

## 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 behavior 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, for example, 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. setup [Singh (1981) and 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., residuals). 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 resampling 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 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), and Nordgaard (1992)]. For obvious

---

Received June 1995; revised March 1996.

AMS 1991 *subject classifications*. Primary 62M10; secondary 62E20.

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

reasons this method may be denoted as a frequency domain bootstrap, whereas all procedures above resample in the time domain. Franke and Härdle (1992) 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 individual 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 in more detail the class of estimators for which this bootstrap with periodogram ordinates works. 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 these 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 notation 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 deferred to the Appendix.

**2. Preliminaries.** Consider a real-valued stationary time series  $\{X_t\}_{t \in \mathbb{Z}}$

with  $\mathbb{E} X_t = 0$  and spectral density (sd)  $f$ . Let us denote by

$$(2.1) \quad \begin{aligned} & A(\phi, f) \\ & \equiv \left( \int_0^\pi \phi^{(1)}(\alpha) f(\alpha) d\alpha, \dots, \int_0^\pi \phi^{(d)}(\alpha) f(\alpha) d\alpha \right)' \quad \left( \equiv \int \phi f \right) \end{aligned}$$

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

$$(2.2) \quad \begin{aligned} & A(\phi, I_T) \\ & \equiv \left( \int_0^\pi \phi^{(1)}(\alpha) I_T(\alpha) d\alpha, \dots, \int_0^\pi \phi^{(d)}(\alpha) I_T(\alpha) d\alpha \right)' \quad \left( \equiv \int \phi I_T \right), \end{aligned}$$

where  $I_T(\alpha)$  is the *tapered periodogram* that is,

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

where

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

denotes the *tapered finite Fourier transform* with data taper  $h_t$  (see Assumptions 5 and 6),

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

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

The following special cases are covered by this class.

EXAMPLE 1 (Autocovariance estimate). Let  $\phi(\alpha) = 2 \cos(\alpha u)$ ,  $u \in \mathbb{Z}$ . Then

$$\begin{aligned} A(\phi, I_T) &= \int_{-\pi}^\pi I_T(\alpha) \exp(-i\alpha u) d\alpha \\ &= (H_{2,T})^{-1} \sum_{t=1}^T h_t X_t h_{t+u} X_{t+u} \equiv c_T(u) \end{aligned}$$

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

EXAMPLE 2 (Spectral distribution function estimate). With  $\phi(\alpha) = \chi_{[0, \lambda]}(\alpha)$ , where  $\chi_{[0, \lambda]}$  is the indicator function on  $[0, \lambda]$ , 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 (sdf)  $F(\lambda) = \int_0^\lambda f(\alpha) d\alpha$ .

EXAMPLE 3 (Whittle estimate). Let  $\mathcal{F} = \{f_\theta: \theta \in \Theta\}$ ,  $\Theta \subset \mathbb{R}^p$ , be a parametric family of spectral densities. Then the parameter  $\theta$  may be estimated by minimizing the Whittle likelihood

$$\mathcal{L}_T(\theta) = (2\pi)^{-1} \int_0^\pi \left\{ \log f_\theta(\alpha) + \frac{I_T(\alpha)}{f_\theta(\alpha)} \right\} d\alpha.$$

With  $\phi \equiv \nabla f_\theta^{-1}$  [where  $\nabla = (\partial/\partial\theta_1, \dots, \partial/\partial\theta_p)'$  and  $f_\theta^{-1} = 1/f_\theta$ ], we have

$$\nabla \mathcal{L}_T(\theta) = 0 \quad \Leftrightarrow \quad A(\theta, I_T) - A(\phi, f) = 0,$$

where  $f$  is the true spectral density.

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  $\alpha_j \neq 0 \pmod{\pi}$  asymptotically independent exponential variables [cf. Brillinger (1981), Theorem 5.2.6]. This suggests the following bootstrap procedure. Let  $n = \lfloor T/2 \rfloor$  and  $I_j = I_T(2\pi j/T)$ .

*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  $\hat{\varepsilon}_\bullet = (1/n)\sum_{j=1}^n \hat{\varepsilon}_j$ .
5. Draw independent bootstrap replicates  $\{\varepsilon_j^*\}$  from the empirical distribution of the  $\tilde{\varepsilon}_j$ .
6. Define bootstrap periodogram values by  $\{I_j^*\} \equiv \{\hat{f}_j \varepsilon_j^*\}$ .

REMARK 1. The rescaling in step 4 avoids an unnecessary bias at the resampling stage.

REMARK 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

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

and  $\phi_j \equiv \phi(2\pi j/T)$ .

To get an idea on the quality of this bootstrap approximation, we study the asymptotic behavior 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  $\{X_t\}$  (cf. Assumption 1) and  $h_t \equiv 1$ , the asymptotic variance is given by

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

where  $\kappa_4$  is the fourth cumulant and  $\sigma^2$  is the variance of the innovations  $\varepsilon_t$ .

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

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

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 0.

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

CASE 2.  $\kappa_4 = 0$ . This condition is fulfilled, for example, 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 be used successfully for the important class of ratio statistics defined below. Denote the normalized spectral density by

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

where  $F$  is the sdf. We consider the functionals  $A(\phi, g) \equiv \int \phi g$ , where  $\phi = (\phi^{(1)}, \dots, \phi^{(d)})$  and  $\phi^{(r)}: [0, \pi] \rightarrow \mathbb{R}$  are functions of bounded variation ( $r = 1, \dots, d$ ).

Note that  $A(\phi, g)$  is the *normalized spectral mean*. The corresponding *normalized spectral mean estimate* is defined by

$$(2.10) \quad A(\phi, J_T) \equiv \int_0^\pi \phi(\alpha) J_T(\alpha) d\alpha \quad \left( \equiv \int \phi J_T \right),$$

where  $J_T$  is the normalized periodogram, that is,  $J_T(\alpha) \equiv I_T(\alpha)/F_T(\pi)$  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

$$(2.11) \quad A(\phi, J_T) = \int_0^\pi \phi(\alpha) I_T(\alpha) d\alpha \bigg/ \int_0^\pi I_T(\alpha) d\alpha,$$

and is therefore denoted as the *ratio statistic*.

EXAMPLE 4 (Autocorrelation estimate). Let  $\phi(\alpha) = \cos(\alpha u)$ , where  $u \in \mathbb{Z}$ . Then

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

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

EXAMPLE 5 (Normalized sdf estimate). 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 of autoregressive parameters 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

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

with  $\psi = \phi f f - \int \phi f$ . 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

$$(2.12) \quad B(\phi, J_T^*) \equiv \frac{\pi}{n} \sum_{j=1}^n \phi_j J_j^*,$$

where

$$J_j^* = I_j^* \bigg/ \left( \frac{\pi}{n} \sum_{k=1}^n I_k^* \right),$$

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.

ASSUMPTION 1.  $\{X_t\}_{t \in \mathbb{Z}}$  is a real-valued *linear process*, that is,

$$X_t = \sum_{u \in \mathbb{Z}} a_u \xi_{t-u},$$

where  $\{\xi_t\}_{t \in \mathbb{Z}}$  are i.i.d. random variables satisfying  $\mathbb{E} \xi_1 = 0$ ,  $\mathbb{E} \xi_1^2 = 1$ ,  $\mathbb{E} \xi_1^8 < \infty$  and  $\sum_u u^2 |a_u| < \infty$ . Denote by  $A(\alpha) \equiv \sum_{u \in \mathbb{Z}} a_u \exp(i \alpha u)$  the *transfer function* and by  $f(\alpha) \equiv (2\pi)^{-1} |A(\alpha)|^2$  the *spectral density* of  $\{X_t\}$ . It is assumed that  $\inf_{\alpha \in [0, \pi]} f(\alpha) > 0$ .

ASSUMPTION 2. The third moment of the innovations vanishes, that is  $\mathbb{E} \xi_t^3 = 0$ .

ASSUMPTION 3.  $\hat{f}$  is an estimate of  $f$ , which is uniformly strongly consistent, that is,

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

ASSUMPTION 4.  $\phi \equiv (\phi^{(1)}, \dots, \phi^{(d)})'$  is a  $d$ -dimensional vector of bounded functions  $\phi^{(r)}: [0, \pi] \rightarrow \mathbb{R}$  having bounded variation. For convenience, we assume that  $\phi^{(r)}$  is extended to the real line with  $\phi^{(r)}(-\alpha) = \phi^{(r)}(\alpha)$  and  $\phi^{(r)}(\alpha + 2\pi) = \phi^{(r)}(\alpha)$ ,  $r = 1, \dots, d$ .

ASSUMPTION 5. The taper  $h_t$  is of the form  $h_t = h(t/T)$ , where  $h: \mathbb{R} \rightarrow [0, 1]$  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 Assumption 5, we assume the following.

ASSUMPTION 6. The function  $h$  is given by  $h_t = h(t/T) = h^{(\rho_T)}(t/T)$ , where

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

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

Most of the tapers used in practice are of the form  $h^{(\rho)}(x)$ . The assumption  $\rho_T = T^{-\delta}$  implies that the sequence of tapers fulfills  $h^{(\rho_T)}(x) \rightarrow \chi_{(0, 1)}(x)$  pointwise, which is called ‘‘asymptotically vanishing.’’

To derive Edgeworth expansions, we need the following assumptions.

ASSUMPTION 7. The filter coefficients  $\{a_u\}$  and the Fourier coefficients  $\{\hat{\phi}(u)\}$  of  $\phi$  decrease exponentially that is, for all large  $u$ ,

$$|a_u| \leq \tau^{|u|}, \quad \|\hat{\phi}(u)\| \leq \tau^{|u|},$$

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

ASSUMPTION 8.  $(\xi_1, \xi_1^2)$  satisfies Cramér’s condition that is, there exist  $\delta > 0$  and  $d > 0$  such that, for all  $\|t\| > d$ ,

$$|\mathbb{E} \exp(it'(\xi_1, \xi_1^2)')| \leq 1 - \delta.$$

ASSUMPTION 9. Denote by  $S_T$  the eight-dimensional finite Fourier transform

$$T^{-1/2} \left( d_T \left( \frac{2\pi}{T} j(1) \right), \dots, d_T \left( \frac{2\pi}{T} j(8) \right) \right)', \quad (j(1), \dots, j(8)) \in \{1, \dots, T/2 - 1\}$$

or the  $(d + 1)$ -dimensional spectral mean estimate  $\int (\phi', 1)' I_T$ . In both cases  $\Sigma \equiv \lim_{T \rightarrow \infty} D(S_T)$  exists and is positive definite, where  $D$  denotes the dispersion matrix. Furthermore, the matrix  $W = \int (\phi', 1)' (\phi', 1) f^2$  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  $B(\phi, J_T^*) - B(\phi, \hat{g})$ , where

$$\hat{g}_j = \hat{f}_j / \left( \frac{\pi}{n} \sum_{k=1}^n \hat{f}_k \right).$$

Furthermore, let  $D_T^2 = V_T^{-1}$ , where  $V_T$  is the dispersion matrix of  $\sqrt{T}A(\phi, J_T)$  and  $\hat{D}_T^2 = \hat{V}_T^{-1}$ , where  $\hat{V}_T$  is the dispersion matrix of  $\sqrt{T}B(\phi, J_T^*)$ .

By  $P^*$  we denote the conditional distribution given the data and by  $\mathbb{E}^*$  the corresponding conditional expectation.

**THEOREM 1.** *Suppose Assumptions 1–9 hold. Then for almost all samples  $\{I_j\}$ ,*

$$\sup_{C \in \mathcal{C}^d} \left| P(\sqrt{T}D_T(A(\phi, J_T) - A(\phi, g)) \in C) - P^*(\sqrt{T}\hat{D}_T(B(\phi, J_T^*) - B(\phi, \hat{g})) \in C) \right| = o(T^{-1/2}).$$

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

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

REMARK 3. The theorem says that the bootstrap approximation holds for the distribution of ratio statistics that fulfill Assumptions 1–9. 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 unexpectedly strong result, since the method does not even imitate the covariance structure of the underlying periodogram sample.

REMARK 4. If the mean is unknown we may use  $X_t - \bar{X}$  instead of  $X_t$  for the calculation of  $I_T$ . Unfortunately, this causes an extra bias of the periodogram and we have only  $\mathbb{E}A(\phi, I_T) = A(\phi, f) + O(T^{-1})$ . This implies that Theorem 2 holds only with  $O(T^{-1/2})$ . Furthermore, (3.2) no longer holds—we have only  $\text{cum}(V_{T,r}) = O(T^{-1/2})$  and  $\text{cum}^*(V_{T,r}^*) = O(T^{-1/2})$  a.s., which leads to the same result as in Theorem 1 with  $o(T^{-1/2})$  replaced by  $O(T^{-1/2})$ . Thus the bootstrap approximation is in this case at least as good as the normal approximation and we may hope that it still does better in practice. At least it is better in certain misspecified situations (cf. the second example in Section 5).

REMARK 5. As indicated in the proof of Theorem 4, the above result does not hold if Assumption 2 is violated, that is, if  $\mathbb{E}\xi_t^3 \neq 0$ . In this case (3.4) no longer holds. Both sides of (3.4) are of order  $O(T^{-1/2})$  which leads to the result of Theorem 1 with  $o(T^{-1/2})$  replaced by  $O(T^{-1/2})$ .

REMARK 6. 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 spectral density estimate influences the method heavily. It seems reasonable to avoid this danger and to use the information on the asymptotic distribution of the residuals.

REMARK 7. Returning to the examples, it is to say that Theorem 1 is applicable for the autocorrelation estimate (Example 4). If, for example, the underlying process is an ARMA( $p, q$ ) process, then all assumptions can be fulfilled including the technical one, Assumption 7. 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  $\phi$ . On the other hand, the authors conjecture that the Edgeworth expansion is also valid under a weaker condition than Assumption 7. 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

$$V_T = \sqrt{T} (A(\phi, J_T) - A(\phi, g))$$

and

$$V_T^* = \sqrt{T} (B(\phi, J_T^*) - B(\phi, \hat{g})).$$

We need Edgeworth expansions for  $D_T V_T$  and  $\hat{D}_T V_T^*$  as proved in Theorems 2 and 3 below. We then have only to verify for the occurring expansion terms

$$(3.1) \quad \Lambda_{T,3}(C) - \Lambda_{T,3}^*(C) = o(T^{-1/2})$$

uniformly in  $C$ . These expressions are the usual terms in an Edgeworth expansion [cf. Bhattacharya and Ranga Rao (1976), pages 51–57]. In this special situation  $\Lambda_{T,3}(\cdot)$  is the signed measure with density

$$\left[ 1 - \sum_{r=1}^d \text{cum}((D_T V_T)_r) \frac{\partial}{\partial x_r} - \frac{1}{6} \sum_{r,s,t=1}^d \text{cum}((D_T V_T)_r, (D_T V_T)_s, (D_T V_T)_t) \frac{\partial}{\partial x_r} \frac{\partial}{\partial x_s} \frac{\partial}{\partial x_t} \right] \prod_{j=1}^d \varphi(x_j),$$

where  $\varphi(x)$  is the density of the standard Gaussian distribution, and  $\Lambda_{T,3}^*(\cdot)$  is the signed measure with the analogous density with  $\text{cum}^*((\hat{D}_T V_T^*)_r)$  and  $\text{cum}^*((\hat{D}_T V_T^*)_r, (\hat{D}_T V_T^*)_s, (\hat{D}_T V_T^*)_t)$ . In order to verify (3.1) we therefore have to prove

$$(3.2) \quad \text{cum}(V_{T,r}) = \text{cum}^*(V_{T,r}^*) + o(T^{-1/2}) \quad \text{a.s.},$$

$$(3.3) \quad \text{cum}(V_{T,r}, V_{T,s}) = \text{cum}^*(V_{T,r}^*, V_{T,s}^*) + o(1) \quad \text{a.s.},$$

$$(3.4) \quad \text{cum}(V_{T,r}, V_{T,s}, V_{T,t}) = \text{cum}^*(V_{T,r}^*, V_{T,s}^*, V_{T,t}^*) + o(T^{-1/2}) \quad \text{a.s.}$$

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 equations 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 Theorems 4 and 5 below. Several lemmas provide the technical details.

**THEOREM 2.** *Suppose Assumptions 1–9 hold. Then the following approximation holds uniformly over convex measurable  $C \subseteq \mathbb{R}^d$ :*

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

**PROOF.** From Janas [(1993) Theorem 2.3], we obtain the following Edgeworth expansion for the statistic  $A(\tilde{\phi}, I_T)$ , where  $\tilde{\phi}' = (\phi', 1)$  and  $\phi$  is as defined in Assumption 4:

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

(cf. Götze and Hipp [(1983), page 217] for the definition of  $\Psi_{T,s}$ ). Lemma 1 (below) yields  $\mathbb{E}A(\phi, I_T) = A(\phi, f) + o(T^{-1})$  and we can therefore replace  $\mathbb{E}A(\phi, I_T)$  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

$$H(x_1, \dots, x_p, y) = \left( \frac{x_1}{y}, \dots, \frac{x_p}{y} \right).$$

We have to check that this function is sufficiently smooth in a neighborhood of  $\mu = j(\phi', 1)f$  and that  $(\text{grad } H)(\mu)$  has full rank  $p$ . The first statement follows from the positivity of  $F(\pi) = jf$ , the variance of the underlying process  $\{X_t\}$ ; the second is trivial.  $\square$

**THEOREM 3.** *Suppose Assumptions 1–9 hold. Then, for almost all samples  $\{I_j\}$  and uniformly over convex measurable  $C \subseteq \mathbb{R}^d$*

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

**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  $B(\phi, J_T^*)$ . Note that  $B(\phi, I_T^*)$  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), page 71]) and Cramér's condition has to be modified in an obvious way.  $\square$

We now construct stochastic approximations for  $V_{T,r}$  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

$$(3.5) \quad \Delta^{(r)} \equiv \left( \phi^{(r)} - \int \phi^{(r)} g \right) g,$$

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

$$\begin{aligned} V_{T,r} &= T^{1/2} \left( \frac{\int \phi^{(r)} I_T}{\int I_T} - \frac{\int \phi^{(r)} f}{\int f} \right) \\ &= T^{1/2} \frac{\int (\phi^{(r)} - \phi^{(r)} g) g(I_T/f)}{\int g(I_T/f)} \\ &= T^{1/2} \frac{\int \Delta^{(r)}(I_T/f)}{\int g(I_T/f)} \\ &= T^{1/2} \int \Delta^{(r)}(I_T/f) (2 - \int g(I_T/f)) + o_p(T^{-1/2}). \end{aligned}$$

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

$$(3.6) \quad \int \Delta^{(r)} = 0,$$

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

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

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

LEMMA 1. *Suppose  $\psi_j$  are bounded functions and Assumptions 1, 5 and 6 hold. Then we have*

- (i)  $\text{cum} \left( \int \psi_1 I_T \right) = \int \psi_1 f + o(T^{-1}),$
- (ii)  $\text{cum} \left( \int \psi_1 I_T, \int \psi_2 I_T \right) = O(T^{-1}),$
- (iii)  $\text{cum} \left( \int \phi_1 I_T, \dots, \int \psi_l I_T \right) = o(T^{-l/2}) \quad l \geq 3.$

PROOF. We give only a sketch. Assumption 1 implies  $\sum_u |u|^2 |c(u)| < \infty$  and therefore also

$$|f(\alpha + \beta) - f(\alpha) - \beta f'(\alpha)| \leq K \beta^2,$$

with some constant  $K$ . Then we obtain

$$\begin{aligned} &\left| \text{cum} \left( \int \psi_1 I_T \right) - \int \psi_1 f \right| \\ &= \left| \frac{1}{2} \int_{-\pi}^{\pi} \psi_1(\alpha) \int_{-\pi}^{\pi} \{ f(\alpha + \beta) - f(\alpha) \} \frac{|H_{1,T}(\beta)|^2}{2\pi H_{2,T}} d\beta d\alpha \right| \\ &\leq K \int_{-\pi}^{\pi} \beta^2 \frac{|H_{1,T}(\beta)|^2}{H_{2,T}} d\beta \leq O(T^{-2+4\delta}) = o(T^{-1}) \end{aligned}$$

by using Lemma 5.4 of Dahlhaus (1988) [note that the taper of Assumption 6 is of degree  $(1, 2\delta)$  in the terminology of that paper]. The proof of (ii) and (iii) is standard (cf. Dahlhaus [(1983), Lemmas 6 and 7]).  $\square$

The application of this lemma leads to the following expressions:

$$(3.7) \quad \text{cum}(W_{T,r}) = -T^{+1/2} \text{cum}\left(\int \Delta^{(r)}(I_T/f), \int g(I_T/f)\right) + o(T^{-1/2}),$$

$$(3.8) \quad \begin{aligned} & \text{cum}(W_{T,r}, W_{T,s}) \\ &= T \text{cum}\left(\int \Delta^{(r)}(I_T/f), \int \Delta^{(s)}(I_T/f)\right) + o(T^{1/2}), \end{aligned}$$

and

$$(3.9) \quad \begin{aligned} & \text{cum}(W_{T,r}, W_{T,s}, W_{T,t}) \\ &= T^{3/2} \text{cum}\left(\int \Delta^{(r)}(I_T/f), \int \Delta^{(s)}(I_T/f), \int \Delta^{(t)}(I_T/f)\right) \\ &\quad - 2T^{3/2} \text{cum}\left(\int \Delta^{(r)}(I_T/f), \int g(I_T/f)\right) \\ &\quad \quad \times \text{cum}\left(\int \Delta^{(s)}(I_T/f), \int \Delta^{(t)}(I_T/f)\right) \\ &\quad - 2T^{3/2} \text{cum}\left(\int \Delta^{(s)}(I_T/f), \int g(I_T/f)\right) \\ &\quad \quad \times \text{cum}\left(\int \Delta^{(r)}(I_T/f), \int \Delta^{(t)}(I_T/f)\right) \\ &\quad - 2T^{3/2} \text{cum}\left(\int \Delta^{(t)}(I_T/f), \int g(I_T/f)\right) \\ &\quad \quad \times \text{cum}\left(\int \Delta^{(r)}(I_T/f), \int \Delta^{(s)}(I_T/f)\right) + o(T^{-1/2}). \end{aligned}$$

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 Assumptions 1 and 4–6, we have*

$$(i) \quad \begin{aligned} & T \text{cum}\left(\int \phi^{(r)}(I_T/f), \int \phi^{(s)}(I_T/f)\right) \\ &= 2\pi \int \phi^{(r)} \phi^{(s)} + (\kappa_4/\sigma^4) \int \phi^{(r)} \int \phi^{(s)} + o(1); \end{aligned}$$

$$\begin{aligned}
 \text{(ii)} \quad T^2 \text{cum} & \left( \int \phi^{(r)}(I_T/f), \int \phi^{(s)}(I_T/f), \int \phi^{(t)}(I_T/f) \right) \\
 & = 8\pi^2 \int \phi^{(r)}\phi^{(s)}\phi^{(t)} + (\kappa_3^2/\sigma^6)4\pi \\
 & \quad \times \left[ \phi^{(r)}(0) \int \phi^{(s)} \int \phi^{(t)} + \phi^{(s)}(0) \int \phi^{(r)} \int \phi^{(t)} + \phi^{(t)}(0) \int \phi^{(r)} \int \phi^{(s)} \right. \\
 & \quad \left. + \int_0^\pi \int_0^\pi \phi^{(r)}(\alpha_1) \phi^{(s)}(\alpha_2) \{ \phi^{(t)}(\alpha_1 + \alpha_2) + \phi^{(t)}(\alpha_1 - \alpha_2) \right. \\
 & \quad \left. + \phi^{(t)}(-\alpha_1 + \alpha_2) + \phi^{(t)}(-\alpha_1 - \alpha_2) \} d\alpha_1 d\alpha_2 \right] \\
 & \quad + (\kappa_4/\sigma^4)4\pi \left( \int \phi^{(r)}\phi^{(s)} \int \phi^{(t)} + \int \phi^{(r)}\phi^{(t)} \int \phi^{(s)} + \int \phi^{(s)}\phi^{(t)} \int \phi^{(r)} \right) \\
 & \quad + (\kappa_6/\sigma^6) \int \phi^{(r)} \int \phi^{(s)} \int \phi^{(t)} + o(1),
 \end{aligned}$$

where  $\kappa_k$  is the  $k$ th cumulant of the innovations  $\xi_t$ .

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 (1988). We omit the details.  $\square$

For the calculation of the cumulants of  $W_{T,r}$ , we need only three kinds of cumulants of the statistics  $\int \Delta(I_T/f)$  and  $\int g(I_T/f)$ . Since  $\int \Delta^{(r)} = 0$  we obtain for these cumulants from Lemma 2 under Assumption 2 ( $\kappa_3 = 0$ ) the following expressions:

$$(3.10) \quad T \text{cum} \left( \int \Delta^{(r)}(I_T/f), \int g(I_T/f) \right) = 2\pi \int \Delta^{(r)}g + o(1),$$

$$(3.11) \quad T \text{cum} \left( \int \Delta^{(r)}(I_T/f), \int \Delta^{(s)}(I_T/f) \right) = 2\pi \int \Delta^{(r)}\Delta^{(s)} + o(1),$$

$$\begin{aligned}
 (3.12) \quad T^2 \text{cum} & \left( \int \Delta^{(r)}(I_T/f), \int \Delta^{(s)}(I_T/f), \int \Delta^{(t)}(I_T/f) \right) \\
 & = 8\pi^2 \int \Delta^{(r)}\Delta^{(s)}\Delta^{(t)} + o(1).
 \end{aligned}$$

Due to the central equation (3.6) (which holds since we are dealing with ratio statistics), the above cumulants are independent of the cumulants  $\kappa_4$  and  $\kappa_6$  of the innovations (while  $\kappa_3 = 0$  has to be assumed). In particular, all contributions from the dependence structure of the periodogram ordinates disappear. For the above cumulants it makes no difference if we replace the dependent  $\{I_j/f_j\}$  by i.i.d. r.v.'s  $\{E_j\}$  where  $E_j$  is exponentially distributed. Later we will see that the bootstrap counterparts of  $\{I_j/f_j\}$  behave similarly as independent and exponentially distributed variables, and that the corre-

sponding bootstrap cumulants have the same limits as above. We now summarize our results on the cumulants of  $W_{T,r}$ .

THEOREM 4. *Under Assumptions 1, 2, 4 and 6, we have*

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

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

and

$$\begin{aligned} &\text{cum}(W_{T,r}, W_{T,s}, W_{T,t}) \\ &= T^{-1/2}8\pi^2 \left( \int \Delta^{(r)}\Delta^{(s)}\Delta^{(t)} - \int \Delta^{(r)}g \int \Delta^{(s)}\Delta^{(t)} \right. \\ &\quad \left. - \int \Delta^{(s)}g \int \Delta^{(r)}\Delta^{(t)} - \int \Delta^{(t)}g \int \Delta^{(r)}\Delta^{(s)} \right) + o(T^{-1/2}), \end{aligned}$$

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. *Suppose Assumptions 1, 4 and 6–8 hold. Then*

$$(i) \quad \sup_{x \in \mathbb{R}} |G_n(x) - G(x)| \rightarrow 0 \quad a.s.$$

and

$$(ii) \quad \int_0^\infty g(x) dG_n(x) \rightarrow \int_0^\infty g(x) dG(x) \quad a.s.$$

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

$$\sup_{x \in \mathbb{R}} \frac{|g(x)|}{1 + |x|^8} < \infty.$$

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 (1995)  $\square$

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

COROLLARY 1. Under Assumptions 1–3, 5 and 7–9, we have

$$(i) \quad \frac{1}{n} \sum_{j=1}^n \hat{\varepsilon}_j^p \rightarrow \mathbb{E} \chi_1^p \quad a.s.$$

and

$$\frac{1}{n} \sum_{j=1}^n \tilde{\varepsilon}_j^p \rightarrow \mathbb{E} \chi_1^p \quad a.s.$$

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

$$(ii) \quad \hat{F}_n \Rightarrow G \quad a.s. \quad \text{and} \quad \tilde{F}_n \Rightarrow G \quad a.s.$$

The proof is given in the Appendix.

The corollary says that the bootstrap distribution  $\tilde{F}_n$  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 Assumptions 1–5 and 7–9, we have, for all  $p \leq 8$ ,

$$\begin{aligned} T^{p-1} \text{cum}^* \left( \frac{\pi}{n} \sum_{j=1}^n \phi_j^{(r_1)} I_j^*, \dots, \frac{\pi}{n} \sum_{j=1}^n \phi_j^{(r_p)} I_j^* \right) \\ \rightarrow (p-1)! (2\pi)^{p-1} \int \prod_{j=1}^p \phi^{(r_j)} f^p \quad a.s., \end{aligned}$$

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

PROOF. The result follows by using straightforward calculations.  $\square$

The lemma shows that the cumulants of order  $p \geq 2$  of the bootstrapped statistic  $B(\phi, I_T^*)$  do not converge to the same limit as the cumulants of  $A(\phi, I_T)$ . We already know the reason why the bootstrap approximation fails in this situation: independent resampling does not take care of the dependence structure among the random variables in 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  $W_{T,r}$ , that is,

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

where

$$\hat{\Delta}_j^{(r)} \equiv \left( \phi_j^{(r)} - \frac{\pi}{n} \sum_{k=1}^n \phi_k^{(r)} \hat{g}_k \right) \hat{g}_j, \quad \hat{g}_j \equiv \hat{f}_j / \left( \frac{\pi}{n} \sum_{k=1}^n \hat{f}_k \right).$$

Replacing the cumulants in (3.10)–(3.12) 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 [note that  $\mathbb{E}^*(\pi/n)\Sigma \hat{\Delta}_j^{(r)}(I_j^{(r)}/\hat{f}_j) = 0$ ].

**THEOREM 5.** *Under Assumptions 1–5 and 7–9, we have*

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

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

and

$$\begin{aligned} \text{cum}^*(W_{T,r}^*, W_{T,s}^*, W_{T,t}^*) \\ = T^{-1/2}8\pi^2 \left( \int \Delta^{(r)}\Delta^{(s)}\Delta^{(t)} - \int \Delta^{(r)}g \int \Delta^{(s)}\Delta^{(t)} \right. \\ \left. - \int \Delta^{(s)}g \int \Delta^{(r)}\Delta^{(t)} - \int \Delta^{(t)}g \int \Delta^{(r)}\Delta^{(s)} \right) + o(T^{-1/2}) \quad a.s. \end{aligned}$$

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 density. A detailed discussion may be found in Dzhaparidze and Yaglom (1983).

Suppose  $\{X_t\}_{t \in \mathbb{Z}}$  is a linear process with spectral density  $f$  that fulfills Assumptions 1, 2 and 7 and we fit a parametric model  $\mathcal{F} = \{f_\theta: \theta \in \Theta\}$  to the data. Suppose  $\theta = (\sigma^2, \tau)$ ,  $f_\theta(\alpha) = \sigma^2 h_\tau(\alpha)$  and Kolmogorov’s formula holds, that is,

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

This is, for example, true for ARMA models where  $\sigma^2$  is the innovation variance and  $\tau$  contains the ARMA parameters [c.f. Brockwell and Davies (1987), Section 5.8]. We do not assume that  $f \in \mathcal{F}$ , that is we allow the model to be misspecified.

The Whittle estimate  $\hat{\theta}_T = (\hat{\sigma}_T^2, \hat{\tau}_T)$  is determined by minimizing the Whittle function  $\mathcal{L}_T(\theta)$  of Example 3; that is it fulfills the equations

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

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  $\theta_0 = (\sigma_0^2, \tau_0)$  which minimizes the corresponding theoretical function

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

that is,  $\theta_0$  is determined by the equations

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

and

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

The bootstrap version of  $\mathcal{L}_T(\theta)$  is

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

and the bootstrap Whittle estimate  $\theta^* = (\sigma^{*2}, \tau^*)$  is determined by minimizing  $\mathcal{L}_T^*(\theta)$  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

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

It is heuristically obvious that  $\theta^* - \bar{\theta}$  will converge to 0, where  $\bar{\theta} = (\bar{\sigma}^2, \bar{\tau})$  is obtained by minimizing

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

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

The heuristics in Section 2 indicate that the bootstrap is valid for the parameter  $\tau$  (since  $\int f \nabla_\tau f_{\theta_0}^{-1} = 0$ ). This will be proved below. However, the bootstrap does not work for the parameter  $\sigma^2$  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  $\sigma^2$ , we note that  $\hat{\tau}_T(\tau_0, \tau^*, \bar{\tau})$  are also the minima of  $L_T(\tau) \equiv \int I_T h_\tau^{-1} [L(\tau) \equiv \int f h_\tau^{-1}, L_*(\tau) \equiv (\pi/n) \sum_j I_j^* h_\tau(2\pi j/T)^{-1}, \bar{L}(\tau) \equiv (\pi/n) \sum_j \hat{f}_j h_\tau(2\pi j/T)^{-1}]$ , respectively]. We need the following assumption in addition to Assumptions 1-9.

ASSUMPTION 10. The set of parameters  $\mathcal{T} \subset \mathbb{R}$  is compact. The parameters are identifiable; that is,  $\tau_1 \neq \tau_2$  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  $\alpha \in [0, \pi]$ . Here  $h_\tau(\alpha)$  and its derivatives are uniformly bounded, that is,  $\exists 0 < \underline{c} \leq \bar{c} < \infty \forall \tau \in \mathcal{T}, \alpha \in [0, \pi]$ ,

$$\underline{c} \leq h_\tau(\alpha) \leq \bar{c}, \quad \left| \left( \frac{\partial}{\partial \tau} \right)^i h_\tau^{-1}(\alpha) \right| \leq \bar{c}, \quad i = 1, \dots, 4,$$

and

$$\left| \left( \frac{\partial}{\partial \alpha} \right)^j h_\tau(\alpha) \right| \leq \bar{c}, \quad j = 1, 2.$$

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

Furthermore, in Assumption 9 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  $J(\tau) \equiv 2\pi \int (\phi_\tau^{(1)} f)^2$ ,  $J_*(\tau) \equiv T \mathbb{E}^*(L_*^{(1)}(\tau))^2$  and  $K_*(\tau) \equiv \mathbb{E}^* L_*^{(2)}(\tau)$ .

THEOREM 6. *Suppose Assumptions 1–10 hold. Then, for almost all samples  $\{I_j\}$ ,*

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

The proof is deferred to the Appendix.

REMARK 8. Without proof we remark that the bootstrap also works for  $\sigma^2$  if  $\kappa_4 = 0$ . In the case of an AR( $p$ ) model,  $\theta^*$  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 simulation 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 candidate 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 neighborhood of peaks in the spectrum. We are convinced that this trade-off is the reason for our bad results.

The behavior of our estimates became much better when smoothing the log periodogram. Although this method is less efficient than smoothing the periodogram itself it is obvious that the bias is strongly improved in the neighborhood of peaks. Furthermore, the log transformation is asymptotically variance stabilizing for the periodogram. In the simulations below we used a kernel estimate with Epanečnikov kernel. A bias correction is obtained from the following heuristic consideration. Suppose  $Z_j$  are i.i.d. exponentially distributed random variables with (constant) mean  $f(\lambda)$  (this is the asymptotic distribution of the periodogram ordinates in a local neighborhood) and  $w_j$  are the kernel weights. Then

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

where  $\Gamma$  is the Gamma function. We therefore estimate  $f_k$  by

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

where

$$w_j = \frac{1}{b} K\left(\frac{1}{b} \frac{2\pi j}{T}\right)$$

and

$$K(x) = \frac{3}{4} \pi \left(1 - \left(\frac{x}{\pi}\right)^2\right) \quad \{|x| \leq \pi\}.$$

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 hold only 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

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

as the bootstrap estimate for the distribution of

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

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  $\varepsilon_t$  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.

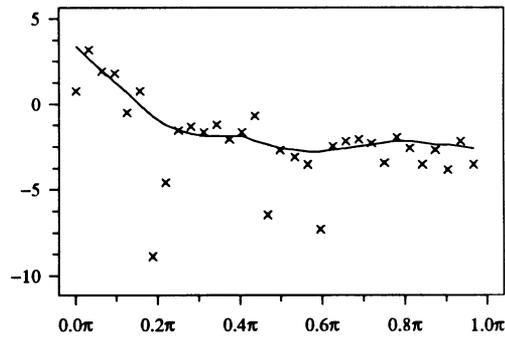


FIG. 1. Log periodogram and kernel estimate for an AR(1) process.

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

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

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 3. 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. In this example  $T = 64$  observations of an ARMA(4, 2) process with AR roots  $0.9^{-1}e^{i0.2\pi}$ ,  $0.9^{-1}e^{-i0.2\pi}$ ,  $0.9^{-1}e^{i0.5\pi}$ ,  $0.9^{-1}e^{-i0.5\pi}$ , MA roots  $0.8^{-1}e^{i0.35\pi}$ ,  $0.8^{-1}e^{-i0.35\pi}$  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

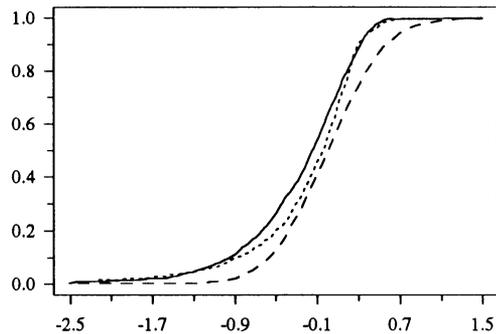


FIG. 2. True (solid), asymptotic (dashed) and bootstrap (dotted) distribution of the first-order correlation for an AR(1) process.

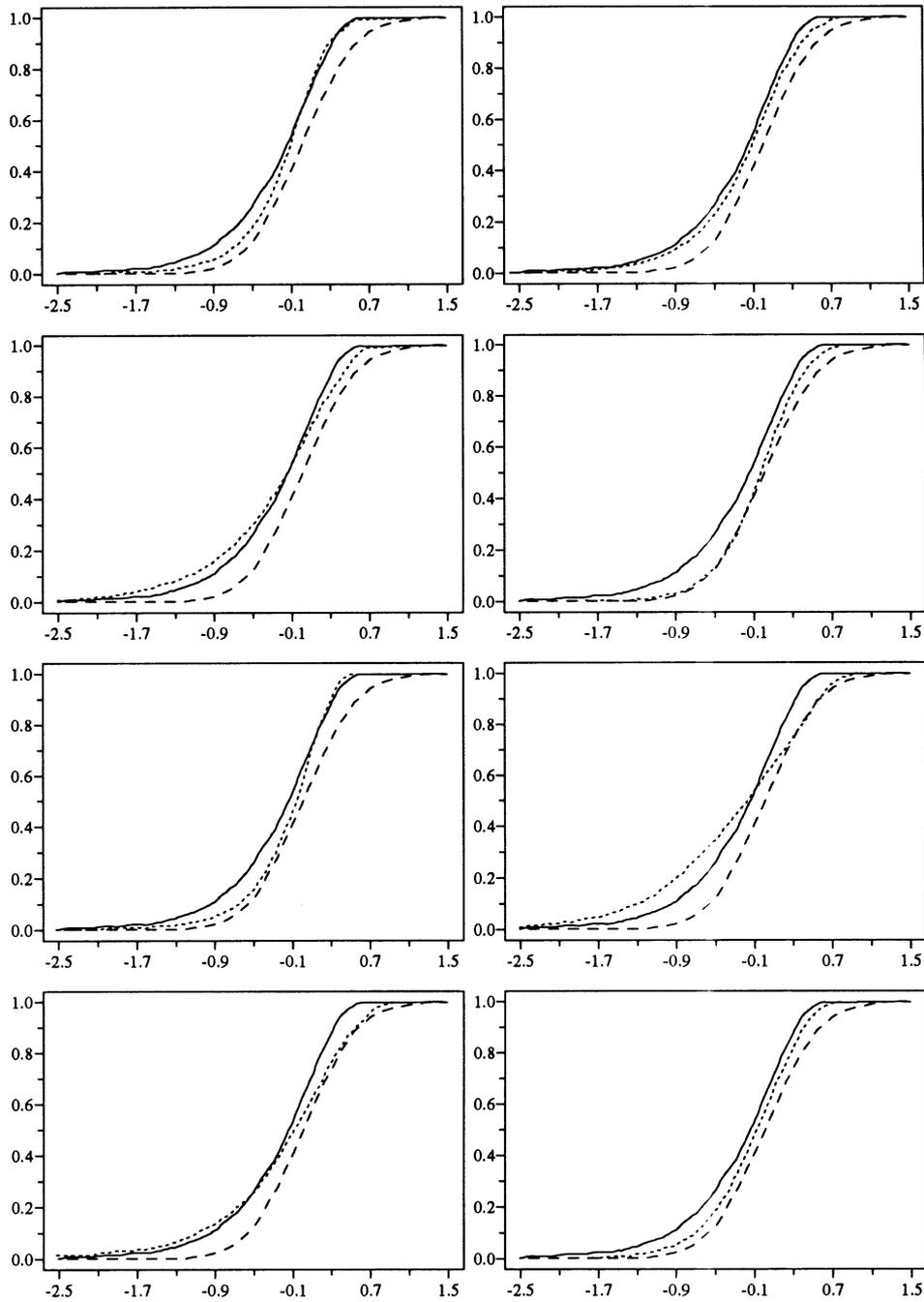


FIG. 3. True (solid), asymptotic (dashed) and bootstrap (dotted) distribution of the first-order correlation for eight different realizations of an AR(1) process.

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

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

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

Suppose we want to construct a confidence set for  $a_0$ . In the correct specified case the above result implies for the Mahalanobis distance

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

Replacing  $(1/\sigma^2)\Sigma$  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. Furthermore, we may estimate

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

and

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

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

Figure 4 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 5 shows the tapered periodogram with the kernel estimate as discussed above where  $b = 0.05$ . In Figure 6 the asymptotic  $\chi_4^2$ -distribution of the above distance is shown as the dashed line. The solid

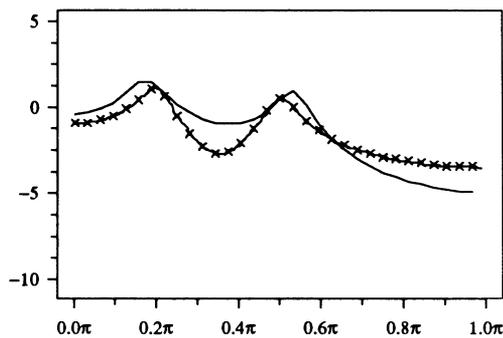


FIG. 4. Log spectrum of an ARMA(4, 2) process (crosses) and an AR(4) fit (solid line).

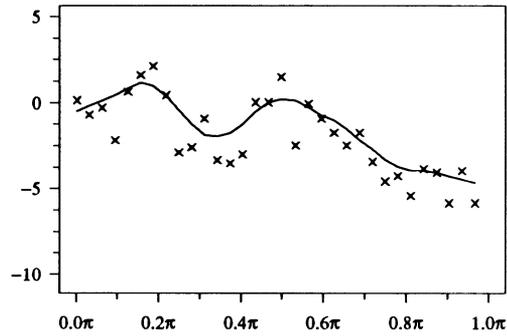


FIG. 5. Log periodogram and kernel estimate for an ARMA(4, 2) process.

line is the true distribution of the statistic with  $(1/\sigma^2)\Sigma$  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  $AR(\infty)$  bootstrap or to a block bootstrap.

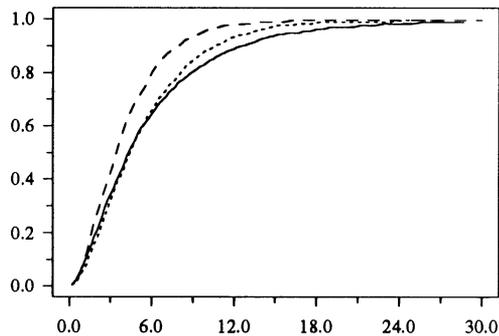


FIG. 6. True (solid), asymptotic (dashed) and bootstrap (dotted) distribution of the Mahalanobis distance for an AR(4) model.

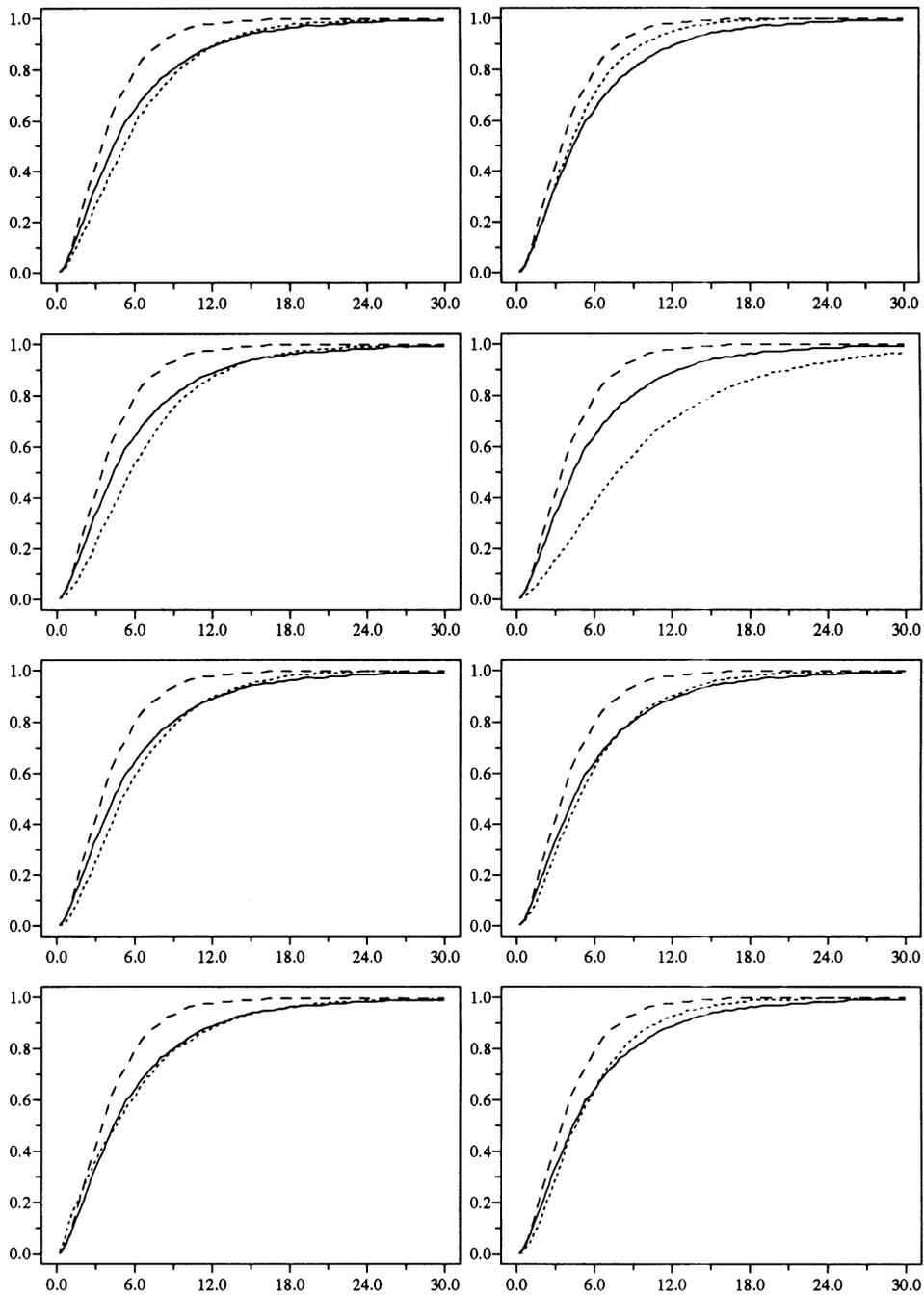


FIG. 7. True (solid), asymptotic (dashed) and bootstrap (dotted) distribution of the Mahalanobis distance of an AR(4) model for eight different realizations of an ARMA(4, 2) process.

## 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:

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

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

We show that

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

By the triangle inequality we have

$$(A.3) \quad d_p(\chi_1, \varepsilon_1^*) \leq d_p(\chi_1, \varepsilon_1^0) + d_p(\varepsilon_1^0, \hat{\varepsilon}_1) + d_p(\hat{\varepsilon}_1, \varepsilon_1^*),$$

where the df of  $\varepsilon_1^0(\hat{\varepsilon}_1)$  is the edf  $G_n(\hat{F}_n)$  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 (A.3) converge to 0 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  $n^{-1}$ ,  $j = 1, \dots, n$ . Then

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

By Lemma 3(ii) we obtain

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

Therefore, the second term in (A.3) converges to 0 by the convergence of the estimate  $\hat{f}$  to  $f$  (Assumption 2).

Using exactly the same argument as above, we get also

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

by Lemma 3(ii) and Assumption 2.  $\square$

**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 occurring 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 mention only the essential steps.

We set down

$$(A.4) \quad V_* \equiv \sqrt{T}(\tau^* - \bar{\tau}),$$

$$(A.5) \quad Z_*^{(i)}(\tau) \equiv \sqrt{T}(L_*^{(i)}(\tau) - \mathbb{E}L_*^{(i)}(\tau)),$$

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

and

$$(A.7) \quad \begin{aligned} U_*(\tau) &\equiv -\frac{Z_*^{(1)}(\tau)}{K_*(\tau)} + \frac{1}{\sqrt{T}} \frac{Z_*^{(1)}(\tau) Z_*^{(2)}(\tau)}{K_*(\tau)} \\ &\quad - \frac{1}{2\sqrt{T}} \frac{Z_*^{(3)}(\tau)}{K_*(\tau)} \left( \frac{Z_*^{(1)}(\tau)}{K_*(\tau)} \right)^2. \end{aligned}$$

We will show that the following stochastic expansion holds:

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

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

By a lemma of Chibisov (cf. Janas [(1993), Lemma 4.5]), the Edgeworth expansions for  $V_*$  and  $U_*(\bar{\tau})$  match up to order  $T^{-1/2}$ . However, 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 (A.8) we consider the following Taylor expansion. Since  $L_*^{(1)}(\tau^*) = 0$ , we have

$$(A.9) \quad \begin{aligned} 0 &= \sqrt{T}L_*^{(1)}(\bar{\tau}) + \frac{1}{\sqrt{T}}Z_*^{(2)}(\bar{\tau})V_* + K_*(\bar{\tau})V_* \\ &+ \frac{1}{2\sqrt{T}}L_*^{(3)}(\bar{\tau})V_*^2 + \frac{1}{6T}L_*^{(4)}(\bar{\tau})V_*^3, \end{aligned}$$

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

$$(A.10) \quad \begin{aligned} V_* &= -\frac{\sqrt{T}L_*^{(1)}(\bar{\tau})}{K_*(\bar{\tau})} - \frac{Z_*^{(2)}(\bar{\tau})}{K_*(\bar{\tau})\sqrt{T}}V_* \\ &- \frac{L_*^{(3)}(\bar{\tau})}{2K_*(\bar{\tau})\sqrt{T}}V_*^2 - \frac{L_*^{(4)}(\bar{\tau})}{6K_*(\bar{\tau})T}V_*^3. \end{aligned}$$

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

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

$$(A.11) \quad P^*(|\tau^* - \bar{\tau}| > d_1T^{\alpha-1/2}) = o(T^{-1/2}) \quad \text{a.s.}$$

$$(A.12) \quad P^*(|L_*^{(i)}(\bar{\tau}) - \mathbb{E}^*L_*^{(i)}(\bar{\tau})| > d_2T^{\alpha-1/2}) = o(T^{-1/2}) \quad \text{a.s. for } i = 1, 2, 3,$$

$$(A.13) \quad P^*\left(\sup_{\tau \in \mathcal{T}} |L_*^{(4)}(\tau)| > d_3T^\alpha\right) = o(T^{-1/2}) \quad \text{a.s.}$$

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

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

where  $P^*(|\tilde{\xi}_*| > d_4T^{2\alpha}) = o(T^{-1/2})$  a.s. for some  $d_4 > 0$ .

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

$$\begin{aligned} V_* &= -\frac{Z_*^{(1)}(\bar{\tau})}{K_*(\bar{\tau})} + \frac{1}{\sqrt{T}}\frac{Z_*^{(1)}(\bar{\tau})}{K_*(\bar{\tau})}\frac{Z_*^{(2)}(\bar{\tau})}{K_*(\bar{\tau})} \\ &- \frac{1}{2\sqrt{T}}\frac{L_*^{(3)}(\bar{\tau})}{K_*(\bar{\tau})}\left(\frac{Z_*^{(1)}(\bar{\tau})}{K_*(\bar{\tau})}\right)^2 + \frac{1}{T}\tilde{\xi}_*, \end{aligned}$$

where  $P^*(|\tilde{\xi}_*| > d_5T^{3\alpha}) = o(T^{-1/2})$  for some  $d_5 > 0$ .  $\square$

**Acknowledgment.** 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. *Sankhya Ser. A* **46** 219–232.
- BHATTACHARYA, R. N. and GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 434–451.
- BHATTACHARYA, R. N. and RANGA RAO, R. (1976). *Normal Approximation and Asymptotic Expansions*. Wiley, New York.
- BICKEL, P. J. and FRIEDMAN, 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, New York.
- 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. Multivariate 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 *Developments in Statistics* (P. R. Krishnaiah, ed.) **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.
- FRIEDMAN, D. A. (1984). On bootstrapping two-stage least-squares estimates in stationary linear models. *Ann. Statist.* **12** 827–842.
- GÖTZE, F. and HIPF, C. (1983). Asymptotic expansions for sums of weakly dependent random vectors. *Z. Wahrsch. Verw. Gebiete* **64** 211–239.
- HURVICH, C. M. and ZEGER, S. (1987). Frequency domain bootstrap methods for time series. Statistics and Operations Research Working Paper, New York Univ.
- 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*. In *Proceedings of the 26th Symposium on the Interface* (J. Sall and A. Lehman, eds.) 423–425. Interface Foundation of North America, Fairfax Station, VA.
- JANAS, D. and VON SACHS, R. (1995). Consistency for non-linear functions of the periodogram of tapered data. *J. Time Ser. Anal.* **16** 585–606.
- 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 and L. Billard, eds.) 225–248. Wiley, New York.
- NORDGAARD, A. (1992). Resampling stochastic processes using a bootstrap approach. In *Bootstrapping and Related Techniques. Lecture Notes in Econom. Math. Sys.* **376**. Springer, 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.

INSTITUT FÜR ANGEWANDTE MATHEMATIK  
UNIVERSITÄT HEIDELBERG  
IM NEUENHEIMER FELD 294  
D-69120 HEIDELBERG  
GERMANY