mail address, L. M. Sakhno: `lms@univ.kiev.ua`