A FREQUENCY DOMAIN BOOTSTRAP FOR RATIO STATISTICS IN TIME SERIES ANALYSIS

by

R. DAHLHAUS and D. JANAS

Universität Heidelberg

The asymptotic properties of the bootstrap in the frequency domain based on studentized periodogram ordinates are studied. It is proved that this bootstrap approximation is valid for ratio statistics such as autocorrelations. By using Edgeworth expansions it is shown that the bootstrap approximation even outperforms the normal approximation. The results carry over to Whittle estimates. In a simulation study the behaviour of the bootstrap is studied for empirical correlations and Whittle estimates.

1. Introduction. The bootstrap (Efron, 1979) is generally accepted as a powerful tool for approximating certain characteristics, e.g. bias, variance or the distribution of statistics that cannot at all or only with excessive effort be calculated by analytical means. For example, the bootstrap provides second order corrected approximations to sampling distributions in the i.i.d. set up (Singh, 1981; Babu and Singh, 1984). In time series analysis, where the data obey a certain dependence structure, this kind of difficulty quite often comes up, particularly, if one is not willing to assume Gaussianity of the data. In principle, one has with a time series only one observation of a multivariate random variable and it is obvious that a bootstrap can only be applied to parts of the data or to certain transformations (e.g. residual). Very often this requires additional assumptions on the dependence structure (e.g. mixing-assumptions) or on the underlying model (e.g. for a residual-based bootstrap).

Künsch (1989) and Liu and Singh (1992) propose to resample whole blocks of consecutive observations. Instead of resampling from the data themselves another idea is to resample from residuals that are approximately i.i.d.. Freedman (1984), Efron and Tibshirani (1986), Swanepoel and van Wyk (1986) and Kreiss and Franke (1992) consider resampling the estimated

---

AMS 1990 *subject classifications*: Primary 62 M 10; secondary 62 E 20.

Key words and phrases: Autocorrelations, bootstrap, periodogram ordinates, ratio statistics, spectral mean, time series, Whittle estimators.

innovations of parametric time series models.

A different approach is to apply Efron's bootstrap method to periodogram ordinates, more precisely to studentized periodogram ordinates, where the periodogram is studentized by a spectral density estimate (cf. Franke and Härdle, 1992; Hurvich and Zeger, 1987; Nordgaard, 1992). For obvious reasons this method may be denoted as a frequency domain bootstrap, whereas all procedures above resample in the time domain. Franke and Härdle apply this procedure to kernel spectral density estimates and show a consistency result. They motivate the approach by interpreting the spectral estimation problem as an approximate regression problem. Unfortunately, the periodogram ordinates are only approximately independent. This causes trouble for other estimates such as estimates of the autocovariance function. The dependence between different periodogram ordinates leads for non Gaussian processes to an additional contribution to the asymptotic variance of this estimate. Since the bootstrap replicates are independent the additional part of the variance cannot be imitated. Therefore the method fails in such cases.

In this paper we study the class of estimators for which this bootstrap with periodogram ordinates works, more detailed. As mentioned above it works for all spectral mean estimates if the data are assumed to be Gaussian. The procedure keeps working without this assumption for the kernel spectral density estimate (Franke and Härdle, 1992) since those estimators have a rate of convergence less than $T^{-1/2}$. However, in other cases the validity of this bootstrap is not obvious. The main result of this paper is that there exists an important class of statistics for which the bootstrap works: *ratio statistics*. These statistics may be represented as ratios of spectral mean estimates and the integrated periodogram. For example, the usual moment estimator for the autocorrelation is a ratio statistic. This estimate is a normalized version of the autocovariance estimate for which the procedure fails. An inspection of the cumulants reveals that the method does not only approximate the mean and the variance of ratio statistics, but also leads to the correct skewness. Besides, by means of Edgeworth expansions for the statistics of interest and their bootstrapped versions we find that the error of the bootstrap approximation is of order less than $T^{-1/2}$ and therefore outperforms the normal approximation.

A different approach was considered in Janas and Dahlhaus (1994). There we have suggested a modification of the frequency bootstrap which imitates the weak dependence structure of the periodogram and leads to a consistent bootstrap approximation for general spectral mean estimates in the non-Gaussian case. However, one can check that this procedure does not lead to a correct estimate of the skewness.

The paper is organized as follows: In section 2 we discuss the problem. The bootstrap procedure based on the sample of the studentized periodogram ordinates is presented and ratio statistics are introduced . At the end of the section we summarize the assumptions and notations needed throughout the paper. The main results are presented in section 3 and applied to Whittle estimates in section 4 . In section 5 some simulation examples illustrate the performance of the method. To make the paper more convenient for the reader some of the proofs are transferred to an appendix.

2. Preliminaries. Consider a real-valued stationary time series $\{X_t\}_{t\in \mathbf{Z}}$ with $\mathbf{E}X_t = 0$ and spectral density (≡ sd) f. Let us denote by

(2.1) $$A(\phi,f) \equiv \Big(\int_0^\pi \phi^{(1)}(\alpha)\, f(\alpha)\, d\alpha, \dots, \int_0^\pi \phi^{(d)}(\alpha)\, f(\alpha)\, d\alpha\Big)' \; (\equiv \textstyle\int\phi\, f)$$

the *spectral mean*, where $\phi^{(r)}$ are functions of bounded variation, r = 1, … , d. The canonical estimate of $A(\phi,f)$ is

(2.2) $$A(\phi,I_T) \equiv \Big(\int_0^\pi \phi^{(1)}(\alpha)\, I_T(\alpha)\, d\alpha, \dots, \int_0^\pi \phi^{(d)}(\alpha)\, I_T(\alpha)\, d\alpha\Big)' \; (\equiv \textstyle\int\phi\, I_T),$$

where $I_T(\alpha)$ is the *tapered periodogram*, i.e.

(2.3) $$I_T(\alpha) \equiv (2\pi H_{2,T})^{-1}\, d_T(\alpha)\, d_T(-\alpha) ,$$

where

(2.4) $$d_T(\alpha) \equiv \sum_{t=1}^{T} h_t\, X_t \exp(-i\,\alpha\, t)$$

denotes the *tapered finite Fourier transform*,

(2.5) $$H_{k,T}(\alpha) \equiv \sum_{t=1}^{T} h_t^k \exp(-i\alpha t)$$

is the *spectral window, and* $H_{k,T} = H_{k,T}(0)$.

The following special cases are covered by this class.

<u>EXAMPLE 1 *(autocovariance estimate)*:</u> Let $\phi(\alpha) = 2\cos(\alpha u)$, $u \in \mathbf{Z}$. Then

$$A(\phi, I_T) = \int_{-\pi}^{\pi} I_T(\alpha) \exp(-i\alpha u)\, d\alpha$$

$$= (H_{2,T})^{-1} \sum_{t=1}^{T} h_t X_t h_{t+u} X_{t+u} \equiv c_T(u)$$

(with $h_t = 0$ for $t \le 0$ and $t > T$) is the usual moment estimator with tapered data for the autocovariance $c(u) \equiv \mathbf{E}(X_t X_{t+u})$.

<u>EXAMPLE 2 *(spectral distribution function (≡ sdf) estimate)*:</u> With $\phi(\alpha) = \chi_{[0,\lambda]}(\alpha)$, $\lambda \in [0,\pi]$, we get the integrated periodogram

$$A(\phi, I_T) = \int_0^{\lambda} I_T(\alpha)\, d\alpha \equiv F_T(\lambda) ,$$

which is an estimate for the spectral distribution function $F(\lambda) = \int_0^{\lambda} f(\alpha) d\alpha$ (sdf).

<u>EXAMPLE 3 *(Whittle estimate)*:</u> Let $F = \{f_\theta : \theta \in \Theta\}$ be a parametric family of spectral densities. Then, the parameter θ may be estimated by minimizing the Whittle likelihood

$$L_T(\theta) = (2\pi)^{-1} \int_0^{\pi} \{\log f_\theta(\alpha) + f_\theta^{-1}(\alpha) I_T(\alpha)\}\, d\alpha .$$

With $\phi(\alpha) \equiv \nabla f_\theta^{-1}$ we have

$$\nabla L_T(\theta) = 0 \Leftrightarrow A(\phi, I_T) - A(\phi, f) = 0$$

where f is the true spectral densitiy.

The basic idea of a bootstrap for $A(\phi,I_T)$ relies on the fact that $I_T(\alpha) / f(\alpha)$ are for a fixed set of frequencies $\{\alpha_1, \dots, \alpha_K\}$ with $\alpha_j \neq 0 \bmod \pi$ asymptotically independent exponential variables (cf. Brillinger, 1981, Theorem 5.2.6). This suggests the following bootstrap procedure. Let $n = [T/2]$ and $I_j = I_T(\frac{2\pi j}{T})$.

Bootstrap procedure

(1) Obtain the sample of periodogram ordinates $\{I_j\}$ for $j = 1, \dots, n$.

(2) Obtain an estimate $\hat{f}$ of the spectral density f (e.g. a kernel estimate) .

(3) Form the studentized periodogram ordinates $\{\hat{\varepsilon}_j\} \equiv \{I_j / \hat{f}_j\}$.

(4) Rescale $\hat{\varepsilon}_j$ and consider $\{\tilde{\varepsilon}_j\} \equiv \{\hat{\varepsilon}_j / \hat{\varepsilon}_\bullet\}$, where $\hat{\varepsilon}_\bullet = \frac{1}{n} \sum_{j=1}^{n} \hat{\varepsilon}_j$.

(5) Draw independent bootstrap replicates $\{\varepsilon_j^*\}$ from the empirical distribution of the $\tilde{\varepsilon}_j$.

(6) Define bootstrap periodogram values by $\{I_j^*\} \equiv \{\hat{f}_j \, \varepsilon_j^*\}$.

<u>REMARKS.</u> (1) The rescaling in step (4) avoids an unneccessary bias at the resampling stage.
(2) Exploiting our knowledge about the asymptotic distribution of $I_T(\alpha) / f(\alpha)$, we may modify the procedure by replacing $\{\varepsilon_j^*\}$ by independent and standard exponentially distributed variables $\{E_j^*\}$. As in step (6) we get modified bootstrap periodogram values $\{I_j^+\} \equiv \{\hat{f}_j E_j^*\}$. We see in the next section that all results hold for both the original procedure as well as the modified one.

We now try to approximate the distribution of $A(\phi,I_T) - A(\phi,f)$ by the distribution of $B(\phi,I_T^*) - B(\phi,\hat{f})$ where

$$B(\phi,I_T^*) \equiv \frac{\pi}{n} \sum_{j=1}^{n} \phi_j \, I_j^* \tag{2.6}$$

and $\phi_j \equiv \phi(\frac{2\pi j}{T})$.

To get an idea on the quality of this bootstrap approximation we study the asymptotic behaviour of both statistics.

It is well known (cf. Dahlhaus, 1983, 1985a) that $\sqrt{T}(A(\phi,I_T) - A(\phi,f))$ is asymptotically normal. For a linear process $\{X_t\}$ and $h_t \equiv 1$ the asymptotic variance is given by

(2.7) $$2\pi\int \phi^2 f^2 + (\kappa_4/\sigma^4)\,(\int\phi f)^2,$$

where κ_4 is the fourth cumulant and σ^2 is the variance of the innovations ε_t .

Under appropiate assumptions $\sqrt{T}(B(\phi,I_T^*) - B(\phi,\hat{f}))$ is also asymptotically normal, but with a variance proportional to

(2.8) $$2\pi\int\phi^2 f^2 \ .$$

The difference in the two asymptotic distributions relies on the fact that the I_T^* are independent, while the dependence structure of the I_T cannot be neglected completely.

Therefore, this bootstrap can only work, if the additional term in (2.7) vanishes. There are two cases in which this term is zero.

<u>CASE 1.</u> $\int\phi f = 0$. If $\nabla\int\log f_\theta = 0$ this is fulfilled for $A(\phi,I_T)$ in the case of the Whittle estimate (Example 3). Note, that $\nabla\int\log f_\theta = 0$ holds for several parametrizations. This can be deduced from Kolmogorov's formula (cf. Brockwell and Davis, 1987, chapter 5.8).

<u>CASE 2.</u> $\kappa_4 = 0$. This condition is e.g. fulfilled if the innovations are assumed to be Gaussian.

In these cases the procedure leads to a correct approximation of the variance. In general, there is no hope for this. For instance, consider the examples of the autocovariance estimate (Example 1) and the sdf estimate (Example 2) in the non Gaussian case.

In this paper we prove that the above bootstrap can successfully be used for the important class of ratio statistics defined below. Denote the normalized spectral density by

(2.9) $$g(\alpha) \equiv f(\alpha) / F(\pi) ,$$

where F is the sdf. Analogous to definition (2.1) we consider functionals of g of the form

$$(2.10) \qquad A(\phi,g) \equiv \int_0^{\pi} \phi(\alpha)\, g(\alpha)\, d\alpha \qquad \left(\equiv \int \phi g\right)$$

where $\phi = (\phi^{(1)}, \dots, \phi^{(d)})$ and $\phi^{(r)}: [-\pi,\pi] \to \mathbf{R}$ are functions of bounded variation, $r = 1, \dots, d$.

$A(\phi,g)$ is denoted as a *normalized spectral mean*. The corresponding *normalized spectral mean estimate* is defined by

$$(2.11) \qquad A(\phi,J_T) \equiv \int_0^{\pi} \phi(\alpha)\, J_T(\alpha)\, d\alpha \qquad \left(\equiv \int \phi\, J_T\right),$$

where J_T is the normalized periodogram, i.e. $J_T(\alpha) \equiv I_T(\alpha) / F_T(\pi)$ with F_T being the integrated periodogram (see Example 2).

The estimate $A(\phi,J_T)$ can be written as a ratio of two spectral mean estimates

$$(2.12) \qquad A(\phi,J_T) = \int_0^{\pi} \phi(\alpha)\, I_T(\alpha)\, d\alpha \Big/ \int_0^{\pi} I_T(\alpha)\, d\alpha,$$

and is therefore denoted as *ratio statistic*.

<u>EXAMPLE 4 *(autocorrelation estimate)*</u>: *Let* $\phi(\alpha) = \cos(\alpha u)$, *where* $u \in \mathbf{Z}$. *Then*

$$A(\phi,J_T) = \int_{-\pi}^{\pi} I_T(\alpha) \exp(-i\alpha u)\, d\alpha \Big/ \int_{-\pi}^{\pi} I_T(\alpha)\, d\alpha \equiv c_T(u) / c_T(0) \equiv \rho_T(u)$$

is an estimator for the autocorrelation $\rho(u) \equiv c(u) / c(0)$ of lag u.

<u>EXAMPLE 5 *(normalized sdf estimate).*</u> *With* $\phi(\alpha) = \chi_{[0,\lambda]}(\alpha)$, *where* $\lambda \in [0,\pi]$, *we get*

$$A(\phi, J_T) = F_T(\lambda) / F_T(\pi),$$

the normalized integrated periodogram which represents an estimate for the normalized sdf $F(\lambda) / F(\pi)$.

Often, only the information about the normalized quantities is needed. For instance, the Yule-Walker-estimates are based on estimates of the autocorrelations and not on the autocovariances, and Bartlett's U_p-statistic for a goodness-of-fit test is based on the normalized version of F_T (cf. Dahlhaus, 1985b).

Easy calculation shows that

$$\sqrt{T}(A(\phi,J_T) - A(\phi,g)) = \frac{\sqrt{T}}{\int f \int I_T} \int \psi\, I_T$$

with $\psi = \phi \int f - \int \phi f$. Since $\int \psi f = 0$ it follows that the asymptotic distribution of $\sqrt{T}(A(\phi,J_T) - A(\phi,g))$ does not depend on the fourth order cumulant. Furthermore, it is equal to the asymptotic distribution of the corresponding bootstrap statistic

$$B(\phi,J_T^*) \equiv \frac{\pi}{n} \sum_{j=1}^{n} \phi_j\, J_j^* \qquad (2.13)$$

where $J_j^* = I_j^* / (\frac{\pi}{n} \sum_{k=1}^{n} I_k^*)$ with I_j^* defined as above. This is a necessary property for the bootstrap to be valid. In section 3 we will prove that the above bootstrap really works for ratio statistics.

We now set down the assumptions.

(A1) $\{X_t\}_{t\in \mathbf{Z}}$ is a real-valued *linear process*, i.e.

$$X_t = \sum_{u\in \mathbf{Z}} a_u\, \varepsilon_{t-u} \ ,$$

where $\{\varepsilon_t\}_{t\in \mathbf{Z}}$ are i.i.d. random variables satisfying $\mathbf{E}\,\varepsilon_1 = 0$, $\mathbf{E}\,\varepsilon_1^2 = 1$, $\mathbf{E}\,\varepsilon_t^3 = 0$, $\mathbf{E}\,\varepsilon_1^8 < \infty$. Denote by $A(\alpha) \equiv \sum_{u\in \mathbf{Z}} a_u \exp(i\alpha u)$ the *transfer function* and by $f(\alpha) \equiv (2\pi)^{-1} |A(\alpha)|^2$ the *spectral density* of $\{X_t\}$. f is non vanishing, i.e.

$$\inf_{\alpha \in [0,\pi]} f(\alpha) > 0 .$$

(A2) $\hat{f}$ is an estimate of f, which is uniformly strong consistent, i.e.

$$\sup_{\alpha \in [0,\pi]} |\hat{f}(\alpha) - f(\alpha)| \rightarrow 0 \quad \text{almost surely } (\equiv \text{ a.s.}) .$$

(A3) $\phi \equiv (\phi^{(1)}, \dots, \phi^{(d)})'$ is a d-dimensional vector of bounded functions $\phi^{(r)}: [-\pi,\pi] \to \mathbf{R}$ having bounded variation. Furthermore, $\phi^{(r)}$ is symmetric, i.e.

$$\phi^{(r)}(-\alpha) = \phi^{(r)}(\alpha), \quad r = 1, \dots, d.$$

(A4) The taper h_t is of the form $h_t = h\left(\frac{t}{T}\right)$ where h: $\mathbf{R} \to [0,1]$ is a function of bounded variation, $h(x) = 0$ for $x \notin (0,1]$ and $H_2 \equiv \int_0^1 h^2(x)\, dx > 0$.

Such a taper h is introduced in practice to reduce leakage effects (cf. Dahlhaus, 1988). In addition to (A4), we assume:

(A5) The function h is given by

$$h_\rho(x) \equiv u(x/\rho)\, \chi_{(0,\rho/2)}(x) + \chi_{[\rho/2,1-\rho/2]}(x) + u((1-x)/\rho)\, \chi_{(1-\rho/2,1]}(x),$$

where $u: [0,1/2] \to [0,1]$ is twice differentiable with bounded second derivative and $u(0) = 0$, $u(1/2) = 1$, and $0 < \rho \le 1$ denotes the proportion of the data which is tapered. Furthermore, ρ depends on T, such that $\rho_T \sim T^{-\delta}$, where $\delta < 1/6$.

Most of the tapers used in practise are of the form $h_\rho(x)$. The assumption $\rho_T = T^{-\delta}$ implies that the sequence of tapers $h_T \equiv h_{\rho_T}$ fulfills $h_T(x) \to \chi_{(0,1)}(x)$ pointwise, which is called 'asymptotically vanishing'.

To derive Edgeworth expansions we need the following assumptions in addition.

(A6) The filter coefficients $\{a_u\}$ and the Fourier coefficients $\{\widehat{\phi}(u)\}$ of ϕ decrease exponentially, i.e. for all large u,

$$|a_u| \le \rho^{|u|}\,,\;\; \|\,\widehat{\phi}(u)\,\| \le \rho^{|u|}\,,$$

where ρ is a fixed number with $0 < \rho < 1$.

(A7) $(\varepsilon_1,\varepsilon_1^2)$ satisfies Cramér's condition, i.e.

$$\exists\, \delta > 0\,,\;\; d > 0 \qquad \forall\, \|\,t\,\| \;>\; d$$
$$|\,\mathbf{E}\,\exp\,(i\,t'(\varepsilon_1,\varepsilon_1^2)')\,| \le 1 - \delta\,.$$

(A8) Denote by S_T the 8-dimensional finite Fourier transform $T^{-1/2}(d_T(\frac{2\pi}{T}\; j(1)),\,\dots\,,\,d_T(\frac{2\pi}{T}\; j(8)))'$, $(j(1),\,\dots\,,\,j(8) \in \{1,...,T/2-1\})$ or of the $(d+1)$-dimensional spectral mean estimate $\int(\phi',\,1)'\,I_T$. In both cases $\Sigma \equiv \lim_{T\to\infty} D(S_T)$ exists and is positive definite, where D denotes the dispersion matrix. Further, the matrix $W \equiv \int(\phi',1)'\,(\phi',1)\,f^2$ is positive definite.

3. The validity of the bootstrap procedure. We now prove that the bootstrap approximation holds for ratio statistics. In particular, the following theorem states that the bootstrap approximation is even better than the normal approximation. The result is proved by using Edgeworth expansions for the original and the bootstrapped statistic and by comparison of the cumulants in both expansions. The evaluation of the cumulants will give additional insight into the approximation. In particular, we will see that the skewness of the distribution is correctly approximated.

To bootstrap the distribution of $A(\phi,J_T) - A(\phi,g)$ we use the statistic $B(\phi,J_T^*) - B(\phi,\widehat{g}\,)$ where $\widehat{g}_j \;=\; \hat{f}_j \,/\, (\frac{\pi}{n}\sum_{k=1}^{n} \hat{f}_k\,)$. Furthermore, let $D_T^2 \;=\; V_T^{-1}$ where V_T is the dispersion matrix of $\sqrt{T}\,A(\phi,J_T)$ and $\widehat{D}_T^2 \;=\; \widehat{V}_T^{-1}$ where $\widehat{V}_T$ is the dispersion matrix of $\sqrt{T}\;B(\phi,J_T^*)$.

By P^* we denote the conditional distribution given the data and by E^* the corresponding conditional expectation.

THEOREM 1. *Assume* (A1) – (A8). *Then for almost all samples* $\{I_j\}$

$$\sup_{C \in \mathcal{C}^d} \Big| P\Big(\sqrt{T}\, D_T\, (A(\phi,J_T) - A(\phi,g)) \in C\Big)$$

$$- P^*\Big(\sqrt{T}\, \widehat{D}_T\, (B(\phi,J_T^*) - B(\phi,\widehat{g}\,)) \in C\Big) \Big| = o(T^{-1/2}),$$

where $\mathcal{C}^d$ *denotes the class of convex measurable* $C \subseteq \mathbf{R}^d$.

Before proving the theorem we make some comments on the result.

REMARKS. (1) The theorem says that the bootstrap approximation holds for the distribution of ratio statistics that fulfill the assumptions (A1) – (A8). Thus, the method of resampling from standardized periodogram ordinates is consistent. Furthermore, Theorem 1 gives an upper bound for the rate of convergence of the bootstrap estimate: The accuracy of the bootstrap approximation is of order less than $T^{-1/2}$ and therefore outperforms the normal approximation. This is an unexpected strong result, since the method not even imitates the covariance structure of the underlying sample.

(2) If the mean is unknown we may use $X_t - \overline{X}$ instead of X_t for the calculation of I_T . In this case we only have $\mathbf{E}\, A(\phi,I_T) = A(\phi,f) + O(T^{-1})$ which leads to the same result as in Theorem 1 with $o(T^{-1/2})$ replaced by $O(T^{-1/2})$. In this case the bootstrap approximation is at least as good as the normal approximation.

(3) The proof of the validity of the bootstrap procedure in this section reveals that all results hold, if we replace the variables $\{\varepsilon_j^*\}$ by variables drawn from the known asymptotic distribution . By Theorem 1 we know that this modified procedure is accurate up to order $o(T^{-1/2})$ as well as the original one. On the other hand the formal Edgeworth expansions show that for both methods the approximation is not better than $O(T^{-1})$, since the fourth order cumulants of bootstrapped and unbootstrapped terms do not match. Unfortunately, higher order asymptotics does not detect differences between both approaches (as conjectured in Franke and Härdle (1992)) and none of them is preferable. However, there exists a significant difference between both methods. The modified procedure avoids one potential error source: resampling from the *studentized* periodogram ordinates. At this stage the bias of the sd estimate influences the method heavily. It seems reasonable to avoid this danger and to use the information about the asymptotic distribution of the residuals.

(4) Returning to the examples it is to say that Theorem 1 is applicable for the autocorrelation

estimate (Example 4) If e.g. the underlying process is an ARMA(p,q)process, then all assumptions can be fulfilled including the technical assumption (A6). This assumption causes some trouble for the normalized sdf estimate (Example 5). The reason is that for $\phi = \chi_{[0,\lambda]}$ the Fourier coefficients do not decrease exponentially. This problem can be solved by modifying the estimate with a smoothed function ϕ. On the other hand the authors conjecture that the Edgeworth expansion is also valid under a weaker condition than (A6). But the proof seems to be rather complicated. In section 4 we consider Whittle estimates (Example 3) in some detail.

PROOF OF THEOREM 1. Let

$$V_{T,r} = \sqrt{T}(A(\phi^{(r)},J_T) - A(\phi^{(r)},g))$$

and

$$V^*_{T,r} = \sqrt{T}(B(\phi^{(r)},J^*_T) - B(\phi^{(r)}, \hat{g})) \quad .$$

We need Edgeworth expansions for $V_{T,r}$ and $V^*_{T,r}$ as proved in Theorem 2 and 3 below. We then only have to verify for the occuring expansion terms

$$\Lambda_{T,3}(C) - \Lambda^*_{T,3}(C) = o(T^{-1/2})$$

uniformly in C (see Bhattacharya and Ranga Rao, 1976, pp. 51 - 57, for the definition of $\Lambda^*_{T,3}$. Götze and Hipp, 1983, p. 217 may be consulted for a definition of $\Lambda_{T,3}$). Since $\Lambda_{T,3}$ and $\Lambda^*_{T,3}$ only differ from each other by coefficients that depend on the underlying distributions through their cumulants up to and including order three, this holds, if

$$\mathbf{cum}\,(V_{T,r}) = \mathbf{cum}^*(V^*_{T,r}) + o(T^{-1/2}) \qquad \text{a.s.}$$
$$\mathbf{cum}\,(V_{T,r}, V_{T,s}) = \mathbf{cum}^*(V^*_{T,r}, V^*_{T,s}) + o(1) \qquad \text{a.s.}$$
$$\mathbf{cum}\,(V_{T,r}, V_{T,s}, V_{T,t}) = \mathbf{cum}^*(V^*_{T,r}, V^*_{T,s}, V^*_{T,t}) + o(T^{-1/2}) \text{ a.s. } .$$

Since V_T and V^*_T are ratio statistics these cumulants are difficult to calculate. However, due to Bhattacharya and Ghosh (1978, Theorem 2(b)) , it is sufficient if we prove these equation for stochastic approximations $W_{T,r}$ and $W^*_{T,r}$ with

$$W_{T,r} = V_{T,r} + o_p(T^{-1/2})$$
$$W^*_{T,r} = V^*_{T,r} + o_{p*}(T^{-1/2})$$

This is done in Theorem 4 and 5 below. Several lemmata provide the technical details.

THEOREM 2. *Assume* (A1) – (A8). *Then the following approximation holds uniformly over convex measurable* $C \subseteq \mathbf{R}^d$:

$$P\left(\sqrt{T}\, D_T(A(\phi,J_T) - A(\phi,g)) \in C\right) = \Lambda_{T,3}(C) + o(T^{-1/2}).$$

PROOF. From Janas (1993, Theorem 2.3) we obtain the following Edgeworth expansion for the statistic $A(\tilde{\phi},I_T)$ where $\tilde{\phi}' = (\phi',1)$.

$$P(\sqrt{T}\,(A(\tilde{\phi},I_T) - \mathbf{E}\, A(\tilde{\phi},I_T)) \in C) = \Psi_{T,s}(C) + o(T^{-(s-2)/2}).$$

(cf. Götze and Hipp, 1983, p. 217, for the definition of $\Psi_{T,s}$). Lemma 1 (below) yields $\mathbf{E}A(\phi,I_T) = A(\phi,f) + o(T^{-1})$ and we can therefore replace $\mathbf{E}A(\phi,I_T)$ by $A(\phi,f)$ in this expansion for s = 3. We now apply the Transformation-Lemma of Bhattacharya and Ghosh (1978, Lemma 2.1) with the transforming function $H(x_1, \dots, x,y) = (\frac{x_1}{y}, \dots, \frac{x_p}{y})$. We have to check that this function is sufficiently smooth in a neighborhood of $\mu = \int(\phi',1)'\, f$ and that (grad H) (μ) has full rank p. The first statement follows from the positivity of $F(\pi) = \int f$, the variance of the underlying process $\{X_t\}$, the second is trivial. ❑

THEOREM 3. *Assume* (A1) – (A8) . *Then for almost all samples* $\{I_j\}$ *and uniformly over convex measurable* $C \subseteq \mathbf{R}^d$:

$$P^*(\sqrt{T}\, \hat{D}_T\left(B(\phi,J^*_T) - B(\phi,\hat{g})\right) \in C) = \Lambda^*_{T,3}(C) + o(T^{-1/2}).$$

PROOF. We only sketch the proof. As in Theorem 2 we first establish an expansion for the statistic $B(\tilde{\phi},I^*_T)$ with $\tilde{\phi}' = (\phi', 1)$ and then apply the Transformation-Lemma to get the expansion for the ratio $B(\phi, J_T^*)$. $B(\phi, I_T^*)$ is a weighted mean of independent and identically distributed random variables with common distribution function $\tilde{F}_n$, the empirical distribution function of the rescaled studentized periodogram ordinates $\{\tilde{\varepsilon}_j\}$ defined in step (4) of the bootstrap procedure. Corollary 1 (below) shows the weak convergence of $\tilde{F}_n$ to an exponential distribution. Now we can prove the edgeworth expansion as it was done for the ordinary sample mean in Babu and Singh (1984). Only two changes are required: The cumulants have to be replaced by averaged cumulants (cf. Bhattacharya and Ranga Rao, 1976, p. 71) and Cramér's condition has to be modified in an obvious way. ❑

We now construct stochastic approximations of $V_{T,r}$ and $V_{T,r}^*$ for which we check afterwards the required equality of the cumulants. We start with an approximation for $V_{T,r}$. Let

(3.1) $$\Delta^{(r)} \equiv (\phi^{(r)} - \int \phi^{(r)} g) g$$

where $g = f / \int f$ is the normalized spectral density. Then we have

$$V_{T,r} = T^{1/2} \left(\frac{\int \phi^{(r)} I_T}{\int I_T} - \frac{\int \phi^{(r)} f}{\int f} \right)$$

$$= T^{1/2} \frac{\int (\phi^{(r)} - \int \phi^{(r)} g) \, g(I_T / f)}{\int g(I_T / f)}$$

$$= T^{1/2} \frac{\int \Delta^{(r)} (I_T / f)}{\int g(I_T / f)}$$

$$= T^{1/2} \int \Delta^{(r)} (I_T / f) (2 - \int g(I_T / f)) + o_p(T^{-1/2})$$

The last equation follows since $\frac{1}{x} = 2 - x + o(|x - 1|)$. Note the equality

(3.2) $$\int \Delta^{(r)} = 0$$

which is of particular importance as is seen later. Define the above approximation of $V_{T,r}$ as $W_{T,r}$, i.e.

(3.3) $$W_{T,r} \equiv T^{1/2} \int \Delta^{(r)}(I_T / f)\,(2 - \int g(I_T / f)).$$

To calculate the first three cumulants of $W_{T,r}$ we need the following Lemma.

LEMMA 1. Suppose ψ_j are bounded functions and (A.1), (A.4) and (A.5) hold. Then we have

(i) $\mathbf{cum}\,(\int \psi_1 I_T) = \int \psi f + o(T^{-1})$

(ii) $\mathbf{cum}\,(\int \psi_1 I_T , \int \psi_2 I_T) = O(T^{-1})$

(iii) $\mathbf{cum}\,(\int \psi_1 I_T , \ldots, \int \psi_l I_T) = o(T^{-l/2}) \quad (l \geq 3).$

PROOF. We only give a sketch. The exponential decrease of a_u implies $\sum_u |u|^{1+\gamma} |c(u)| < \infty$ for any $\gamma \in (0,1)$ and therefore also

$$|f(\alpha + \beta) - f(\alpha) - \beta f'(\alpha)| \leq K|\beta|^{1+\gamma}$$

with some constant K. Then we obtain

$$\left|\mathbf{cum}\,(\int\psi_1 I_T) - \int \psi f \right| = \left| \frac{1}{2}\int_{-\pi}^{\pi} \psi_1(\alpha)\int_{-\pi}^{\pi} \{f(\alpha + \beta) - f(\alpha)\} \frac{|H_{1,T}(\beta)|^2}{2\pi H_{2,T}} d\beta d\alpha \right|$$

$$\leq K\int_{-\pi}^{\pi} |\beta|^{1+\gamma} \frac{|H_{1,T}(\beta)|^2}{H_{2,T}} d\beta \leq O(T^{-1-\gamma+4\delta}) = o(T^{-1})$$

by using Lemma 5.4 of Dahlhaus (1988) (note that the taper of (A5) is of degree (1,2δ) in the terminology of that paper). The proof of (ii) and (iii) is standard (cf. Dahlhaus, 1983, Lemma 6 and Lemma 7).

The application of this lemma leads to the following expressions.

(3.4) $\quad \mathbf{cum}\,(W_{T,r}) = -\,T^{+1/2}\,\mathbf{cum}\,(\int \Delta^{(r)}\,(I_T / f)\,,\int g(I_T / f)) + o(T^{-1/2})$

(3.5) $\quad \mathbf{cum}\,(W_{T,r}\,,W_{T,s}) \;=\; T\,\mathbf{cum}\,(\int \Delta^{(r)}\,(I_T / f)\,,\int \Delta^{(s)}\,(I_T / f)) + o(T^{-1/2})$

(3.6) $$\mathbf{cum}\,(W_{T,r},W_{T,s},W_{T,t}) = T^{3/2}\,\mathbf{cum}\,(\int \Delta^{(r)}(I_T / f)\,,\int \Delta^{(s)}(I_T / f)\,,\int \Delta^{(t)}(I_T / f))$$

$$-\,2\,T^{3/2}\,\mathbf{cum}\,(\int \Delta^{(r)}(I_T / f),\int g(I_T / f))\;\;\mathbf{cum}\,(\int \Delta^{(s)}(I_T / f)\,,\int \Delta^{(t)}(I_T / f)\,)$$

$$-\,2\,T^{3/2}\mathbf{cum}\,(\,\int \Delta^{(s)}(I_T / f)\,,\int g(I_T / f)\,)\;\;\mathbf{cum}\,(\,\int \Delta^{(r)}(I_T / f)\,,\int \Delta^{(t)}(I_T / f))$$

$$-\,2\,T^{3/2}\mathbf{cum}\,(\int \Delta^{(t)}(I_T / f),\int g(I_T / f)\,)\,\mathbf{cum}\,(\,\int \Delta^{(r)}(I_T / f),\int \Delta^{(s)}(I_T / f)) + o(T^{-1/2}))$$

Observe that all cumulants can be expressed in terms of cumulants of second and third order of statistics of the form $\int \phi^{(r)}(I_T / f)$ with $\phi^{(r)} \in \{\Delta^{(r)}, g\}$. For an asymptotically vanishing taper these cumulants are given in the next lemma.

LEMMA 2. Under (A1), (A3) and (A5) we have

(i) $\quad T \cdot \mathbf{cum}\,(\int \phi^{(r)}(I_T / f)\,,\int \phi^{(s)}(I_T / f))$

$$= 2\pi \int \phi^{(r)}\,\phi^{(s)} \;+ (\kappa_4/\sigma^4)\int \phi^{(r)}\int \phi^{(s)} \;+\; o(1)\;,$$

(ii) $\quad T^2 \cdot \mathbf{cum}\,(\int \phi^{(r)}(I_T / f)\,,\int \phi^{(s)}(I_T / f)\,,\int \phi^{(t)}(I_T / f))$

$$= 8\,\pi^2 \int \phi^{(r)}\,\phi^{(s)}\,\phi^{(t)}$$

$$+ (\kappa_4/\sigma^4)\,\cdot\,4\pi\,(\int \phi^{(r)}\,\phi^{(s)}\int \phi^{(t)} + \int \phi^{(r)}\,\phi^{(t)}\int \phi^{(s)} + \int \phi^{(s)}\,\phi^{(t)}\int \phi^{(r)})$$

$$+ (\kappa_6/\sigma^6)\int \phi^{(r)}\int \phi^{(s)}\int \phi^{(t)} + o(1)\;,$$

where $\kappa_4(\kappa_6)$ is the fourth (sixth) cumulant of the innovations ε_t .

PROOF. The lemma is proved by using the product theorem for cumulants (cf. Brillinger, 1981, Theorem 2.3.2) and applying again Lemma 5.4 of Dahlhaus (1988a). We omit details.

For the calculation of the cumulants of $W_{T,r}$ we only need three kinds of cumulants of the statistics $\int \Delta(I_T / f)$ and $\int g(I_T / f)$. For those cumulants we obtain by Lemma 2 and the equation $\int \Delta^{(r)} = 0$ the following expressions:

$$(3.7) \qquad T \cdot \mathbf{cum}\,\left(\int \Delta^{(r)}(I_T / f), \int g(I_T / f)\right) = 2\pi \int \Delta^{(r)} g + o(1),$$

$$(3.8) \qquad T \cdot \mathbf{cum}\,\left(\int \Delta^{(r)}(I_T / f)\,, \int \Delta^{(s)}(I_T / f)\right) = 2\pi \int \Delta^{(r)}\,\Delta^{(s)} + o(1)\,,$$

$$(3.9) \qquad T^2 \cdot \mathbf{cum}\,\left(\int \Delta^{(r)}(I_T / f)\,, \int \Delta^{(s)}(I_T / f), \int \Delta^{(t)}(I_T / f)\,\right) = 8\pi^2 \int \Delta^{(r)}\,\Delta^{(s)}\Delta^{(t)} + o(1)\,.$$

Note that these cumulants are independent from the cumulants κ_4 and κ_6 of the innovations. Furthermore, all contributions which stem from the dependence structure of the periodogram ordinates cancel by means of the central equation (3.2). For the above cumulants it makes no difference if we replace the dependent rvs $\{I_j / f_j\}$ by i.i.d. rvs $\{E_j\}$ where E_j is exponentially distributed. Later, we will see that the bootstrap counterparts of $\{I_j / f_j\}$ behave similar as independent and exponentially distributed variables, and that the corresponding bootstrap cumulants have the same limits as above. We now summarize our results on the cumulants of $W_{T,r.}$

THEOREM 4. *Under* (A1), (A3) *and* (A5) *we have*

$$\mathbf{cum}(W_{T,r}) = -\,T^{-1/2}\, 2\pi \int \Delta^{(r)}\, g + o(T^{-1/2})\,,$$

$$\mathbf{cum}(W_{T,r}\,, W_{T,s}) = 2\pi \int \Delta^{(r)}\,\Delta^{(s)} + o(1)\,,$$

$$\mathbf{cum}\,(W_{T,r}\,, W_{T,s}, W_{T,t}) = T^{-1/2}\;8\pi^2 \left(\int \Delta^{(r)}\,\Delta^{(s)}\,\Delta^{(t)} - \int \Delta^{(r)} g \int \Delta^{(s)}\Delta^{(t)}\right.$$

$$- \int \Delta^{(s)} g \int \Delta^{(r)} \Delta^{(t)} - \int \Delta^{(t)} g \int \Delta^{(r)} \Delta^{(s)} \,) \; + \; o(T^{-1/2}),$$

where $\Delta^{(r)} \equiv (\phi^{(r)} - \int \phi^{(r)} g)\, g$, $g \equiv f / F(\pi)$.

Before calculating the bootstrap cumulants of the corresponding approximation of $V^*_{T,r}$ we set down some auxiliary results. Let G_n denote the empirical distribution function of $\{I_j / f_j\}$and $G(x) = 1 - \exp(-x)$. The next lemma provides a Glivenko-Cantelli lemma for $\{I_j / f_j\}$.

LEMMA 3. *Assume* (A1), (A4) *and* (A6) – (A8). *Then*

(i) $$\sup_{x \in \mathbf{R}} |G_n(x) - G(x)| \to 0 \qquad \text{a.s.} \;,$$

and

(ii) $$\int_0^\infty g(x)\, dG_n(x) \to \int_0^\infty g(x)\, dG(x) \qquad \text{a.s.},$$

for every function g(x) *which is piecewise uniformly continuous and satisfies*

$$\sup_{x \in \mathbf{R}} |g(x)| / (1 + |x|^8) < \infty \;.$$

PROOF. The proof is analogous to the proof of Theorem 1 in Chen and Hannan (1980). However, due to the data taper we have to replace the required Edgeworth expansion by the expansion given in Theorem 4.3 of Janas and von Sachs (1993).

From Lemma 3 we deduce the following corollary. Let $\widehat{F}_n$ ($\widetilde{F}_n$) denote the empirical distribution function of $\{\hat{\varepsilon}_j\}$ ($\{\tilde{\varepsilon}_j\}$) . $F_n \Rightarrow F$ means that the distribution F_n converges weakly to F (F_n may be random).

COROLLARY 1. *Under* (A1), (A2), (A4) *and* (A6) – (A8) *we have*

(i) $$\frac{1}{n} \sum_{j=1}^{n} \hat{\varepsilon}_j^p \to \mathbf{E}\, \chi_1^p \quad \text{a.s.} \quad \text{and} \quad \frac{1}{n} \sum_{j=1}^{n} \tilde{\varepsilon}_j^p \to \mathbf{E}\, \chi_1^p \qquad \text{a.s.} \;,$$

for all $p \le 8$ *where* $\chi_1 \sim G$, *and*

(ii) $\widehat{F}_n \Rightarrow G$ a.s. *and* $\widetilde{F}_n \Rightarrow G$ a.s..

The proof is given in the appendix.

The corollary says that the bootstrap distribution $\widetilde{F}_n$ converges to the exponential distribution in the Mallow's metric d_p defined as in Bickel and Freedman (1981). The exponential distribution is absolutely continuous and fulfills Cramér's condition, which is important in the context of Edgeworth expansions. From the first part of the corollary and by the linearity of the cumulants we obtain the next lemma on the cumulants of the bootstrapped statistic $B(\phi, I_T^*)$.

LEMMA 4. *Under* (A1) – (A4), *and* (A6) – (A8) *we have for all* $p \le s$,

$$T^{p-1} \cdot \mathbf{cum}^*\Big(\frac{\pi}{n} \sum_{j=1}^{n} \phi_j^{(r_1)}\, I_j^*, \ldots , \frac{\pi}{n} \sum_{j=1}^{n} \phi_j^{(r_p)}\, I_j^* \Big) \to (p-1)!\,(2\pi)^{p-1} \int \prod_{j=1}^{p} \phi^{(r_j)} f^p \qquad \text{a.s.}$$

where $r_1, \ldots , r_p \in \{1, \ldots , d\}$.

PROOF. The result follows by using straightforward calculations.

The lemma shows that the cumulants of order $p \ge 2$ of the bootstrapped statistic $B(\phi, I_T^*)$ do not converge to the same limit as the cumulants of $A(\phi, I_T)$. We know already the reason why the bootstrap approximation fails in this situation: Independent resampling does not take care of the dependence structure among the basic sample.

However, for ratio statistics we now prove that the bootstrap approximation with independent resampling is sufficient. Let $W_{T,r}^*$ denote the bootstrap version of $W_{T,r}$, i.e.

$$W_{T,r}^* \equiv T^{1/2} \frac{\pi}{n} \sum_{j=1}^{n} \widehat{\Delta}_j^{(r)} \frac{I_j^*}{\widehat{f}_j} \Big(2 - \frac{\pi}{n} \sum_{j=1}^{n} \widehat{g}_j \frac{I_j^*}{\widehat{f}_j}\Big)$$

where $\widehat{\Delta}_j^{(r)} \equiv \Big(\phi_j^{(r)} - \frac{\pi}{n} \sum_{k=1}^{n} \phi_k^{(r)}\, \widehat{g}_k\Big)\, \widehat{g}_j$, $\widehat{g}_j \equiv \widehat{f}_j \Big/ \Big(\frac{\pi}{n} \sum_{k=1}^{n} \widehat{f}_k\Big)$.

Replacing the cumulants in (3.7) - (3.9) by the corresponding bootstrap cumulants from Lemma 4 and proceeding as in the proof of Theorem 4 (with the integrals replaced by sums) leads to the following result on the bootstrap cumulants.

THEOREM 5. *Under* (A1) – (A4) and (A6) - (A8) *we have*

$$\mathbf{cum}^*(W^*_{T,r}) = -\; T^{-1/2}\, 2\pi \int \Delta^{(r)}\, g + o(T^{-1/2}) \text{ a.s. },$$

$$\mathbf{cum}^*(W^*_{T,r}\,,\, W^*_{T,s}) = 2\pi \int \Delta^{(r)}\, \Delta^{(s)} + o(1) \qquad \text{a.s. },$$

$$\mathbf{cum}^*(W^*_{T,r}\,,\, W^*_{T,s}\,,\, W^*_{T,t}) = T^{-1/2}\; 8\pi^2 \Big(\int \Delta^{(r)}\Delta^{(s)}\Delta^{(t)} - \int \Delta^{(r)}\; g \int \Delta^{(s)}\, \Delta^{(t)}$$

$$- \int \Delta^{(s)}\; g \int \Delta^{(r)}\Delta^{(t)} - \int \Delta^{(t)}\; g \int \Delta^{(r)}\, \Delta^{(s)} \Big) + o(T^{-1/2}) \text{ a.s.}$$

Since the first three cumulants of $W_{T,r}$ and $W^*_{T,r}$ are the same we have established Theorem 1.

4. Whittle estimates. Whittle estimates are based on the periodogram (Whittle, 1953). They are obtained by minimizing the distance $\mathcal{L}_T(\theta)$ of Example 3 between the periodogram and the parametric form of the spectral densitiy. A detailed discussion may be found in Dzhaparidze and Yaglom (1983).

Suppose $\{X_t\}_{t\in \mathbf{Z}}$ is a linear process with spectral density f that fulfills the assumptions (A1) and (A6) and we fit a parametric model $\mathcal{F} = \{f_\theta : \theta \in \Theta\}$ to the data. Suppose $\Theta = (\sigma^2, \tau)$, $f_\theta = \sigma^2 h_\tau$ and Kolmogorov's formula holds, i.e.

$$\int_{-\pi}^{\pi} \log f_\theta\,(\alpha)\, d\alpha = 2\pi \log \frac{\sigma^2}{2\pi}$$

(which is e.g. true for ARMA-models - c.f. Brockwell and Davies, 1987, chapter 5.8). We do not assume that $f \in \mathcal{F}$, i.e. we allow that the model is misspecified.

The Whittle estimate $\hat{\theta}_T = (\hat{\sigma}^2_T, \hat{\tau}_T)$ is determined by minimizing the Whittle function $L_T(\theta)$ of Example 3, i.e. it fulfills the equations

$$\int_0^\pi I_T(\alpha)\ \nabla_\tau f^{-1}_{\hat{\theta}_T}(\alpha)\, d\alpha = 0$$

and

$$\frac{1}{\pi} \int_0^\pi f_{\hat{\theta}_T}(\alpha)^{-1}\, I_T(\alpha)\, d\alpha = 1\ .$$

It is known that $\hat{\theta}_T$ converges to $\theta_0 = (\sigma^2_0, \tau_0)$ which minimizes the corresponding theoretical function

$$L(\theta) = \frac{1}{2\pi} \int_0^\pi \{\log f_\theta(\alpha) + f_\theta(\alpha)^{-1}\, f(\alpha)\}\, d\alpha\ ,$$

i.e. θ_0 is determined by the equations

$$\int_0^\pi f(\alpha)\ \nabla_\tau f^{-1}_{\theta_0}(\alpha)\, d\alpha = 0$$

and

$$\frac{1}{\pi} \int_0^\pi f_{\theta_0}(\alpha)^{-1}\, f(\alpha)\, d\alpha = 1.$$

The bootstrap version of $L_T(\theta)$ is

$$L^*_T(\theta) = \frac{1}{2} \log \frac{\sigma^2}{2\pi} + \frac{1}{2n} \sum_{j=1}^{n} f_\theta^{-1}\left(\frac{2\pi j}{T}\right) I^*_j$$

and the bootstrap Whittle estimate $\theta^* = (\sigma^{2*},\tau^*)$ is determined by minimizing $L_T^*(\theta)$ which leads to the equations

$$\frac{1}{n}\sum_{j=1}^{n} I_j^* \nabla_\tau f_{\theta^*}^{-1}\left(\frac{2\pi j}{T}\right) = 0$$

and

$$\frac{1}{n}\sum_{j=1}^{n} f_{\theta^*}^{-1}\left(\frac{2\pi j}{T}\right) I_j^* = 1 \quad .$$

It is heuristically obvious that $\theta^* - \bar{\theta}$ will converge to zero where $\bar{\theta} = (\bar{\sigma}^2,\bar{\tau})$ is obtained by minimizing

$$\bar{L}_T(\theta) = \frac{1}{2}\log\frac{\sigma^2}{2\pi} + \frac{1}{2n}\sum_{j=1}^{n} f_\theta^{-1}\left(\frac{2\pi j}{T}\right)\hat{f}_j$$

where $\hat{f}$ is the (nonparametric) estimate of the bootstrap procedure. $\bar{\theta}$ fulfills the same equations as θ^* with I^* replaced by $\hat{f}$. (Intuitively, one might expect $\hat{\theta}$ as the limit of θ^*. The limit $\bar{\theta}$ is a consequence of the bootstrap which implies that $\mathbf{E}^* I_j^*$ is equal to $\hat{f}$ and not equal to I_j). Note that θ^* and $\bar{\theta}$ depend on T.

The heuristics in section 2 indicates that the bootstrap is valid for the parameter τ (since $\int f\ \nabla_\tau f_{\theta_0}^{-1} = 0$). This will be proved below. However, the bootstrap does not work for the parameter σ^2 unless $\kappa_4 = 0$ (e.g. if the innovations are Gaussian).

We restrict ourselves to the one-dimensional case. However, we conjecture that an analogous result also holds in the general case. To eliminate the dependence of the parameter σ^2 we note that $\hat{\tau}_T$ $(\tau_0, \tau^*, \bar{\tau})$ are also the minima of $L_T(\tau) \equiv \int I_T\, h_\tau^{-1}$ $(L(\tau) \equiv \int f\ h_\tau^{-1}$, $L_*(\tau) \equiv \frac{\pi}{n}\Sigma_j I_j^*\, h_\tau\left(\frac{2\pi j}{T}\right)^{-1}$, $\bar{L}(\tau) \equiv \frac{\pi}{n}\Sigma_j \hat{f}_j\ h_\tau\left(\frac{2\pi j}{T}\right)^{-1}$ respectively). We need the following assumption in addition to (A1) to (A8).

(A9) The set of parameters $\mathcal{T} \subset \mathbf{R}$ is compact. The parameters are identifiable, i.e. $\tau_1 \neq \tau_2$ implies $h_{\tau_1} \neq h_{\tau_2}$ on a set with positive Lebesgue measure.The function $h_\tau(\alpha)$ is four times continuously differentiable with respect to $\tau \in \mathcal{T}$ and two times continuously differentiable with respect to $\alpha \in [0,\pi]$. $h_\tau(\alpha)$ and its derivatives are uniformly bounded, i.e. $\exists\ 0 < \underline{c} \le \bar{c} < \infty\ \ \forall\ \tau \in \mathcal{T}\ , \alpha \in [0,\pi]$

$$\underline{c} \le h_\tau(\alpha) \le \bar{c}\ ,\ \left| \left(\frac{\partial}{\partial \tau}\right)^i h_\tau^{-1}(\alpha) \right| \le \bar{c}\ (i = 1, \ldots, 4) \text{ and } \left| \left(\frac{\partial}{\partial \alpha}\right)^j h_\tau(\alpha) \right| \le \bar{c}\ (j = 1,2).$$

Let $\phi_\tau = (\phi_\tau^{(1)}, \phi_\tau^{(2)}, \phi_\tau^{(3)})$ with $\phi_\tau^{(i)} \equiv \left(\frac{\partial}{\partial \tau}\right)^i h_\tau^{-1}$ (i = 1,2,3) .There exists $d_0 > 0$ such that $K(\tau) \equiv L^{(2)}(\tau) = \int \phi_\tau^{(2)} f \ge d_0$ for all $\tau \in \mathcal{T}$.

Furthermore, in (A8) we have to replace $\int (\phi',1)'\, I_T$ by $\int \phi_\tau\, I_T$ and to define the weight matrix W as $\int \phi_\tau \phi_\tau' f^2$. In addition let $J(\tau) \equiv 2\pi \int (\phi_\tau^{(1)} f)^2$, $J_*(\tau) \equiv T\mathbf{E}^*(L_*^{(1)}(\tau))^2$ and $K_*(\tau) \equiv \mathbf{E}^* L_*^{(2)}(\tau)$.

THEOREM 6. *Assume that* (A1) – (A9) *hold*. *Then for almost all samples* $\{I_j\}$

$$\sup_{x \in \mathbf{R}} \left| \mathbf{P}((TK^2(\tau_0)/J(\tau_0))^{1/2}(\hat{\tau}_T - \tau_0) \le x) - \mathbf{P}^*((TK_*^2(\bar{\tau})/J_*(\bar{\tau}))^{1/2}(\tau^* - \bar{\tau}) \le x) \right| = o(T^{-1/2})$$

The proof is transferred to the appendix.

REMARK. Without proof we remark that the bootstrap also works for σ^2 if $\kappa_4 = 0$. In the case of an AR(p)-model θ^* is the Yule Walker estimate with the covariances

$$c^*(u) = \frac{2\pi}{n} \sum_{j=1}^{n} I_j^* \cos\left(\frac{2\pi j}{T} u\right) .$$

5. Practical considerations and simulations examples. We now report on two simulation examples and make remarks on the design of the bootstrap with respect to the estimate $\hat{f}$ and to data tapers.

A natural condidate for $\hat{f}$ seems to be a kernel estimate as suggested in Franke and Härdle (1992). However, our simulations with kernel estimates were not convincing. It is usually recommended for a bootstrap in nonparametric regression to choose a bandwidth which is a bit larger than the optimal one (cf. Franke and Härdle, 1992). However, choosing a large bandwidth leads to a strong bias in the neighbourhood of peaks in the spectrum. We are convinced that this trade-off is the reason for our bad results.

The behaviour of the estimates becomes much better if one smoothes the log periodogram (note that the log transformation is asymptotically variance stabilizing). In the simulations below we used a kernel estimate with Epanechnikov kernel. A bias correction is obtained from the following heuristic consideration. Suppose Z_j are iid exponentially distributed random variables with (constant) mean $f(\lambda)$ (this is the asymptotic distribution of the periodogram ordinates in a local neighbourhood) and w_j are the kernel weights. Then

$$E\left(\exp\left(\Sigma\, w_j \log Z_j\right)\right) = \prod_j E\, Z_j^{w_j} = f(\lambda) \prod_j \Gamma\left(1 + w_j\right)$$

where Γ is the Gamma function. We therefore estimate f_k by

$$\exp\left(\sum_j \left[w_j \log I_{k+j} - \log \Gamma\left(1 + w_j\right)\right]\right)$$

where $w_j = \frac{1}{b} K\left(\frac{1}{b}\,\frac{2\pi j}{T}\right)$ and $K(x) = \frac{3}{4}\pi\left(1 - \left(\frac{x}{\pi}\right)^2\right)\ \{|x| \leq \pi\}$.

With this estimate the results turned out to be quite good. In particular they were insensitive with respect to the choice of b. We therefore chose b by "eye inspection" having in mind that the bandwidth in the bootstrap step should be a bit larger than the optimal one (with respect to the mean square error).

Our theoretical results only hold for an asymptotically vanishing taper (which is a realistic assumption). Since a taper is very often essential to obtain reasonable results for small samples we recommend to correct for the taper by using

$$\sqrt{T H_{4,T}}\ /\ H_{2,T}\left(B\left(\phi, J_T^*\right) - B\left(\phi, \hat{g}\right)\right)$$

as the bootstrap estimate for the distribution of

$$A\left(\phi, J_T\right) - A\left(\phi, g\right) .$$

The additional factor tends to 1 for an asymptotically vanishing taper and to a correct (first order!) bootstrap approximation in the general case.

In the first example we considered the estimate for the autocorrelation function from Example 4. Samples of size 64 of the AR(1) process

$$X_t = aX_{t-1} + \varepsilon_t$$

with $a = 0.9$ and ε_t uniform on $[-\sqrt{3}, \sqrt{3}\,]$ were considered. A 10% Tukey Hanning taper was applied and the bandwidth of the above estimate was chosen to be $b = 0.1$. Figure 1 shows the logarithm of the periodogram of the sample together with the logarithm of the kernel estimate.

$\rho_T(1)$ is also the Yule-Walker estimate for a. It is known that

$$\sqrt{T}\left(\rho_T(1) - \rho(1)\right) \xrightarrow{D} N\left(0, \left(H_4 / H_2^2\right)\left(1 - a^2\right)\right)$$

where $H_k = \lim T^{-1} H_{k,T}$. In Figure 2 this asymptotic distribution is shown as the dashed line. The solid line is the true distribution (simulated with 2000 replications). The dotted line is a "typical" bootstrap approximation with the frequency bootstrap as described above calculated from 2000 bootstrap samples. The corresponding plots for eight additional original processes can be found in Figure 6. The bootstrap approximation is always better than the asymptotic distribution. In particular it gives a good bias correction.

The second example shows the bootstrap in a much more complicated situation. T = 64 observations of an ARMA(4,2) process with AR-roots $0.9^{-1}e^{i\,0.2\pi}$, $0.9^{-1}e^{-i\,0.2\pi}$, $0.9^{-1}e^{i\,0.5\pi}$, $0.9^{-1}e^{-i\,0.5\pi}$, MA-roots $0.8^{-1}e^{i\,0.35\pi}$, $0.8^{-1}e^{-i\,0.35\pi}$ and uniform innovations on $[-\sqrt{3}, \sqrt{3}\,]$ were generated. A (misspecified) AR(4)-model was fitted to the data and the Whittle estimate $\hat{a}_T$ for the parameters with a 10% Tukey-Hanning taper was calculated (in this case the Whittle estimate is identical to the Yule-Walker estimate). Our goal is now to estimate the distribution of the Mahalanobis distance in this misspecified situation.

If the AR(4)-model were correct we would have

$$\sqrt{T}\left(\hat{a}_T - a_0\right) \xrightarrow{D} N\left(0, \left(H_4 / H_2^2\right)\sigma^2\Sigma^{-1}\right)$$

where Σ is the covariance matrix. In the misspecified case a similar result holds where a_0 now is the minimizer of $\mathcal{L}(\theta)$ (cf. section 4) and a different limit covariance matrix. In that case a_0 is the best approximating value.

Suppose we want to construct a confidence set for a_O. In the correct specified case the above result implies for the Mahalanobis distance

$$\frac{T\, H_{2,T}^2}{H_{4,T}} (\hat{a}_T - a_0)' \frac{1}{\sigma^2}\Sigma\, (\hat{a}_T - a_0) \xrightarrow{D} \chi_4^2.$$

Replacing $\frac{1}{\sigma^2}\Sigma$ by a consistent estimate leads to an asymptotic confidence set for a_0. As proved in Section 4 $\hat{a}_T$ is approximately a ratio statistic. Futhermore, we may estimate

$$\frac{1}{c(0)}\Sigma_{ij} \qquad \text{by} \qquad \rho_T(i-j)$$

and

$$\sigma^2 / c(0) \qquad \text{by} \qquad 1 - \sum_{j=1}^{p} \hat{a}_j \rho_T(j)$$

which again are ratio statistics. Therefore, it is heuristically clear that the frequency bootstrap also works for the above Mahalanobis distance.

Figure 3 shows the logarithm of the true spectral density of the ARMA(4,2)-process (connected crosses) and the parametric AR(4)-spectral density estimate (solid line). Figure 4 shows the tapered periodogram with the kernel estimate as discussed above where b = 0.05. In Figure 5 the asymptotic χ_4^2 distribution of the above distance is shown as the dashed line. The solid line is the true distribution of the statistic with $\frac{1}{\sigma^2}\Sigma$ replaced by the above estimates (in the misspecified situation!). It was obtained by simulation with 2000 samples. The dotted line again is a "typical" bootstrap approximation with the frequency bootstrap calculated from 2000 bootstrap samples. The corresponding plots for eight additional original processes can be found in Figure 7.

Only the fourth picture of Figure 7 shows a bad result. In this case the nonparametric estimate showed a third (small) peak and one of the two peaks of the fitted bootstrap-AR(4)-model sometimes fell on that small peak resulting in a large Mahalanobis distance.

In the other cases the bootstrap distribution is quite close to the true one. Since the bootstrap is a nonparametric bootstrap it can be used to estimate also the effects due to model-misspecification.

It is obvious that more simulation studies are needed. In particular it would be interesting to see how the above bootstrap compares to an AR(∞) bootstrap or to a block-bootstrap.

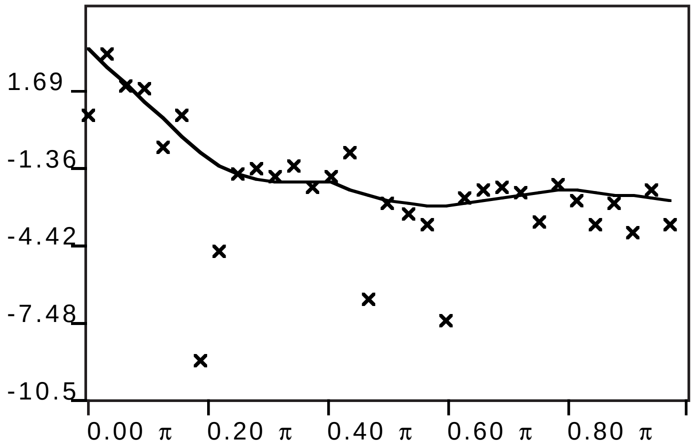

Figure 1. Log-periodogram and kernel estimate for an AR(1)-process

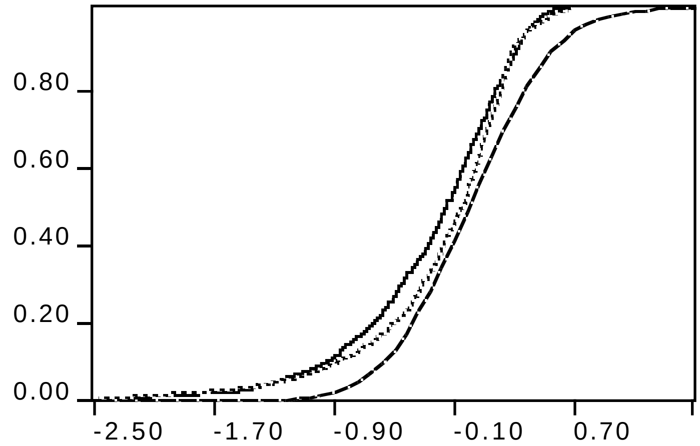

Figure 2. True (solid), asymptotic (dashed) and bootstrap (dotted) distribution of the first order correlation for an AR(1)-process

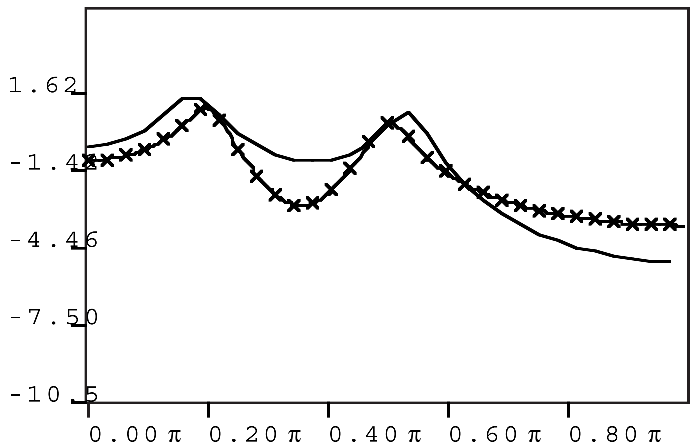

Figure 3. Log-spectrum of an ARMA(4,2)-process (crosses) and of an AR(4)-fit (solid line)

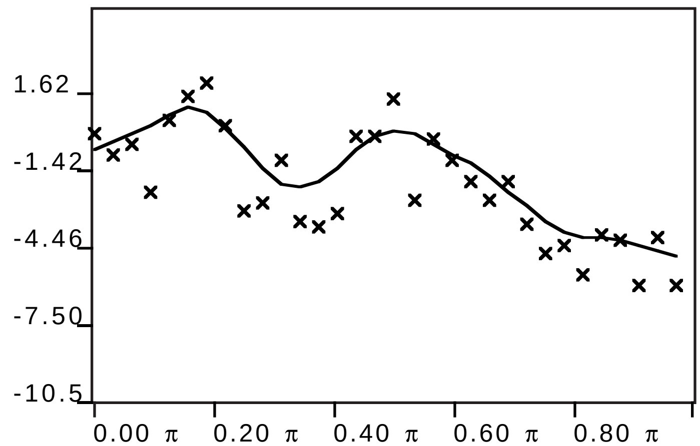

Figure 4. Log-periodogram and kernel estimate for an ARMA(4,2)-proces

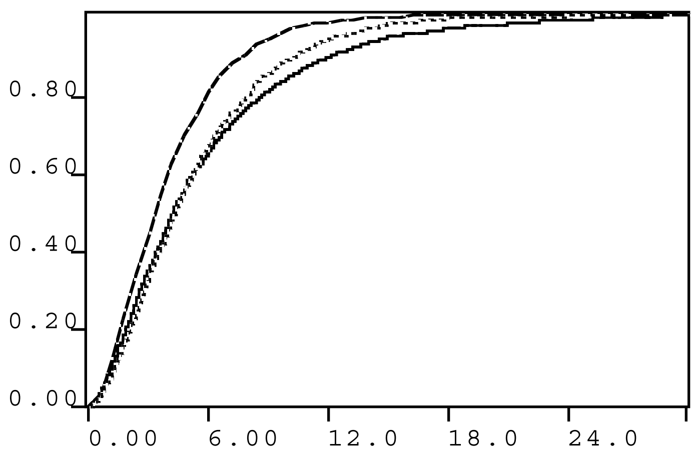

Figure 5. True (solid), asymptotic (dashed) and bootstrap (dotted) distribution of the Mahalanobis-distance for an AR(4)-model

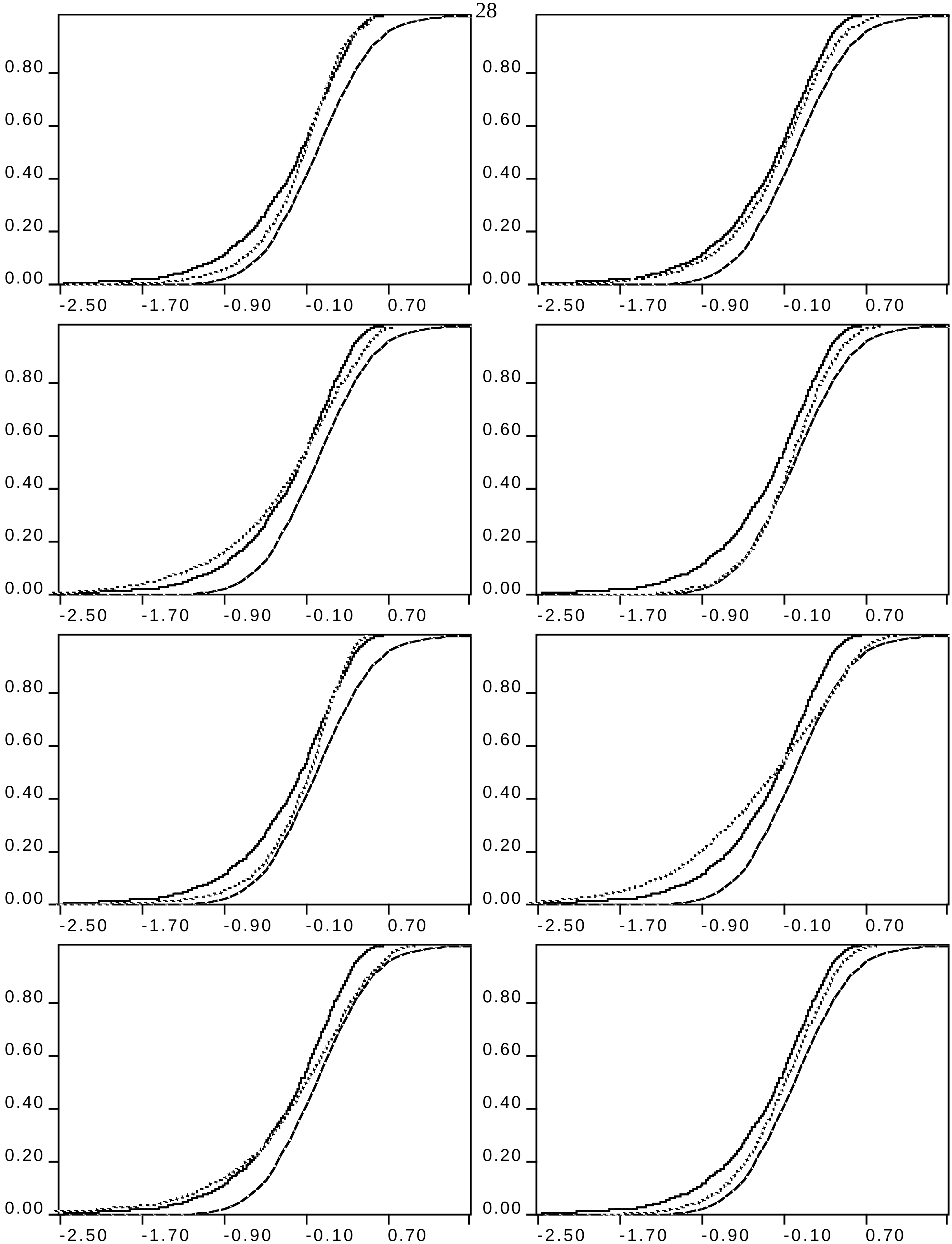

Figure 6. True (solid), asymptotic (dashed) and bootstrap (dotted) distribution of the first order correlation

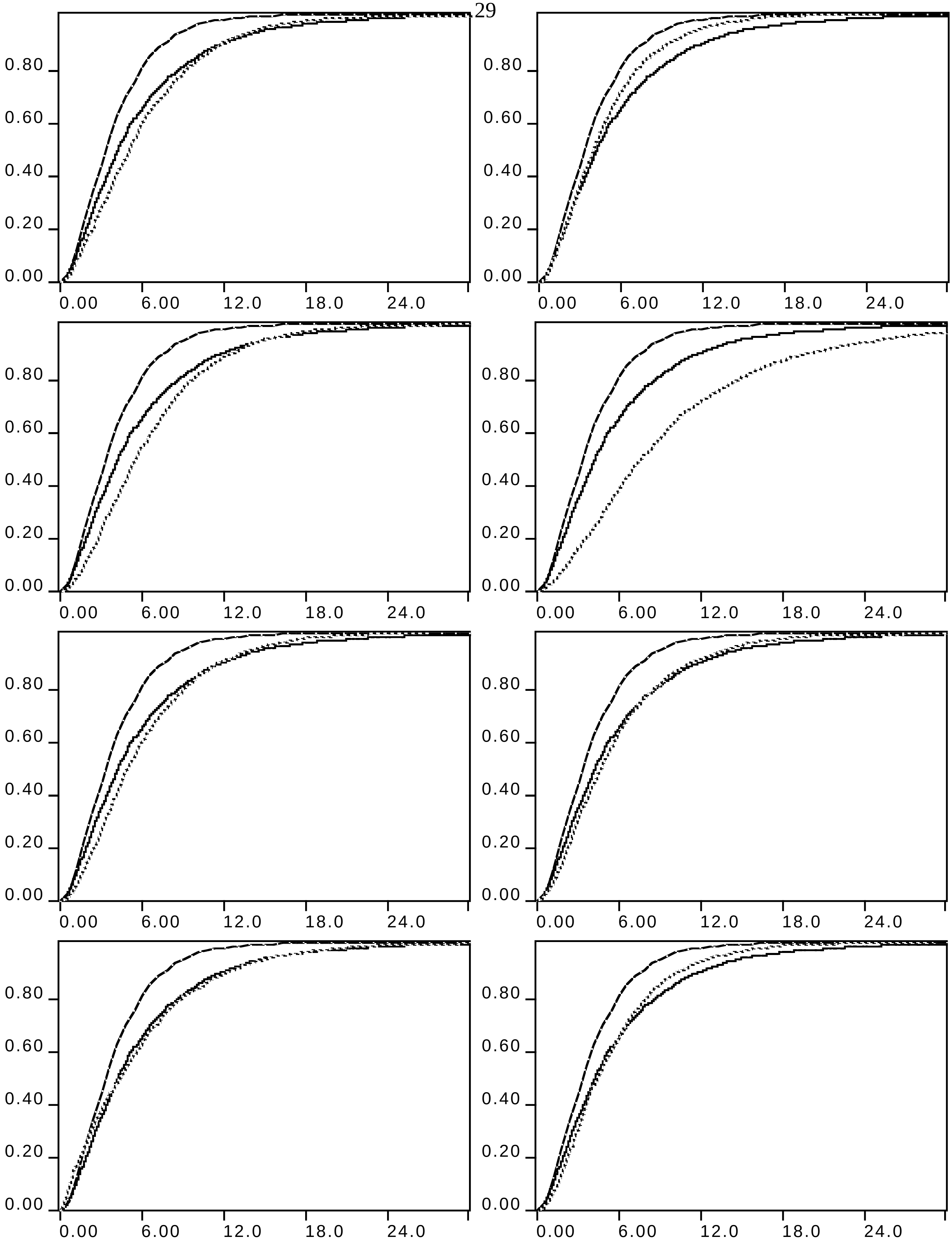

Figure 7. True (solid), asymptotic (dashed) and bootstrap (dotted) distribution of the Mahalanobis-distance of an

Appendix. PROOF OF COROLLARY 1. As in Bickel and Freedman (1981) we introduce the metric d_p as a measure for the distance between distributions F and G, for which the p-th absolute moment exists.

(1) $$d_p(F,G) \equiv \inf\{\mathbf{E}\,|X - Y|^p\}^{1/p} ,$$

where the infinum is taken over all pairs of random variables X and Y having marginal distributions F and G, respectively. For p = 2 this metric is the Mallow's metric. We write $d_p(X,Y)$ instead of $d_p(F,G)$.

We show that

(2) $$d_p(\chi_1,\varepsilon_1^*) \rightarrow 0 \qquad \text{a.s.} \; .$$

By the triangle-inequality we have

(3) $$d_p(\chi_1,\varepsilon_1^*) \le d_p(\chi_1,\varepsilon_1^0) + d_p(\varepsilon_1^0,\hat{\varepsilon}_1) + d_p(\hat{\varepsilon}_1,\varepsilon_1^*),$$

where the df of $\varepsilon_1^0(\hat{\varepsilon}_1)$ is the edf $G_n(\widehat{F}_n)$ of the true residuals $\{I_j / f_j\}$ (of the unscaled empirical residuals $\{I_j / \hat{f}_j\}$). We prove that all three terms on the right-hand side of (3) converge to zero almost surely.

For the first term the assertion follows from Lemma 3. To get an upper bound for $d_p(\varepsilon_1^0,\hat{\varepsilon}_1)$ we choose the joint distribution of $(\varepsilon_1^0,\hat{\varepsilon}_1)$ such that it assumes the value $(I_j / f_j, I_j / \hat{f}_j)$ with probability n^{-1}, j = 1, … , n . Then,

$$\begin{aligned} d_p(\varepsilon_1^0,\hat{\varepsilon}_1)^p &\le \frac{1}{n}\sum_{j=1}^{n}\left|\frac{I_j}{f_j} - \frac{I_j}{\hat{f}_j}\right|^p \\ &= \frac{1}{n}\sum_{j=1}^{n}\left(\frac{I_j}{f_j}\right)^p\left|1 - \frac{f_j}{\hat{f}_j}\right|^p \\ &\le \sup_j\left|1 - \frac{f_j}{\hat{f}_j}\right|^p \cdot \frac{1}{n}\sum_{j=1}^{n}\left(\frac{I_j}{f_j}\right)^p . \end{aligned}$$

By Lemma 3 (ii) we obtain

$$\frac{1}{n} \sum_{j=1}^{n} \left(\frac{I_j}{f_j}\right)^p \rightarrow \mathbf{E}\ \chi_1^p \ < \ \infty \quad \text{a.s..}$$

Therefore, the second term in (3) converges to zero by the convergence of the estimate $\hat{f}$ to f (Assumption A2).

Using exactly the same argument as above, we also get

$$\begin{aligned} d_p\ (\hat{\varepsilon}_1, \varepsilon_1^*)^p \ &\leq \ \frac{1}{n} \sum_{j=1}^{n} |\hat{\varepsilon}_j - \tilde{\varepsilon}_j|^p \\ &= \ \frac{1}{n} \sum_{j=1}^{n} \left| \frac{\hat{\varepsilon}_j(\hat{\varepsilon}_\bullet - 1)}{\hat{\varepsilon}_\bullet} \right|^p \\ &= \ \left| \frac{1}{n} \sum_{j=1}^{n} \hat{\varepsilon}_j \ - \ 1 \right|^p \cdot \frac{1}{n} \sum_{j=1}^{n} \hat{\varepsilon}_j^p \ / \left(\frac{1}{n} \sum_{j=1}^{n} \hat{\varepsilon}_j \right)^p \\ &\rightarrow \ 0 \quad \text{a.s.,} \end{aligned}$$

by Lemma 3 (ii) and Assumption A2. ❐

PROOF OF THEOREM 6 . As in the proof of Theorem 1 we derive Edgeworth expansions for the distribution of the Whittle estimate and for its bootstrapped version. Then the result follows by a comparison of the corresponding coefficients of the polynomials occuring in these expansions. The Edgeworth expansion for the Whittle estimate is given in Janas (1993, Theorem 3.1) for the case where the model is correctly specified ($f = f_{\theta_0}$). The proof for the more general case discussed here is exactly the same. The expansion for the bootstrap counterpart can be deduced in a similar way. Therefore we only mention the essential steps.

We set down

$$(4) \qquad V_* \ \equiv \ \sqrt{T}(\tau^* - \bar{\tau})$$

(5) $$Z_*^{(i)}(\tau) \equiv \sqrt{T}(L_*^{(i)}(\tau) - \mathbf{E}\, L_*^{(i)}(\tau))$$

(6) $$K_*(\tau) \equiv - \mathbf{E}^* L_*^{(2)}(\tau)$$

and

(7) $$U_*(\tau) \equiv - \frac{Z_*^{(1)}(\tau)}{K_*(\tau)} + \frac{1}{\sqrt{T}} \frac{Z_*^{(1)}(\tau)}{K_*(\tau)} \frac{Z_*^{(2)}(\tau)}{K_*(\tau)} - \frac{1}{2\sqrt{T}} \frac{Z_*^{(3)}(\tau)}{K_*(\tau)} \left(\frac{Z_*^{(1)}(\tau)}{K_*(\tau)}\right)^2$$

We will show that the following stochastic expansion holds:

(8) $$V_* = U_*(\bar{\tau}) + \frac{1}{T} \xi_* ,$$

where ξ_* satisfies $\mathbf{P}^*(|\xi_*| > \rho_T \sqrt{T}) = o(T^{-1/2})$ a.s. for some sequence $\rho_T \to 0$, $\rho_T \sqrt{T} \to \infty$ as $T \to \infty$.

By a lemma of Chibisov (cf. Janas , 1993, Lemma 4.5), the Edgeworth expansions for V_* and $U_*(\bar{\tau})$ match up to order $T^{-1/2}$. But the Edgeworth expansion for $U_*(\bar{\tau})$ follows from Theorem 1 by the Transformation-Lemma of Bhattacharya and Ghosh (cf. Janas , 1993).

For the proof of (8) we consider the following Taylor expansion. Since $L_*^{(1)}(\tau^*) = 0$, we have

(9) $$0 = \sqrt{T}\, L_*^{(1)}(\bar{\tau}) + \frac{1}{\sqrt{T}} Z_*^{(2)}(\bar{\tau})\, V_* + K_*(\bar{\tau})\, V_* + \frac{1}{2\sqrt{T}} L_*^{(3)}(\bar{\tau}) V_*^2 + \frac{1}{6T} L_*^{(4)}(\tilde{\tau}) V_*^3 ,$$

where $|\tilde{\tau} - \bar{\tau}| \le |\tau^* - \bar{\tau}|$. We rewrite (9) as

(10) $$V_* = - \frac{\sqrt{T}\, L_*^{(1)}(\bar{\tau})}{K_*(\bar{\tau})} - \frac{Z_*^{(2)}(\bar{\tau})}{K_*(\bar{\tau})\sqrt{T}} V_* - \frac{L_*^{(3)}(\bar{\tau})}{2K_*(\bar{\tau})\sqrt{T}} V_*^2 - \frac{L_*^{(4)}(\tilde{\tau})}{6K_*(\bar{\tau})T} V_*^3 .$$

The following bounds for tail probabilities can be derived analogous to the corresponding bounds in Janas (1993).

For every $\alpha > 0$ there exist positive constants d_1, d_2 and d_3 such that

(11) $$\mathbf{P}^*(|\tau^* - \bar{\tau}| > d_1\, T^{\alpha - 1/2}) = o(T^{-1/2}) \quad \text{a.s.},$$

(12) $\mathbf{P}^*(|L_*^{(i)}(\tau) - \mathbf{E}^* L_*^{(i)}(\tau)| > d_2\, T^{\alpha-1/2}) = o(T^{-1/2})$ a.s. , for i = 1,2,3

(13) $$\mathbf{P}^*(\sup_{\tau \in \mathcal{T}} | L_*^{(4)}(\tau) | > d_3\, T^{\alpha}) = o(T^{-1/2}) \quad \text{a.s..}$$

By (11) – (13) with $0 < \alpha < 1/10$, we can write (10) as

(14) $$V_* = -\frac{\sqrt{T}\, L_*^{(1)}(\tau)}{K_*(\tau)} + \frac{1}{\sqrt{T}}\, \tilde{\xi}_* \ ,$$

where $\mathbf{P}^*(| \tilde{\xi}_* | > d_4\, T^{2\alpha}) = o(T^{-1/2})$ a.s. , for some $d_4 > 0$.

Substituting (14) for the right-hand side of (10), and noting $\mathbf{E}^* L_*^{(1)}(\tau) = 0$ a.s., we have

$$V_* = -\frac{Z_*^{(1)}(\tau)}{K_*(\tau)} + \frac{1}{\sqrt{T}}\, \frac{Z_*^{(1)}(\tau)}{K_*(\tau)}\, \frac{Z_*^{(2)}(\tau)}{K_*(\tau)} - \frac{1}{2\sqrt{T}}\, \frac{L_*^{(3)}(\tau)}{K_*(\tau)} \left(\frac{Z_*^{(1)}(\tau)}{K_*(\tau)} \right)^2 + \frac{1}{T}\, \xi_* \ ,$$

where $\mathbf{P}^*(| \xi_* | > d_5\, T^{3\alpha}) = o(T^{-1/2})$, for some $d_5 > 0$.

Acknowledgement. The computations were done together with M. Diller by using a version of SPECSIM which is based on the framework of RANDOM&TEMPLATE.

References.

Babu, G. J. and Singh, K. (1984). On one-term Edgeworth correction by Efron's bootstrap. *Sankhyà*, Ser. A **46** 219 - 232.

Bhattacharya, R. N. and Ranga Rao, R. (1976). Normal Approximation and Asymptotic Expansions. Wiley, New York.

Bhattacharya, R. N. and Ghosh, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 434 - 451.

Bickel, P. J. and Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* **9** 1196 - 1217.

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

Brockwell, P. J. and Davis, R. A. (1987). Time Series: Theory and Methods. Springer Verlag.

Chen, Z. - G. and Hannan, E. J. (1980). The distribution of periodogram ordinates. *J. Time Ser. Anal.* **1** 73 - 82.

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

Dahlhaus, R. (1985a). Asymptotic normality of spectral estimates. *J. Mult. Anal.* **16** 412 - 431.

Dahlhaus, R. (1985b). On the asymptotic distribution of Bartlett's U_p-statistic. *J. Time Ser. Anal.* **6** 213 - 227.

Dahlhaus, R. (1988). Small sample effects in time series analysis: a new asymptotic theory and a new estimate. *Ann. Statist.* **16** 808-841.

Dzhaparidze, K. O. and Yaglom, A. M. (1983). Spectrum parameter estimation in time series analysis. In: *Krishnaiah, P. R.. Developments in Statistics.* Vol. **4** (Academic Press, New York).

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.* **7** 1 - 26.

Efron, B. and Tibshirani, R. J. (1986). Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy (with discussion). *Statist. Sci.* **1** 54 - 77.

Franke, J. and Härdle, W. (1992). On bootstrapping kernel spectral estimates. *Ann. Statist.* **20** 121 - 145.

Freedman, D. A. (1984). On bootstrapping two-stage least-squares estimates in stationary linear models. *Ann. Statist.* **12** 827 - 842.

Götze, F. and Hipp, C. (1983). Asymptotic expansions for sums of weakly dependent random vectors. *Zeit. Wahr. verw. Gebiete* **64** 211 - 239.

Hurvich, C.M. and Zeger, S. (1987). Frequency domain bootstrap methods for time series. New York University. Statistics and Operations Research area working paper.

Janas, D. (1993). Edgeworth expansions for spectral mean estimates with applications to Whittle estimates. Ann. Inst. Statist. Math. **46** 667 - 682.

Janas, D. and Dahlhaus, R. (1994). A frequency domain bootstrap for time series. In: Computationally Intensive Statistical Methods. Proc. of the 26th Symp. on the Interface (John Sall, Ann Lehman, eds.), 423 - 425.

Janas, D. and von Sachs, R. (1993). Consistency for non-linear functions of the periodogram of tapered data. Preprint, Universität Heidelberg.

Kreiss, J. P. and Franke, J. (1992). Bootstrapping stationary autoregressive moving-average models. *J. Time Ser. Anal.* **13** 297 - 317.

Künsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *Ann. Statist.* **17** 1217 - 1241.

Liu, R. Y. and Singh, K. (1992). Moving blocks jackknife and bootstrap capture weak dependence. In: Exploring the limits of bootstrap (R. LePage, L. Billard, eds.), 225 - 248, Wiley, New York.

Nordgaard, A. (1992). Resampling stochastic processes using a bootstrap approach. In: Bootstrapping and Related Techniques (K.H. Jöckel, G. Rothe, W. Sendler, eds.) Lecture Notes in Econometrics and Mathematical Systems 376, Springer-Verlag, Berlin.

Singh, K. (1981). On the asymptotic accuracy of Efron's bootstrap. *Ann. Statist.* **9** 1187 - 1195.

Swanepoel, J. W. H. and van Wyk, J. W. Y. (1986). The bootstrap applied to power spectral density function estimation. *Biometrika* **73** 135 - 141.

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