*The Annals of Statistics*
2003, Vol. 31, No. 6, 1923–1955


# AUTOREGRESSIVE-AIDED PERIODOGRAM BOOTSTRAP FOR TIME SERIES[1]

BY JENS-PETER KREISS AND EFSTATHIOS PAPARODITIS

*Technische Universität Braunschweig and University of Cyprus*

A bootstrap methodology for the periodogram of a stationary process is proposed which is based on a combination of a time domain parametric and a frequency domain nonparametric bootstrap. The parametric fit is used to generate periodogram ordinates that imitate the essential features of the data and the weak dependence structure of the periodogram while a nonparametric (kernel-based) correction is applied in order to catch features not represented by the parametric fit. The asymptotic theory developed shows validity of the proposed bootstrap procedure for a large class of periodogram statistics. For important classes of stochastic processes, validity of the new procedure is also established for periodogram statistics not captured by existing frequency domain bootstrap methods based on independent periodogram replicates.

**1. Introduction.** Consider a strictly stationary univariate process $\mathbf{X} = (X_t : t \in \mathbb{Z} = \{0, \pm 1, \pm 2, \dots\})$ and assume that $X_t$ has the representation

$$X_t = \sigma \sum_{\nu=-\infty}^{\infty} \alpha_\nu \varepsilon_{t-\nu}, \qquad t \in \mathbb{Z}, \tag{1.1}$$

where $\{\alpha_\nu\}$, $\alpha_0 = 1$, is an absolutely summable sequence, $\{\varepsilon_t\}$ is a sequence of independent, identically distributed random variables with mean 0 and unit variance and $\sigma$ is a positive constant. Assume that we have observations $X_1, X_2, \dots, X_n$ of the process $\mathbf{X}$ at hand. Statistical inference in the frequency domain is commonly based on the so-called periodogram $I_n(\lambda)$,

$$I_n(\lambda) = \frac{1}{2\pi n} \left| \sum_{t=1}^{n} X_t e^{-i\lambda t} \right|^2, \qquad \lambda \in [0, \pi], \tag{1.2}$$

which is known to be an asymptotically unbiased but not consistent estimator of the spectral density $f$ of the process $\mathbf{X}$. Because of (1.1) and assumptions (A1) and (A2) below, $f$ has the representation

$$f(\lambda) = \frac{\sigma^2}{2\pi} \left| \sum_{\nu=-\infty}^{\infty} \alpha_\nu e^{-i\nu\lambda} \right|^2, \qquad \lambda \in [0, \pi]. \tag{1.3}$$

---

Received January 2001; revised February 2003.

[1]Supported by Sonderforschungsbereich 373 at Humboldt-Universität zu Berlin.

*AMS 2000 subject classifications.* Primary 62G09; secondary 62M10.

*Key words and phrases.* Bootstrap, periodogram, nonparametric estimators, ratio statistics, spectral means.

Later we will require in some cases causality and invertibility of $\{X_t\}$ which narrows the class (1.1) a little bit. This ensures that the process $\{X_t\}$ can be represented as a one-sided infinite-order autoregressive process

$$X_t = \sum_{\nu=1}^{\infty} a_\nu X_{t-\nu} + \sigma \cdot \varepsilon_t, \qquad t \in \mathbb{Z}, \tag{1.4}$$

where $\{a_\nu\}$ is an absolute summable sequence.

Methods of bootstrapping the periodogram $I_n(\lambda)$ have attracted considerable attention in recent years. Compared to time domain bootstrap methods, the appeal of frequency domain methods lies in the fact that for a huge class of stochastic processes the observed series $X_1, X_2, \ldots, X_n$ can be transformed into a set of $N = [n/2]$ nearly independent statistics, the periodogram ordinates at the so-called Fourier frequencies $\lambda_j = 2\pi j/n,\ j = 0, 1, 2, \ldots, N$. Since for $\lambda_j \in (0, \pi)$ the mean and the variance of $I_n(\lambda_j)$ are approximately equal to $f(\lambda_j)$ and $f^2(\lambda_j)$, respectively, bootstrap methods designed for a nonparametric regression setup with independent errors can be potentially applied to bootstrap the periodogram.

For Gaussian processes, frequency domain bootstrap methods have been considered, among others, by Nordgaard (1992) and Theiler, Paul and Rubin (1994). Using the property that the relation between periodogram and spectral density can be approximately described by means of a multiplicative regression model, Hurvich and Zeger (1987) for Gaussian processes and Franke and Härdle (1992) in a more general context, proposed a nonparametric residual-based bootstrap method. The method uses an initial (nonparametric) estimate of the spectral density $f$ and i.i.d. resampling of (appropriately defined) frequency domain residuals. Franke and Härdle established asymptotic validity of this method for nonparametric (kernel) estimators of the spectral density while Dahlhaus and Janas (1996) extended the validity of this bootstrap procedure to the class of the so-called ratio statistics and to Whittle estimators. An alternative idea to bootstrap the periodogram has been proposed by Paparoditis and Politis (1999). Their method uses smoothness properties of the spectral density $f$ and the periodogram replicates are obtained by locally resampling from adjacent periodogram ordinates. An overview of the different methods to bootstrap time series in the frequency domain is given by Paparoditis (2002).

The independence of the bootstrap periodogram ordinates is an essential feature of the bootstrap procedures mentioned above that restricts the classes of statistics to which the existing methods can be successfully applied. Loosely speaking, validity of the above nonparametric bootstrap procedures can be established only for periodogram statistics for which the weak and asymptotically vanishing dependence of the periodogram ordinates does not affect their large-sample distribution. Nonparametric estimators and ratio statistics have this property; compare Franke and Härdle (1992) and Dahlhaus and Janas (1996). However, there are other interesting classes of periodogram statistics for which the dependencies of the periodogram ordinates sum up to a nonvanishing contribution. For instance, Dahlhaus (1985) investigated the following class of integrated periodogram estimators:

$$M(I_n, \varphi) = \int_0^\pi \varphi(\lambda) I_n(\lambda)\, d\lambda \tag{1.5}$$

for some appropriately defined functions $\varphi$ on $[0, \pi]$. The following are some special cases covered by (1.5).

EXAMPLES. (i) $\varphi(\lambda) = 2\cos(\lambda h)$, $h \in \mathbb{N}_0$, leads to the empirical autocovariance $\hat{\gamma}_n(h)$, since

$$\int_0^\pi 2\cos(\lambda h) I_n(\lambda)\, d\lambda = \frac{1}{n} \sum_{t=1}^{n-h} X_t X_{t+h} =: \hat{\gamma}_n(h),$$

which is an $\sqrt{n}$-consistent estimate of the true underlying autocovariance $\gamma(h) = E X_t X_{t+h}$.

(ii) $\varphi(\lambda) = \mathbb{1}_{[0,x]}(\lambda)$, $x \in [0, \pi]$, leads to the integrated periodogram

$$F_n(x) = \int_0^x I_n(\lambda)\, d\lambda,$$

which consistently estimates the spectral distribution function $F(x) = \int_0^x f(\lambda)\, d\lambda$.

Again from Dahlhaus (1985), it is known that under suitable assumptions the asymptotic distribution of

$$\sqrt{n}\big(M(I_n, \varphi) - M(f, \varphi)\big) = \sqrt{n}\left(\int_0^\pi \varphi(\lambda) I_n(\lambda)\, d\lambda - \int_0^\pi \varphi(\lambda) f(\lambda)\, d\lambda\right) \tag{1.6}$$

is Gaussian with mean 0 and variance given by

$$2\pi \int_0^\pi \varphi^2(\lambda) f^2(\lambda)\, d\lambda + \kappa_4 \left(\int_0^\pi \varphi(\lambda) f(\lambda)\, d\lambda\right)^2, \tag{1.7}$$

where $\kappa_4$ is the fourth cumulant of $\varepsilon_t$. Note that instead of (1.5) a discretized version may also be used. Discretization of the integral is usually done along the Fourier frequencies $\lambda_j = 2\pi j/n$, $j = 0, 1, \ldots, N$. This leads to the following discrete version of (1.6):

$$\frac{2\pi}{\sqrt{n}} \sum_{j=0}^{N} \varphi(\lambda_j)\{I_n(\lambda_j) - f(\lambda_j)\}. \tag{1.8}$$

Under some smoothness assumptions on $\varphi$, the difference between (1.6) and (1.8) is asymptotically negligible.

A modification of the nonparametric residual-based frequency domain bootstrap has been proposed by Janas and Dahlhaus (1994) in order to deal with the periodogram estimators (1.5). However, their method is based on a direct estimation of the fourth-order cumulant of the error process which requires nonparametric estimation of functionals of the spectral density of $\{X_t\}$ and of the squared process $\{X_t^2\}$.

In this paper, we introduce a new bootstrap procedure for the periodogram which is based on a combination of a parametric time domain and a nonparametric frequency domain bootstrap. The essential feature of the new bootstrap proposal is the following: for periodogram statistics for which the dependence of the periodogram ordinates does not affect their asymptotic distribution, the bootstrap procedure proposed "works" under the same set of process assumptions as those required for the aforementioned fully nonparametric methods. Furthermore, for stochastic processes possessing representation (1.4), our procedure also leads to asymptotically valid approximations for more general classes of periodogram statistics, including those given by (1.5).

It is mentioned above that in case we are interested in ratio statistics, that is,

$$R(I_n, \varphi) = \frac{\int_0^\pi \varphi(\lambda) I_n(\lambda)\, d\lambda}{\int_0^\pi I_n(\lambda)\, d\lambda}, \tag{1.9}$$

only, frequency domain bootstrap methods based on independent periodogram replicates work asymptotically. The reason is that the asymptotic distribution of a ratio statistic, which is again a normal distribution, does not depend on the fourth-order cumulant of $\sigma\varepsilon_1$. In this case, the asymptotic variance for ratio statistics is equal to

$$2\pi \int_0^\pi \psi^2 f^2\, d\lambda \Big/ \left(\int f\, d\lambda\right)^4, \tag{1.10}$$

where $\psi = \varphi \int f - \int \varphi f$ [cf. Dahlhaus and Janas (1996), page 1939]. Nevertheless, we think that the autoregressive-aided frequency domain bootstrap presented in this paper may also outperform the finite-sample behavior of the existing frequency domain bootstrap methods for ratio statistics because the dependence structure of the periodogram ordinates is mimicked to a certain extent in our bootstrap proposal. The frequency domain nonparametric residual-based bootstrap treats the periodogram ordinates as independent random variables, which they are only asymptotically. The same is also true for the local bootstrap of the periodogram proposed by Paparoditis and Politis (1999).

To describe the basic idea behind our procedure, recall first that under certain assumptions on the moment structure of the error process and the rate of decrease of the coefficients $\{\alpha_\nu\}$ in (1.1) [see assumptions (A1) and (A2) in Section 2], we have

$$E\big(I_n(\lambda_j)\big) = f(\lambda_j) + O(n^{-1}) \tag{1.11}$$

and

$$\operatorname{Cov}\big(I_n(\lambda_j), I_n(\lambda_k)\big) = \begin{cases} f^2(\lambda_j) + O(n^{-1}), & \text{for } j = k, \\ n^{-1} f(\lambda_j) f(\lambda_k)\Big(\dfrac{E\varepsilon_1^4}{\sigma^4} - 3\Big) + o(n^{-1}), & \text{for } j \neq k; \end{cases} \tag{1.12}$$

compare Brillinger (1981) and Janas and Dahlhaus (1994).

Consider next the autoregressive process $\widetilde{\mathbf{X}} = \{\widetilde{X}_t : t \in \mathbb{Z}\}$ defined by

$$\widetilde{X}_t = \sum_{\nu=1}^{p} a_\nu(p)\widetilde{X}_{t-\nu} + \sigma(p)\widetilde{\varepsilon}_t, \tag{1.13}$$

where $\mathbf{a}(p) = (a_1(p), a_2(p), \dots, a_p(p))' = \boldsymbol{\Gamma}(p)^{-1}\boldsymbol{\gamma}_p$, $\boldsymbol{\Gamma}(p) = (\gamma(i - j))_{i,j=1,2,\dots,p}$, $\boldsymbol{\gamma}_p = (\gamma(1), \gamma(2), \dots, \gamma(p))'$ and $\widetilde{\varepsilon}_t$ is an i.i.d. sequence with mean 0 and unit variance. Let $\sigma^2(p) = \gamma(0) - \mathbf{a}(p)'\boldsymbol{\Gamma}^{-1}(p)\mathbf{a}(p)$ and assume that $E(\widetilde{\varepsilon}_t^4) < \infty$. Note that $\boldsymbol{\Gamma}(p)^{-1}$ exists for every $p \in \mathbb{N}$ since $\gamma(0) > 0$ and $\gamma(h) \to 0$ as $h \to \infty$; compare Brockwell and Davis (1991), Proposition 5.1.1. Furthermore, $\mathbf{a}(p)$ is the vector of coefficients of the best autoregressive fit in $L_2$-distance; that is, the coefficients $(a_1(p), a_2(p), \dots, a_p(p))'$ are defined uniquely as the arg min of the norm $E(X_t - \sum_{\nu=1}^{p} c_\nu X_{t-\nu})^2$.

Let $f_{\mathrm{AR}}(\lambda) = \sigma^2(p)|\Psi_{\mathrm{AR}}(e^{-i\lambda})|^2$ be the spectral density of $\widetilde{\mathbf{X}}$, where $\Psi_{\mathrm{AR}}(z) = 1/(1 - \sum_{\nu=1}^{p} a_\nu(p)z^\nu) =: 1/A_p(z)$ and consider random variables $Y_n(\lambda_j)$, $j = 0, 1, 2, \dots, N$, defined by

$$Y_n(\lambda_j) = q(\lambda_j)\widetilde{I}_n(\lambda_j), \tag{1.14}$$

where

$$q(\lambda) = \frac{f(\lambda)}{f_{\mathrm{AR}}(\lambda)} \tag{1.15}$$

and $\widetilde{I}_n(\lambda) = (2\pi n)^{-1}|\sum_{t=1}^{n} \widetilde{X}_t \exp\{-i\lambda t\}|^2$; that is, $\widetilde{I}_n(\lambda)$ is the periodogram based on observations $\widetilde{X}_1, \widetilde{X}_2, \dots, \widetilde{X}_n$ from $\widetilde{\mathbf{X}}$. Since the periodogram $\widetilde{I}_n(\lambda)$ satisfies (1.11) and (1.12) with $f$ replaced by $f_{\mathrm{AR}}$ and $E(\varepsilon_1^4)/\sigma^4$ replaced by $E(\widetilde{\varepsilon}_1)^4/\sigma^4(p)$, we get using (1.15) that

$$E\big(Y_n(\lambda_j)\big) = f(\lambda_j) + O(n^{-1}) \tag{1.16}$$

and

$$\operatorname{Cov}\big(Y_n(\lambda_j), Y_n(\lambda_k)\big) = \begin{cases} f^2(\lambda_j) + O(n^{-1}), & \text{for } j = k, \\ n^{-1} f(\lambda_j) f(\lambda_k)\Big(\dfrac{E\widetilde{\varepsilon}_1^4}{\sigma^4(p)} - 3\Big) + o(n^{-1}), & \text{for } j \neq k. \end{cases} \tag{1.17}$$

Furthermore, and as for the ordinary periodogram ordinates, if $f(\lambda) > 0$ for all $\lambda \in [0, \pi]$, then for a set of frequencies $0 < \lambda_1 < \lambda_2 < \cdots < \lambda_m < \pi$ the random vector $(Y_n(\lambda_1), Y_n(\lambda_2), \ldots, Y_n(\lambda_m))'$ is asymptotically distributed as a vector of independent and exponentially distributed variables, the $s$th component of which has mean $f(\lambda_s)$ and variance $f^2(\lambda_s)$.

Since the random variables $Y_n(\lambda_j)$ resample closely the random behavior of the periodogram ordinates $I_n(\lambda_j)$, the above results suggest the following: if the (asymptotic) distribution of a statistic based on $I_n(\lambda_j)$ is not affected by the dependence of the periodogram, then this distribution can be well approximated by the distribution of the corresponding statistic based on the random variables $Y_n(\lambda_j)$. Furthermore, if $E\widetilde{\varepsilon}_1^4/\sigma^4(p)$ is close to $E\varepsilon_1^4/\sigma^4$ such an approximation will also be valid for periodogram statistics for which the dependence of the periodogram ordinates affects the asymptotic distribution of interest. We expect this to be true since in this case, and as (1.17) shows, the covariance of the random variables $Y_n(\lambda_j)$ also mimics correctly the covariance of the corresponding periodogram ordinates $I_n(\lambda_j)$. An important case where this is true is the case where the underlying process $\mathbf{X}$ belongs to the infinite-order autoregressive class (1.4) and the order $p$ of the autoregressive approximation increases (at an appropriate rate) as the sample size $n$ increases.

An implementation of the above idea for bootstrapping the periodogram is presented in Section 3. We mention here, however, that the basic idea underlying the bootstrap procedure proposed in this paper, which combines a parametric autoregressive approximation of the process $\mathbf{X}$ with a nonparametric "correction" function $q$ to resample the stochastic behavior of the periodogram $I_n(\lambda)$, can also be applied using parametric classes of processes other than the autoregressive one. For instance, we could also have considered a finite-order moving average approximation of the process $\mathbf{X}$ and defined, in a similar way as in (1.15), an appropriate correction function $q$, that is, $q(\lambda) = f(\lambda)/f_{\mathrm{MA}}(\lambda)$, where $f_{\mathrm{MA}}(\cdot)$ denotes the spectral density of the approximating moving average model. This will make a restriction to the process class (1.4) for statistics like those given in (1.5) superfluous. However, we rely in the following on the autoregressive approximation because it is a computationally easier and faster technique which is successful in many situations.

The paper is organized as follows. Section 2 states the technical assumptions needed. Section 3 describes in detail the proposed bootstrap procedure. Section 4 deals with applications of this procedure in approximating the sampling behavior of the so-called ratio statistics and of spectral means, while Section 5 deals with nonparametric estimators in the frequency domain. In Section 6, some practical issues are discussed and a small simulation example is presented. Proofs of the main theorems as well as of some technical lemmas are deferred to Section 7.

**2. Assumptions.** In this section, we precisely state the conditions imposed on the class of stochastic processes and on the class of periodogram-based statistics considered.

We assume that the underlying process $\mathbf{X}$ satisfies the following assumption:

(A1) $\{X_t; t \in \mathbb{Z}\}$ is a real-valued stationary process

$$X_t = \sigma \cdot \sum_{\nu=-\infty}^{\infty} \alpha_\nu \varepsilon_{t-\nu}, \qquad \alpha_0 = 1,$$

where $\sum_{\nu=-\infty}^{\infty} |\nu|\,|\alpha_\nu| < \infty$.

For the error process $\{\varepsilon_t\}$, we assume that:

(A2) $\{\varepsilon_t; t \in \mathbb{Z}\}$ constitutes a sequence of independent and identically distributed random variables with $E\varepsilon_t = 0$, $E\varepsilon_t^2 = 1$ and $E\varepsilon_t^4 < \infty$. Furthermore, $\sigma \in (0, \infty)$ and $\kappa_4$ denotes the fourth-order cumulant of $\varepsilon_t$.

The following condition on the function $\varphi$ appearing in the definition of spectral means $M(I_n, \varphi)$ or ratio statistics $R(I_n, \varphi)$ is common in the literature [cf. also Dahlhaus (1985)]:

(A3) $\varphi : [0, \pi] \to \mathbb{R}$ is a bounded function of bounded variation. We assume that $\varphi$ is extended to the real line with $\varphi(-\lambda) = \varphi(\lambda)$ and $\varphi(\lambda + 2\pi) = \varphi(\lambda)$.

To obtain nonparametric kernel estimators of the spectral density $f$ or of the function $q$, the following basic assumptions (A4)–(A6) are imposed on the smoothing kernel $K$ and the smoothing bandwidths $h$ and $b$, respectively. The strengthening of the smoothness assumptions on $f$ and $K$ stated in (A7) and (A8) is necessary in order to deal with the bias of bootstrap nonparametric estimators.

(A4) $K$ denotes a nonnegative kernel function with compact support $[-\pi, \pi]$. The Fourier transform $k$ of $K$ is assumed to be a symmetric, continuous and bounded function satisfying $k(0) = 2\pi$ and $\int_{-\infty}^{\infty} k^2(u)\,du < \infty$.

Assumption (A4) implies that we have the following representation: $K(x) = (2\pi)^{-1} \int_{-\infty}^{\infty} k(u) e^{-iux}\,du$. Note that $k(0) = 2\pi$ implies that $(2\pi)^{-1} \times \int_{-\infty}^{\infty} K(u)\,du = 1$, while the symmetry of $k$ implies the same property for $K$.

(A5) $h \to 0$ as $n \to \infty$ such that $nh \to \infty$;
(A6) $b \to 0$ as $n \to \infty$ such that $nb \to \infty$;
(A7) the spectral density $f$ of $\mathbf{X}$ is three times continuous differentiable on $[-\pi, \pi]$;
(A8) $K$ is three times continuously differentiable on $[-\pi, \pi]$.

In some applications, we will narrow the class described in (A1) a little bit by requiring invertibility of the underlying process $\mathbf{X}$. In particular, we will assume that:

(B1) $\{X_t; t \in \mathbb{Z}\}$ is a real-valued stationary autoregressive process of infinite order

$$X_t = \sum_{\nu=1}^{\infty} a_\nu X_{t-\nu} + \sigma\varepsilon_t,$$

where $\sum_{\nu=1}^{\infty} \nu|a_\nu| < \infty$ and $1 - \sum_{\nu=1}^{\infty} a_\nu z^\nu \neq 0$ for all complex $z$ with $|z| \leq 1$.

Finally, for the class of processes described in (B1) and for some classes of statistics to be considered later, the order $p$ of the approximating autoregressive process is allowed to increase with the sample size $n$ and to be random. For $p = p(n)$, we will only require that

(B2) $p \in [p_{\min}(n), p_{\max}(n)]$ where $p_{\max}(n) \geq p_{\min}(n) \to \infty$ as $n \to \infty$ and $p_{\max}(n)^5 \log(n)/n$ stays bounded.

Observe that the above assumption implies that the autoregressive order $p$ may be chosen data dependent, that is, $p$ could be random and chosen according to an order selection criterion like AIC or FPE; compare Section 6 for a discussion.

**3. The bootstrap procedure.** Based on the motivation given in the Introduction, the bootstrap procedure investigated in this paper can be described by the following five steps:

STEP 1. Given the observations $X_1, \ldots, X_n$, we fit an autoregressive process of order $p$, where $p$ may depend on the particular sample at hand. This leads to estimated parameters $\hat{a}_1(p), \ldots, \hat{a}_p(p)$ and $\hat{\sigma}(p)$, which are obtained from the common Yule–Walker equations; compare Brockwell and Davis (1991). Consider the estimated residuals

$$\hat{\varepsilon}_t = X_t - \sum_{\nu=1}^{p} \hat{a}_\nu(p) X_{t-\nu}, \qquad t = p+1, \ldots, n,$$

and denote by $\hat{F}_n^c$ the empirical distribution of the standardized quantities $\hat{\varepsilon}_{p+1}, \ldots, \hat{\varepsilon}_n$; that is, $\hat{F}_n^c$ has mean 0 and unit variance.

STEP 2. Generate bootstrap observations $X_1^+, X_2^+, \ldots, X_n^+$, according to the following autoregressive model of order $p$:

$$X_t^+ = \sum_{\nu=1}^{p} \hat{a}_\nu(p) X_{t-\nu}^+ + \hat{\sigma}(p) \cdot \varepsilon_t^+,$$

where $(\varepsilon_t^+)$ constitutes a sequence of i.i.d. random variables with cumulative distribution function $\hat{F}_n^c$ (conditionally on the given observations $X_1, \ldots, X_n$).

The bootstrap process $\mathbf{X}^+ = (X_t^+ : t \in \mathbb{Z})$ possesses the following spectral density:

$$\hat{f}_{\mathrm{AR}}(\lambda) = \frac{\hat{\sigma}^2(p)}{2\pi} \left| 1 - \sum_{\nu=1}^{p} \hat{a}_\nu(p) e^{-i\nu\lambda} \right|^{-2}, \qquad \lambda \in [0, \pi].$$

Note that because we make use of the Yule–Walker parameter estimators in Step 1 it is always ensured that $\hat{f}_{\mathrm{AR}}$ is well defined; that is, the polynomial $1 - \sum_{\nu=1}^{p} \hat{a}_\nu(p) z^\nu$ has no complex roots with magnitude less than or equal to 1. Moreover, the bootstrap autocovariances $\gamma^+(h) = E^+ X_1^+ X_{1+h}^+$, $h = 0, 1, \dots, p$, coincide with the empirical autocovariances $\hat{\gamma}_n(h)$ of the underlying observations. It should be noted that it is convenient, but not necessary, to work with Yule–Walker parameter estimates. Any $\sqrt{n}$-consistent parameter estimates would suffice.

STEP 3. Compute the periodogram of the bootstrap observations, that is,

$$I_n^+(\lambda) = \frac{1}{2\pi n} \left| \sum_{t=1}^{n} X_t^+ e^{-i\lambda t} \right|^2, \qquad \lambda \in [0, \pi].$$

STEP 4. Define the following nonparametric estimator $\hat{q}$:

$$\hat{q}(\lambda) = \frac{1}{n} \sum_{j=-N}^{N} K_h(\lambda - \lambda_j) \frac{I_n(\lambda_j)}{\hat{f}_{\mathrm{AR}}(\lambda_j)} \qquad \text{for } \lambda \in [0, \pi),$$

while, for $\lambda = \pi$, $\hat{q}(\pi)$ is defined as twice the quantity on the right-hand side of the above equation taking into account that no Fourier frequencies greater than $\pi$ exist. Here, and above, the $\lambda_j$'s denote the Fourier frequencies, $K : [-\pi, \pi] \to [0, \infty)$ denotes a probability density (kernel), $K_h(\cdot) = h^{-1} K(\cdot/h)$ and $h > 0$ is the so-called bandwidth; compare (A4) and (A5).

STEP 5. Finally, the bootstrap periodogram $I_n^*$ is defined as follows:

$$I_n^*(\lambda) = \hat{q}(\lambda) I_n^+(\lambda), \qquad \lambda \in [0, \pi].$$

Some remarks are now in order. Although the theory developed in the next sections allows for a data-dependent order $p$ to be as flexible as possible [cf. assumption (B2)], a nonparametric correction in the final step is introduced in order to catch data features, which can be or are not represented by the autoregressive fit. This nonparametric correction is done via the function $\hat{q}$ and there are several reasons justifying its use. First of all, the nonparametric correction in Steps 4 and 5 makes the proposed bootstrap procedure applicable to a more general class of stochastic processes than the purely autoregressive bootstrap. As will be seen in the next sections, this is in particular true for periodogram statistics based on realizations of the process (1.1) which cannot be captured by the purely autoregressive bootstrap, that is, by the corresponding statistics based on the pseudo-periodogram values $I_n^+(\lambda)$. On the other hand, the parametric approximation makes the new procedure more general than the nonparametric residual-based bootstrap or the local periodogram bootstrap procedure. This is true since the new procedure leads to asymptotically valid approximations for a larger class of periodogram statistics than the aforementioned procedures if the underlying process belongs to the important infinite-order autoregressive class (1.4). Finally, and as in the spirit of the so-called prewhitening idea in nonparametric spectral density estimation [cf. Press and Tukey (1956)], we expect an improved behavior of the spectral density estimator $\hat{q} \cdot \hat{f}_{\mathrm{AR}}$ (implicitly) used by our bootstrap procedure. If, for example, the true underlying spectral density has some dominant peaks, then prewhitening leads to a considerable improvement of the estimator. The reason is that an autoregressive fit is really able to catch the peaks of the spectral density rather well and the curve $I_n(\lambda)/\hat{f}_{\mathrm{AR}}(\lambda)$ is much smoother than $I_n(\lambda)$ and thus much easier to estimate nonparametrically.

To elaborate on the differences between the nonparametric residual-based bootstrap procedure of the periodogram and the autoregressive-aided periodogram bootstrap proposed in this paper, recall first that, under the assumptions of the paper and the definition of $q$, we have

$$I_n(\lambda_j) = 2\pi q(\lambda_j) f_{\mathrm{AR}}(\lambda_j) I_{n,\varepsilon}(\lambda_j) + R_n(\lambda_j), \tag{3.1}$$

where $I_{n,\varepsilon}(\lambda)$ denotes the periodogram of the i.i.d. series $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ and the remainder $R_n(\lambda_j)$ satisfies $\max_{\lambda_j \in [0,\pi]} E(R_n(\lambda_j))^2 = O(n^{-1})$; compare Brockwell and Davis (1991), Theorem 10.3.1. Furthermore, by the fact that $I_n^*(\lambda) = \hat{q}(\lambda) I_n^+(\lambda)$, a similar expression can be obtained for $I_n^*(\lambda_j)$; that is, we have

$$I_n^*(\lambda_j) = 2\pi \hat{q}(\lambda_j) \hat{f}_{\mathrm{AR}}(\lambda_j) I_{n,\varepsilon^+}(\lambda_j) + R_n^*(\lambda_j), \tag{3.2}$$

where $I_{n,\varepsilon^+}(\lambda)$ denotes the periodogram of the i.i.d. series $\varepsilon_1^+, \varepsilon_2^+, \dots, \varepsilon_n^+$ and the remainder $R_n^*(\lambda_j)$ satisfies $\max_{\lambda_j \in [0,\pi]} E^*(R_n^*)^2(\lambda_j) = O_P(n^{-1})$; compare Lemma 7.3. Note that, like $I_{n,\varepsilon}(\lambda_j)$, the periodogram ordinates $I_{n,\varepsilon^+}(\lambda_j)$ are not independent. Finally, recall that the bootstrap periodogram ordinates of the nonparametric residual-based periodogram bootstrap are given by $I_n^\star(\lambda_j) = \hat{f}_h(\lambda_j) U_j^\star$, where $\hat{f}_h(\lambda) = n^{-1} \sum_{j=-N}^{N} K_h(\lambda - \lambda_j) I_n(\lambda_j)$ and $U_j^\star$ is an i.i.d. sequence based on the rescaled residuals $\hat{U}_j = I_n(\lambda_j)/\hat{f}(\lambda_j)$. Thus, the nonparametric residual-based periodogram bootstrap differs from the autoregressive-aided periodogram bootstrap not only by the independence of the generated bootstrap periodogram ordinates $I_n^\star(\lambda_j)$ but also by the estimator of the spectral density $f$ used. In particular, in the autoregressive-aided periodogram bootstrap the kernel estimator $\hat{f}_h$ is replaced by the (implicitly used) autoregressive-aided spectral density estimator $\tilde{f} = \hat{q} \cdot \hat{f}_{\mathrm{AR}}$. Note that (7.24) and (7.25) imply that, for every $p \in N$ fixed, the estimator $\tilde{f}$ is uniformly consistent, that is,

$$\sup_{\lambda \in [0,\pi]} |\tilde{f}(\lambda) - f(\lambda)| \to 0$$

in probability.

**4. Spectral means and ratio statistics.** The bootstrap analog of (1.6) now reads as follows:

$$\sqrt{n}\big(M(I_n^*,\varphi) - M(\tilde{f},\varphi)\big) = \sqrt{n}\bigg(\int_0^\pi \varphi(\lambda) I_n^*(\lambda)\,d\lambda - \int_0^\pi \varphi(\lambda)\tilde{f}(\lambda)\,d\lambda\bigg), \tag{4.1}$$

where

$$\tilde{f}(\lambda) = \hat{q}(\lambda)\hat{f}_{\mathrm{AR}}(\lambda). \tag{4.2}$$

Alternatively, we may, as above, consider the following discretized bootstrap statistic:

$$\frac{2\pi}{\sqrt{n}}\sum_{j=0}^{N}\varphi(\lambda_j)\big(I_n^*(\lambda_j) - \tilde{f}(\lambda_j)\big). \tag{4.3}$$

We can now present the following theorem, which states that our bootstrap procedure works.

THEOREM 4.1. (i) *Assume* (B1) *and* (B2) *and* (A2)–(A5). *Then we have* (*in probability*)

$$\begin{aligned}&\mathcal{L}\big[\sqrt{n}\big(M(I_n^*,\varphi) - M(\tilde{f},\varphi)\big)\big|X_1,\dots,X_n\big]\\ &\qquad\Rightarrow \mathcal{N}\bigg(0, 2\pi\int_0^\pi \varphi^2 f^2\,d\lambda + \kappa_4\bigg(\int_0^\pi \varphi f\,d\lambda\bigg)^2\bigg).\end{aligned}$$

(ii) *Assume* (A1)–(A5). *Then we have for all fixed $p \in \mathbb{N}$ that the same assertion as in* (i) *holds true with $\kappa_4$ replaced by* $\kappa_4(p) = E(X_p - \sum_{\nu=1}^{p} a_\nu(p)X_{p-\nu})^4/\sigma(p)^4 - 3$.

From Dahlhaus and Janas (1996) and Paparoditis and Politis (1999), it is known that the residual-based frequency domain bootstrap and the local bootstrap work for the so-called ratio statistics [cf. (1.9)]. Thus, it is worth studying the behavior of our autoregressive-aided frequency domain bootstrap for such statistics. From Theorem 4.1(i), it is clear that under the assumptions of this part of the theorem our bootstrap proposal works for ratio statistics. More interesting is the question whether the autoregressive-aided frequency domain bootstrap works for ratio statistics even if we keep the order $p$ of the autoregressive fit fixed. To this end, observe that we have as in Dahlhaus and Janas (1996)

$$\begin{aligned}\sqrt{n}\big(R(I_n^*,\varphi) - R(\tilde{f},\varphi)\big) &= \sqrt{n}\bigg(\frac{\int_0^\pi \varphi(\lambda)I_n^*(\lambda)\,d\lambda}{\int_0^\pi I_n^*(\lambda)\,d\lambda} - \frac{\int_0^\pi \varphi(\lambda)\tilde{f}(\lambda)\,d\lambda}{\int_0^\pi \tilde{f}(\lambda)\,d\lambda}\bigg)\\ &= \frac{\sqrt{n}}{\int\tilde{f}(\lambda)\,d\lambda\int I_n^*(\lambda)\,d\lambda}\int_0^\pi \tilde{\psi}(\lambda)I_n^*(\lambda)\,d\lambda,\end{aligned} \tag{4.4}$$

with $\tilde{\psi}(\lambda) = \varphi(\lambda) \int \tilde{f}(\lambda)\, d\lambda - \int \varphi(\lambda) \tilde{f}(\lambda)\, d\lambda$. Since $\int \tilde{\psi}(\lambda) \tilde{f}(\lambda)\, d\lambda = 0$ [which also implies that the limit $\int \psi f \, d\lambda$ is equal to 0 and therefore from Theorem 4.1(i) that the asymptotic distribution does not depend on fourth-order cumulants], we immediately obtain the following corollary.

COROLLARY 4.1. (i) *Assume* (B1) *and* (B2) *and* (A2)–(A5). *Then we have* (*in probability*)

$$\mathcal{L}\big[\sqrt{n}\big(R(I_n^*, \varphi) - R(\tilde{f}, \varphi)\big)\big| X_1, \dots, X_n\big]$$
$$\Rightarrow \mathcal{N}\left(0, 2\pi \int_0^\pi \psi^2(\lambda) f^2(\lambda)\, d\lambda \Big/ \left(\int f d\lambda\right)^4\right),$$

*where* $\psi(\lambda) = \varphi(\lambda) \int f(u)\, du - \int \varphi(u) f(u)\, du$.

(ii) *Assume* (A1)–(A5). *Then we have for all fixed* $p \in \mathbb{N}$ *that the same assertion as in* (i) *holds true*.

*Thus*, *in both cases the limiting normal distribution in the bootstrap world coincides with the limiting distribution of ratio statistics* [*cf.* (1.10)].

It is worth mentioning that the results of Theorem 4.1 and Corollary 4.1, which state the asymptotic distributions of spectral means and ratio statistics in the bootstrap world, do not need assumption (A2) [saying that the innovations $(\varepsilon_t)$ are i.i.d. random variables] in its full strength. Actually, it is possible to prove similar central limit theorems under much weaker assumptions on the innovations. For instance, it suffices to assume that the innovations $(\varepsilon_t)$ form a fourth-order stationary and ergodic sequence. The reason we do not work under such an assumption in this paper is that the variance of the asymptotic distribution of spectral means given by formula (1.7) is valid for linear processes with an i.i.d. error structure only and that we are interested in mimicking this distribution in the bootstrap world.

**5. Nonparametric estimators.** An interesting class of consistent spectral density estimators is given by

$$\hat{f}(\lambda) = \frac{1}{n} \sum_{j=-N}^{N} K_b(\lambda - \lambda_j) I_n(\lambda_j), \tag{5.1}$$

where $K(\cdot)$ is a kernel satisfying (A4), $K_b(\cdot) = b^{-1} K(\cdot/b)$ and $b = b(n)$ is a smoothing bandwidth satisfying (A6). In the following, we are interested in estimating the distribution of the statistic

$$\sqrt{nb}\big(\hat{f}(\lambda) - f(\lambda)\big). \tag{5.2}$$

For this, the bootstrap statistic

$$\sqrt{nb}\big(\hat{f}^*(\lambda) - \tilde{f}(\lambda)\big) \tag{5.3}$$

is used, where

$$\hat{f}^*(\lambda) = \frac{1}{n} \sum_{j=-N}^{N} K_b(\lambda - \lambda_j) I_n^*(\lambda_j). \tag{5.4}$$

The following theorem shows that if the underlying process satisfies (1.1) then the proposed bootstrap procedure works. For this and in order to metrize the distance between distributions, we use in the following theorem Mallows' $d_2$ metric on the space $\{\mathcal{P} : \mathcal{P}$ probability measure on $(\mathbb{R}, \mathcal{B})$, $\int |x|^2 \, d\mathcal{P} < \infty\}$. This metric is defined according to $d_2(\mathcal{P}_1, \mathcal{P}_2) = \inf\{E|Y_1 - Y_2|^2\}^{1/2}$, where the infimum is taken over all real-valued random variables $(Y_1, Y_2)$ which have marginal distributions $\mathcal{P}_1$ and $\mathcal{P}_2$, respectively; compare Bickel and Freedman (1981) for more details.

THEOREM 5.1. *Suppose that assumptions* (A1), (A2) *and* (A4)–(A8) *are satisfied. Then, for all fixed* $p \in \mathbb{N}$, *the following are true*:

(i) *If* $nb^5 \to 0$, *then*

$$d_2\{\mathcal{L}(\sqrt{nb}(\hat{f}(\lambda) - f(\lambda))), \mathcal{L}(\sqrt{nb}(\hat{f}^*(\lambda) - \tilde{f}(\lambda))|X_1, X_2, \dots, X_n)\} \to 0$$

*in probability.*

(ii) *If* $b \sim n^{-1/5}$ *and* $nh^3 \to \infty$, *then the same result as in* (i) *holds true.*

To elaborate on the assumption $nh^3 \to \infty$ needed in the second part of the above theorem, note that if $nb^5 \to 1$ then the bias $E(\hat{f}(\lambda)) - f(\lambda)$ of the nonparametric estimator (5.1) is asymptotically not negligible. It rather converges to $(1/4\pi) f''(\lambda) \int u^2 K(u) \, du$ as $n \to \infty$; compare Priestley (1981). To provide a valid approximation of the distribution of $\sqrt{nb}(\hat{f}(\lambda) - f(\lambda))$ in this case, too, the bootstrap has to be able to estimate the bias term correctly. The condition $nh^3 \to \infty$ implies that $h$ tends to 0 more slowly than $b$; that is, $\hat{q}(\lambda)$ should be somewhat smoother than the optimal (with respect to minimizing the mean square error) kernel estimator of $q$. Therefore, the above assumption can be interpreted as an oversmoothing assumption, which is common in applications of the bootstrap to approximate the bias in nonparametric estimation; compare Romano (1988), Franke and Härdle (1992) and Paparoditis and Politis (1999).

## 6. Practical aspects and numerical examples.

6.1. *Some remarks on choosing the bootstrap parameters.* Before proceeding with the investigation of the finite-sample performance of our bootstrap procedure, some remarks on the choice of the parameters $p$ and $h$ are in order. Here we restrict our discussion to some rather heuristic ideas on how to choose these parameters, which may be helpful as guidelines in applications. Clearly, more theoretical work is required in order to make definite recommendations.

Our considerations in the rest of the paper are based on the following observation. For $\theta > 0$ denote by $\mathrm{Exp}(\theta)$ the exponential distribution with parameter $\theta$ and recall that, for $0 < \lambda < \pi$ and $f(\lambda) > 0$,

$$d_2\big(\mathcal{L}\big(I_n(\lambda)\big), \mathrm{Exp}\big(f(\lambda)\big)\big) \to 0.$$

Furthermore, by standard arguments it can be shown that

$$d_2\big(\mathcal{L}\big(I_n^*(\lambda)|X_1, X_2, \dots, X_n\big), \mathrm{Exp}\big(\tilde{f}(\lambda)\big)\big) \to 0$$

in probability. By the triangule inequality, we then observe that the quality of the approximation of the distribution of $I_n(\lambda)$ by the conditional distribution of $I_n^*(\lambda)$ depends heavily on how close the estimator $\tilde{f}(\lambda)$ used in our bootstrap procedure is to the unknown spectral density $f(\lambda)$. A choice of $p$ and $h$ can, therefore, be achieved by minimizing a measure of the stochastic distance of $\tilde{f}(\lambda)$ from $f(\lambda)$.

Consider first the problem of choosing the autoregressive order $p$. If the underlying process obeys the infinite-order autoregressive representation [cf. assumption (B1)] with the error process satisfying (A2), then condition (B2) imposed on the behavior of $p$ covers quite general situations allowing, for instance, $p$ to be chosen data dependent. One way to choose $p$ is according to the AIC criterion, which minimizes the function $\mathrm{AIC}(p) = \arg\min_p\{\hat{\sigma}^2(p)(1 + 2p/n)\}$ over a range of values $p = 1, 2, \dots, p_{\max}(n)$, where, for example, $p_{\max}(n) = 10\log_{10}(n)$, which is the default value in S-PLUS. It is known that under certain regularity conditions such a choice of $p$ leads to an asymptotically optimal autoregressive spectral estimator $\hat{f}_{\mathrm{AR}}$ with respect to minimizing the relative squared error $\int_{-\pi}^{\pi}((\hat{f}_{\mathrm{AR}}(\lambda) - f(\lambda))/f(\lambda))^2\, d\lambda$; compare Shibata (1981). Choosing $p$ according to this criterion is covered by our assumption (B2).

The above discussion on the choice of $p$ is, however, no more valid if the underlying process does not obey an infinite-order autoregressive structure. In this case, an appropriate autoregressive fit can rely on the idea of prewhitening. Understanding such a prewhitening as a graphical device, the autoregressive order $p$ can be selected as the smallest $p$ over the range $1, 2, \dots, p_{\max}(n)$ for which the smoothed rescaled periodogram $I_n(\lambda_j)/\hat{f}_{\mathrm{AR}}(\lambda_j)$, $j = 1, 2, \dots, N$, is closest to a constant.

The second parameter to be selected is the smoothing parameter $h$ used to estimate the unknown function $q$. Given an autoregressive order $p$ and a corresponding autoregressive spectrum $f_{\mathrm{AR}}$, the choice of the smoothing bandwidth $h$ can be based on the following considerations. Recall that the cross-validation criterion for choosing the bandwidth $h$ in the spectral density estimation proposed by Beltrão and Bloomfield (1987) [cf. also Robinson (1991)] is based on a discretized version of Whittle's approximation of minus twice the Gaussian likelihood given by

$$\sum_{j=1}^{N}\left\{\log f(\lambda_j) + \frac{I_n(\lambda_j)}{f(\lambda_j)}\right\}. \tag{6.1}$$

Following Beltrão and Bloomfield (1987), we introduce the leave-one-out estimate of $q(\lambda)$:

$$\hat{q}_{-j}(\lambda_j) = \frac{1}{n} \sum_{j \in \mathbf{N}_j} K_h(\lambda_j - \lambda_s) \frac{I_n(\lambda_j)}{\hat{f}_{\mathrm{AR}}(\lambda_j)}, \tag{6.2}$$

where $\mathbf{N}_j = \{s : -N \le s \le N \text{ and } j - s \ne \pm j \bmod N\}$. That is, $\hat{q}_{-j}$ is the kernel estimator of $q$ when the $j$th periodogram ordinate is deleted. Now, substituting $q(\lambda_j) f_{\mathrm{AR}}(\lambda_j)$ for $f(\lambda_j)$, $\hat{q}_{-j}(\lambda_j)$ for $q(\lambda_j)$ and $\hat{f}_{\mathrm{AR}}(\lambda_j)$ for $f_{\mathrm{AR}}(\lambda_j)$ in (6.1) leads, after ignoring the factor $\log f_{\mathrm{AR}}(\lambda_j)$, to the function

$$\mathrm{CV}(h) = \sum_{j=1}^{N} \left\{ \log \hat{q}_{-j}(\lambda_j) + \frac{I_n(\lambda_j)/\hat{f}_{\mathrm{AR}}(\lambda_j)}{\hat{q}_{-j}(\lambda_j)} \right\}. \tag{6.3}$$

The function (6.3) can now be used to assess the merits of different values of $h$ in estimating the function $q$ for $p \in \mathbb{N}$ and $\hat{f}_{\mathrm{AR}}$ given. According to such a cross-validation-type criterion, $h$ can be selected as the minimizer of (6.3).

6.2. *Numerical examples.* In this section, we compare the performance of the proposed autoregressive-aided periodogram bootstrap to that of the autoregressive bootstrap and that of the nonparametric residual-based periodogram bootstrap by means of simulations. Furthermore, we also include in our comparisons the block bootstrap method. In order to make such a comparison, we choose in the following a statistic for which all methods lead to asymptotically correct approximations. In particular, we study and compare the performance of the four aforementioned bootstrap methods in estimating the standard deviation $\sigma_1$ of the first-order sample autocorrelation $\hat{\rho}_n(1) = \hat{\gamma}_n(1)/\hat{\gamma}_n(0)$, where $\hat{\gamma}_n(h) = n^{-1} \sum_{t=1}^{n-h} (X_t - \overline{X})(X_{t+h} - \overline{X})$ is the sample autocovariance at lag $h$ and $\overline{X} = n^{-1} \sum_{t=1}^{n} X_t$.

Realizations of length $n = 50$ and $n = 400$ from the model

$$X_t = \varepsilon_t + \theta \varepsilon_{t-1},$$

with $\theta = 0.95$ and $\varepsilon_t \sim N(0, 1)$ have been considered. To estimate the exact standard deviation of $\hat{\rho}_n(1)$, 10,000 replications have been used while the bootstrap approximations are based on $B = 300$ bootstrap replications. In particular, the performance of the autoregressive-based and autoregressive-aided method has been studied for three different choices of the autoregressive order $p$ and a choice of $p$ based on the AIC criterion. Similarly, the nonparametric residual-based frequency domain bootstrap has been applied using three different choices of the smoothing parameter $h$ and a choice of $h$ based on the cross-validation criterion obtained by minimizing (6.3) after replacing the rescaled periodogram $\widetilde{I}_n(\lambda)$ appearing in (6.2) and (6.3) by the ordinary periodogram $I_n(\lambda)$. Recall that in the nonparametric residual-based periodogram bootstrap, $h$ is the bandwidth used to obtain $\hat{f}_h$ which is a kernel estimator of $f$. The cross-validation criterion has also been applied to choose the smoothing bandwidth $h$ used to estimate the nonparametric correction function $q(\cdot)$ of the autoregressive-aided periodogram bootstrap. In all cases, the Bartlett–Priestley kernel $K$ has been used. Finally, the block bootstrap procedure has been applied for four different choices of the block size parameter $l$.

The results are summarized in Table 1, where the mean value, the standard deviation and the mean square error of the four bootstrap approximations are reported as sample moments over 200 simulations. A more detailed picture of the behavior of the bootstrap estimators is given in Figure 1 where some boxplots of the distributions of the different bootstrap approximations for the cases $n = 50$ and $n = 400$ are presented.

As the entries in Table 1 show, the results of the autoregressive-aided periodogram bootstrap compare favorably with those of the other three methods. In particular and compared to the nonparametric residual-based periodogram bootstrap, we observe an overall decrease in the mean square error of the new periodogram bootstrap estimator. The table also shows a decrease in the variability of the mean square error of the autoregressive-aided periodogram bootstrap compared to that of the purely autoregressive bootstrap over the different choices of the bootstrap parameters, that is, the autoregressive order $p$. As this table confirms, this decrease is mainly due to a reduction in the bias of the bootstrap estimator which is caused by the nonparametric correction applied. The

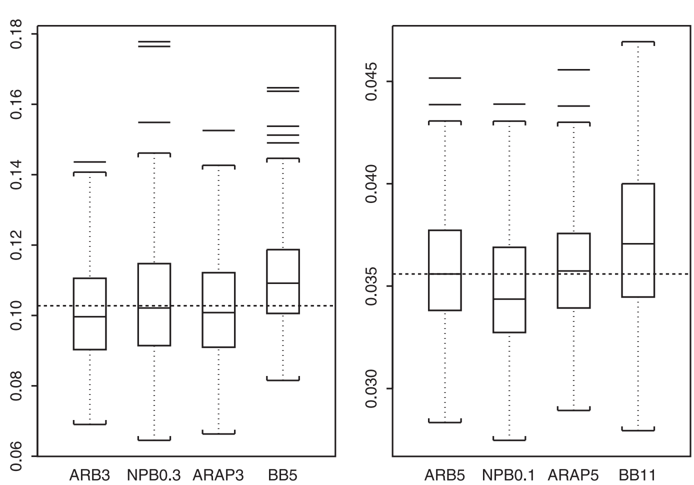


FIG. 1. *Boxplots of the bootstrap distributions with target indicated by the horizontal dashed line*. Left panel: $n = 50$, ARB3, *autoregressive bootstrap with* $p = 3$; NPB0.3, *nonparametric residual-based periodogram bootstrap with* $h = 0.3$; ARAP3, *autoregressive-aided periodogram bootstrap with* $p = 3$; BB5, *block bootstrap with* $l = 5$. Right panel: $n = 400$, ARB5, *autoregressive bootstrap with* $p = 5$; NPB0.1, *nonparametric residual-based periodogram bootstrap with* $h = 0.1$; ARAP5, *autoregressive-aided periodogram bootstrap with* $p = 5$; BB11, *block bootstrap with* $l = 11$.

TABLE 1

*Autoregressive bootstrap (AR-Boot), nonparametric residual-based periodogram bootstrap (NP-Boot), autoregressive-aided periodogram bootstrap (ARAP-Boot) and block bootstrap (Block-Boot) estimates of the standard deviation of the first-order sample autocorrelation*

| | | $n = 50$ | | | | $n = 400$ | | |
|---|---|---|---|---|---|---|---|---|
| | | **Mean($\hat{\sigma}_1^*$)** | **SD($\hat{\sigma}_1^*$) × 10** | **MSE($\hat{\sigma}_1^*$) × $10^3$** | | **Mean($\hat{\sigma}_1^*$)** | **SD($\hat{\sigma}_1^*$) × $10^2$** | **MSE($\hat{\sigma}_1^*$) × $10^5$** |
| Est. exact | | 0.103 | | | | 0.0356 | | |
| AR-Boot | $p = 1$ | 0.133 | 0.0782 | 0.9646 | $p = 5$ | 0.0362 | 0.2856 | 0.8480 |
| | $p = 3$ | 0.101 | 0.1445 | 0.2114 | $p = 7$ | 0.0356 | 0.2623 | 0.6882 |
| | $p = 5$ | 0.103 | 0.1532 | 0.2351 | $p = 9$ | 0.0366 | 0.2529 | 0.7324 |
| | $p_{\mathrm{AIC}}$ | 0.107 | 0.1640 | 0.2847 | $p_{\mathrm{AIC}}$ | 0.0368 | 0.2743 | 0.9070 |
| NP-Boot | $h = 0.1$ | 0.079 | 0.1619 | 0.7995 | $h = 0.075$ | 0.0345 | 0.3369 | 1.2521 |
| | $h = 0.2$ | 0.092 | 0.1779 | 0.4350 | $h = 0.100$ | 0.0348 | 0.3239 | 1.1157 |
| | $h = 0.3$ | 0.101 | 0.1789 | 0.3216 | $h = 0.150$ | 0.0369 | 0.3118 | 1.1401 |
| | $h_{\mathrm{CV}}$ | 0.099 | 0.1763 | 0.4515 | $h_{\mathrm{CV}}$ | 0.0342 | 0.3288 | 1.2690 |
| ARAP-Boot | $p = 1$ | 0.105 | 0.1178 | 0.1459 | $p = 2$ | 0.0356 | 0.2601 | 0.6775 |
| | $p = 3$ | 0.102 | 0.1507 | 0.2278 | $p = 5$ | 0.0362 | 0.2731 | 0.7876 |
| | $p = 5$ | 0.104 | 0.1617 | 0.2632 | $p = 7$ | 0.0356 | 0.2635 | 0.6951 |
| | $p_{\mathrm{AIC}}$ | 0.106 | 0.1617 | 0.2792 | $p_{\mathrm{AIC}}$ | 0.0368 | 0.2747 | 0.9072 |
| Block-Boot | $l = 3$ | 0.121 | 0.0966 | 0.4258 | $l = 9$ | 0.0382 | 0.3719 | 2.0649 |
| | $l = 5$ | 0.111 | 0.1482 | 0.2823 | $l = 11$ | 0.0375 | 0.3590 | 1.6373 |
| | $l = 7$ | 0.107 | 0.1723 | 0.3165 | $l = 13$ | 0.0369 | 0.4156 | 1.9147 |
| | $l = 9$ | 0.101 | 0.2009 | 0.4070 | $l = 15$ | 0.0364 | 0.4161 | 1.7953 |

results based on the new bootstrap procedure also seem to be less sensitive to the choice of the corresponding bootstrap parameters, which is probably due to the frequency domain nonparametric correction via the function $\hat{q}$. The effect of this nonparametric correction is clearly seen in comparing the results of the autoregressive-aided periodogram bootstrap to those of the purely autoregressive bootstrap for the case $p = 1$ and $n = 50$ or $p = 2$ and $n = 400$. Finally, the block bootstrap estimator suffers from a larger mean square error compared to the other three bootstrap-based estimators.

**7. Proofs and technical lemmas.** Let us collect some properties of the process $\mathbf{X}^+$; compare Section 3, Step 2. The process $\mathbf{X}^+$ possesses for all $p \in \mathbb{N}$ the following moving average representation:

$$X_t^+ = \hat{\sigma}(p) \sum_{\nu=0}^{\infty} \hat{\alpha}_\nu(p) \varepsilon_{t-\nu}^+, \qquad t \in \mathbb{Z}.$$

The coefficients $[\hat{\alpha}_\nu(p) : \nu \in \mathbb{N}_0]$ can be computed from $[\hat{\alpha}_0(p) = 1]$:

$$\left(1 - \sum_{\nu=1}^{p} \hat{a}_\nu(p) z^\nu\right)^{-1} = 1 + \sum_{\nu=1}^{\infty} \hat{\alpha}_\nu(p) z^\nu \qquad \text{for all } |z| \le 1. \tag{7.1}$$

Recall that the Yule–Walker estimators $[\hat{a}_\nu(p) : \nu = 1, \dots, p]$ always ensure invertibility of the fitted autoregressive model. Using Theorem 2.1 Hannan and Kavalieris (1986), we obtain under assumption (A1) and (A2) from Cauchy's inequality for holomorphic functions

$$|\hat{\alpha}_\nu(p) - \alpha_\nu(p)| \le \frac{p}{(1 + 1/p)^\nu} \mathcal{O}_P\left(\sqrt{\frac{\log(n)}{n}}\right), \tag{7.2}$$

uniformly in $\nu \in \mathbb{N}$ and $p \le p_{\max}(n)$. Here $[\alpha_\nu(p) : \nu \in \mathbb{N}]$ is exactly defined as $[\hat{\alpha}_\nu(p) : \nu \in \mathbb{N}]$; compare (7.1), with $\hat{a}_\nu(p)$ replaced by $a_\nu(p)$, $\nu = 1, \dots, p$. For details, we refer to Kreiss (1999), Section 8.

Baxter (1962), Theorem 2.2, and a standard Banach algebra argument lead under assumption (B1) for all large $p$ to

$$\sum_{\nu=0}^{\infty} |\alpha_\nu(p) - \alpha_\nu| \le \mathcal{O}_P\left(\sum_{\nu=p}^{\infty} |a_\nu|\right) = o_P(1). \tag{7.3}$$

Equations (7.2) and (7.3) imply for the autocovariances

$$\gamma^+(k) = E^+ X_t^+ X_{t+k}^+ = \sum_{\nu=0}^{\infty} \hat{\alpha}_{\nu+k}(p) \hat{\alpha}_\nu(p) \hat{\sigma}^2(p), \qquad k \in \mathbb{N}_0, \tag{7.4}$$

of the process $(X_t^+)$ that

$$\sum_{k=0}^{n} |\gamma^+(k)| = \mathcal{O}_P(1), \tag{7.5}$$

which, in turn, leads to

$$\frac{1}{n}\sum_{k=0}^{n} k|\gamma^{+}(k)| = o_P(1) \tag{7.6}$$

by Kronecker's lemma; compare Bauer (1974), Lemma 61.1.

In the case that we keep the autoregressive order $p$ fixed, then the bound in (7.2) is $\mathcal{O}_P(n^{-1/2})$ and (7.4)–(7.6) are valid as stated.

LEMMA 7.1. (i) *Assume* (B1), (B2), (A2), (A4) *and* (A5). *Then we have, for all* $\lambda \in [0, \pi]$,

$$\hat{q}(\lambda) \to 1 \qquad \textit{in probability.}$$

*Moreover,*

$$\int_0^{\pi} |\hat{q}(\lambda) - 1|\, d\lambda \to 0 \qquad \textit{in probability.}$$

(ii) *Assume* (A1), (A2), (A4) *and* (A5). *Then we have, for all fixed* $p \in \mathbb{N}$ *and all* $\lambda \in [0, \pi]$,

$$\hat{q}(\lambda) \to q(\lambda) := \frac{f(\lambda)}{f_{\mathrm{AR}}(\lambda)} \qquad \textit{in probability}$$

*and*

$$\int_0^{\pi} |\hat{q}(\lambda) - q(\lambda)|\, d\lambda \to 0 \qquad \textit{in probability.}$$

PROOF. Since the arguments are rather similar to those given for the consistency of the smoothed periodogram, we only give a sketch for the more complicated part (i) and indicate the modifications necessary for part (ii).

From Hannan and Kavalieris [(1986), Theorem 2.1] and Baxter [(1962), Theorem 2.2], we obtain [uniformly in $p \le p_{\max}(n)$] $\sup_{0\le\lambda\le\pi} |\hat{f}_{\mathrm{AR}}^{-1}(\lambda) - f^{-1}(\lambda)| = o_P(1)$. To see this, recall the representations $\hat{f}_{\mathrm{AR}}^{-1}(\lambda) = \hat{\sigma}^2(p)|1 - \sum_{\nu=1}^{p} \hat{a}_\nu(p) \times \exp\{-i\nu\lambda\}|^2/2\pi$ and $f^{-1}(\lambda) = \sigma^2|1 - \sum_{\nu=1}^{\infty} a_\nu \exp\{-i\nu\lambda\}|^2/2\pi$.

Since $f'$ is bounded because of $\sum_\nu \nu|a_\nu| < \infty$ and $f^{-1}$ is bounded because of the invertibility assumption of the underlying autoregressive model [cf. (B1)], we obtain

$$\hat{q}(\lambda) = \frac{1}{n f_{\mathrm{AR}}(\lambda)} \sum_{j=-N}^{N} K_h(\lambda - \lambda_j) I_n(\lambda_j) + o_P(1)\,,$$

where the remainder term is uniform in $\lambda$. Finally, since our assumptions imply smoothness of the spectral density $f$, we have, for all $\lambda \in [0, \pi]$, $n^{-1}\sum_{j=-N}^{N} K_h(\lambda - \lambda_j) I_n(\lambda_j) \to_{n\to\infty} f(\lambda)$ in probability.

To see the second assertion of (i), it suffices to show that $\int_0^\pi n^{-1}|\sum_{j=-N}^N K_h \times (\lambda-\lambda_j)(I_n(\lambda_j)-EI_n(\lambda_j))|\,d\lambda$ converges to 0 in probability, because $EI_n(\lambda)$ converges under our assumptions uniformly to $f(\lambda)$; compare Brockwell and Davis (1991), Proposition 10.3.1. The expectation of the square of this last expression is bounded through

$$\begin{aligned}
&\int_0^\pi \frac{1}{n^2}\sum_{j=-N}^N K_h^2(\lambda-\lambda_j)\operatorname{Var} I_n(\lambda_j)\,d\lambda \\
&\quad+\int_0^\pi \frac{1}{n^2}\sum_{j,k=-N,k\neq j}^N K_h(\lambda-\lambda_j)K_h(\lambda-\lambda_k)\operatorname{Cov}\bigl(I_n(\lambda_j),I_n(\lambda_k)\bigr)\,d\lambda \\
&\quad=\frac{1}{n^2h^2}\mathcal{O}(nh)+\frac{1}{n^2h^2}\mathcal{O}(nh^2)=o(1); \qquad \text{compare (1.2).}
\end{aligned}$$

For a proof of (ii), replace $f$ by $f_{\mathrm{AR}}$. $\square$

LEMMA 7.2. (i) *Assume* (B1), (B2) *and* (A2). *Then we have* $\sum_{k=0}^\infty \sqrt{k}\times|\hat\alpha_k(p)|<\infty$.

(ii) *Assume* (A1) *and* (A2) *and let* $p\in\mathbb{N}$ *be fixed. Let* $\tilde{\mathbf{a}}=(\tilde a_1(p),\tilde a_2(p),\dots,\tilde a_p(p))'$ *be an* $\sqrt{n}$*-consistent estimator of* $\mathbf{a}(p)$ *which satisfies* $1-\sum_{\nu=1}^p\tilde a_\nu(p)\times z^\nu\neq 0$ *for* $|z|\le 1$. *Then* $\sum_{k=1}^\infty k^\delta|\tilde\alpha_k(p)|=O_P(1)$ *for any* $\delta\in(0,\infty)$.

PROOF. From (7.2), we have that

$$\begin{aligned}
\sum_{k=0}^\infty \sqrt{k}|\hat\alpha_k(p)-\alpha_k(p)| &\le \sum_{k=0}^\infty \sqrt{k}(1+1/p)^{-k}\mathcal{O}_P\bigl(\sqrt{p^2\log(n)/n}\bigr)\\
&\le \mathcal{O}_P\bigl(\sqrt{p^5\log(n)/n}\bigr)=\mathcal{O}_P(1),
\end{aligned}$$

uniformly in $p\le p_{\max}(n)$. Thus, it suffices to consider $\sum_{k=0}^\infty\sqrt{k}|\alpha_k(p)-\alpha_k|$ and $\sum_{k=0}^\infty\sqrt{k}|\alpha_k|$. The fact that $\sum_{k=0}^\infty\sqrt{k}|\alpha_k|<\infty$ follows from (B1) and Hannan and Kavalieris (1986). In order to see that the first series is convergent, we obtain exactly along the lines of the proof of Lemma 8.2 in Kreiss (1999) that, for all $p$ large enough, $\sum_{k=0}^\infty\sqrt{k}|\alpha_k(p)-\alpha_k|\le C\sum_{k=0}^\infty\sqrt{k}|a_k(p)-a_k|$. From Baxter (1962), we finally obtain that, again for all $p$ large enough, $\sum_{k=0}^\infty\sqrt{k}|a_k(p)-a_k|\le C'\sum_{k=p}^\infty\sqrt{k}|a_k|=o_P(1)$, since $\sum_{k=0}^\infty\sqrt{k}|a_k|<\infty$. This implies (i).

To see (ii), recall that $1-\sum_{k=1}^p a_k(p)z^k\neq 0$ for $z\le 1$. More exactly, for each fixed $p$, $\epsilon>0$ exists such that the power series $(1-\sum_{k=1}^p a_k(p)z^k)^{-1}=1+\sum_{k=1}^\infty\alpha_k(p)z^k$ converges for $|z|<1+\epsilon$. This implies $\alpha_k(p)(1+\epsilon/2)^k\to 0$ as $k\to\infty$; that is, there exist positive constants $C>0$ and $\rho\in(0,1)$ such that $|\alpha_k(p)|\le C\rho^k$ for $k=1,2,\dots$. Now, $\sum_{k=1}^\infty k^\delta|\tilde\alpha_k(p)|\le\sum_{k=1}^\infty k^\delta|\alpha_k(p)|+$

$\sum_{k=1}^{\infty} k^{\delta}|\tilde{\alpha}_k(p) - \alpha_k(p)|$. Use Lemma 2.2 of Kreiss and Franke (1992) to bound the difference $|\tilde{\alpha}_k(p) - \alpha_k(p)|$ and get that, for some $\eta > 0$,

$$\begin{aligned}\sum_{k=1}^{\infty} k^{\delta}|\alpha_k(p)| &+ \sum_{k=1}^{\infty} k^{\delta}|\tilde{\alpha}_k(p) - \alpha_k(p)| \\ &\le O(1) + \sum_{k=1}^{\infty} k^{\delta}(1+\eta)^{-k} O_P(n^{-1/2}) = O_P(1).\end{aligned}$$

Note that the last equality follows using the fact that the $O_P(n^{-1/2})$ term is uniformly in $k$. $\square$

PROOF OF THEOREM 4.1. We prove part (i) only. First of all, we show that we can restrict consideration to

$$\sqrt{n}\int_0^{\pi} \varphi(\lambda)\big(I_n^{+}(\lambda) - E^{+}I_n^{+}(\lambda)\big)\hat{q}(\lambda)\,d\lambda. \tag{7.7}$$

To see this, observe that $I_n^*(\lambda) = I_n^+(\lambda)\hat{q}(\lambda)$ and that, for all $\lambda \in [0,\pi]$, $I_n^+(\lambda) = (\hat{\gamma}_n^+(0) + 2\sum_{k=1}^{n-1}\hat{\gamma}_n^+(k)\cos(\lambda k))/(2\pi)$, where $\hat{\gamma}_n^+(k) = \frac{1}{n}\sum_{t=1}^{n-k} X_t^+ X_{t+k}^+$, $k = 0, 1, 2, \dots, n-1$.

We have $E^+\hat{\gamma}_n^+(k) = (1-\frac{k}{n})E^+X_1^+X_{1+k}^+ = (1-\frac{k}{n})\gamma^+(k)$, $k = 0, 1, \dots, n-1$, $\hat{f}_{\mathrm{AR}}(\lambda) = (\gamma^+(0) + 2\sum_{k=1}^{\infty}\gamma^+(k)\cos(\lambda k))/(2\pi)$. Thus, in order to see that the difference between expressions (4.1) and (7.7) is $o_P(1)$ it suffices to show (7.8) and (7.9):

$$\sqrt{n}\sum_{k=n}^{\infty}\gamma^+(k)\int_0^{\pi}\varphi(\lambda)\hat{q}(\lambda)\cos(\lambda k)\,d\lambda = o_P(1), \tag{7.8}$$

$$\sum_{k=1}^{n-1}\frac{k}{\sqrt{n}}\gamma^+(k)\int_0^{\pi}\varphi(\lambda)\hat{q}(\lambda)\cos(\lambda k)\,d\lambda = o_P(1). \tag{7.9}$$

The boundedness of $\sum_{k=1}^{n}\sqrt{k}|\gamma^+(k)|$ in probability, Lemma 7.1 and the absolute summability of the Fourier coefficients of $\varphi$ imply (7.8) and (7.9).

Now, (7.7) can be rewritten as

$$\begin{aligned} Z_n := {} & \frac{1}{2\pi}\big(\sqrt{n}\big(\hat{\gamma}_n^+(k) - E^+\hat{\gamma}_n^+(k)\big) : k = 0, 1, \dots, n-1\big) \\ & \times \bigg(\int_0^{\pi}\varphi(\lambda)\hat{q}(\lambda)\,d\lambda, 2\int_0^{\pi}\varphi(\lambda)\hat{q}(\lambda)\cos(\lambda k)\,d\lambda : k = 1, \dots, n-1\bigg)' \\ & + o_P(1). \end{aligned} \tag{7.10}$$

From Kreiss (1999), Theorem 3.1, for each fixed $M \in \mathbb{N}$, we have that

$$\mathcal{L}\big(\sqrt{n}\big(\hat{\gamma}_n^+(k) - E^+\gamma_n^+(k)\big) : k = 0, \dots, M \mid X_1, \dots, X_n\big) \Rightarrow \mathcal{N}(0, V_M), \tag{7.11}$$

where

$$V_M = \left[ (E\varepsilon_1^4 - 3)\gamma(i)\gamma(j) + \sum_{k=-\infty}^{\infty} \big(\gamma(k)\gamma(k-i+j) + \gamma(k+j)\gamma(k-i)\big) \right]_{i,j=0}^{M}.$$

Because of Lemma 7.1, we can replace $\hat{q}(\cdot)$ in (7.10) by its limit 1. Denote by $Z_n'$ the quantity which is defined as $Z_n$, see (7.10), with this replacement. For $M \in \mathbb{N}$, decompose $Z_n'$ into the following two quantities:

$$(7.12)\qquad \begin{aligned} Z_{n,M}' &:= \frac{1}{2\pi}\big(\sqrt{n}(\hat{\gamma}_n^+(k) - E^+\hat{\gamma}_n^+(k)) : k = 0, 1, \dots, M\big) \\ &\qquad \times \left(\int_0^\pi \varphi(\lambda)\, d\lambda, 2\int_0^\pi \varphi(\lambda)\cos(\lambda k)\, d\lambda : k = 1, \dots, M\right)' \end{aligned}$$

and

$$(7.13)\qquad \begin{aligned} Z_n' - Z_{n,M}' &:= \frac{1}{\pi}\big(\sqrt{n}(\hat{\gamma}_n^+(k) - E^+\hat{\gamma}_n^+(k)) : k = M+1, \dots, n-1\big) \\ &\qquad \times \left(\int_0^\pi \varphi(\lambda)\cos(\lambda k)\, d\lambda : k = M+1, \dots, n-1\right)'. \end{aligned}$$

To obtain the asymptotic normality stated in the theorem, we have to show [cf. Brockwell and Davis (1991), Proposition 6.3.9]

$$(7.14)\qquad Z_{n,M}' \Rightarrow \mathcal{N}(0, \tau_M^2) \qquad \text{for all } M \in \mathbb{N},$$

$$(7.15)\qquad \tau_M^2 \to \tau^2 \qquad \text{as } M \to \infty$$

and

$$(7.16)\qquad \lim_{M\to\infty} \limsup_{n\to\infty} P^+\big\{|Z_n' - Z_{n,M}'| > \varepsilon\big\} = 0 \qquad \text{for all } \varepsilon > 0.$$

The result (7.14) is a direct consequence of (7.11). And (7.16) can be seen as follows: $E^+|Z_n' - Z_{n,M}'| \le \sum_{k=M+1}^{n-1} (n \operatorname{Var}^+(\hat{\gamma}_n^+(k)))^{1/2} |\int_0^\pi \varphi(\lambda)\cos(\lambda k)\, d\lambda|$. Since $n \operatorname{Var}^+(\hat{\gamma}_n^+(k))$ is bounded (in probability) uniformly in $k$, we obtain $\limsup_{n\to\infty} E^+|Z_n' - Z_{n,M}'| = o_P(1)$, as $M \to \infty$, because of the absolute summability of the Fourier coefficients of $\varphi$.

It remains to show (7.15). In a first step, it is easy to see that

$$\lim_{M\to\infty} \tau_M^2 = \frac{1}{4\pi^2}\kappa_4 \left(\sum_{r=0}^{\infty} \hat{\varphi}_r \gamma(r)\right)^2 + \frac{1}{4\pi^2} \sum_{k=-\infty}^{\infty} \sum_{r,s=0}^{\infty} \hat{\varphi}_r \hat{\varphi}_s \big\{\gamma(k)\gamma(k-r+s) + \gamma(k+s)\gamma(k-r)\big\},$$

where $\kappa_4 = E\varepsilon_1^4 - 3$, $\hat{\varphi}_0 = \int_0^\pi \varphi(\lambda)\,d\lambda$ and $\hat{\varphi}_r = 2\int_0^\pi \varphi(\lambda)\cos(r\lambda)\,d\lambda$, $r \geq 1$. This concludes the proof of Theorem 4.1, since

$$\begin{aligned}\int_0^\pi \varphi(\lambda) f(\lambda)\,d\lambda &= \int_0^\pi \varphi(\lambda)\frac{1}{2\pi}\left\{\gamma(0) + 2\sum_{r=1}^{\infty}\gamma(r)\cos(\lambda r)\right\}d\lambda \\ &= \frac{1}{2\pi}\sum_{r=0}^{\infty}\hat{\varphi}_r\gamma(r)\end{aligned}$$

and

$$\begin{aligned}&2\pi\int_0^\pi \varphi^2(\lambda) f^2(\lambda)\,d\lambda \\ &\quad = \frac{1}{\pi^2}\sum_{r,s=0}^{\infty}\hat{\varphi}_r\hat{\varphi}_s\int_0^\pi \cos(r\lambda)\cos(s\lambda)\left\{\gamma(0) + 2\sum_{k=1}^{\infty}\gamma(k)\cos(k\lambda)\right\}f(\lambda)\,d\lambda \\ &\quad = \frac{1}{4\pi^2}\sum_{r,s=0}^{\infty}\sum_{k=-\infty}^{\infty}\hat{\varphi}_r\hat{\varphi}_s\{\gamma(k)\gamma(k-r+s) + \gamma(k)\gamma(k+r+s)\} \\ &\quad = \frac{1}{4\pi^2}\sum_{r,s=0}^{\infty}\sum_{k=-\infty}^{\infty}\hat{\varphi}_r\hat{\varphi}_s\{\gamma(k)\gamma(k-r+s) + \gamma(k+s)\gamma(k-r)\},\end{aligned}$$

where, for the last equality, we have used the following addition formula of trigonometric functions: $\cos a \cos b \cos c = \frac{1}{4}(\cos(a+b-c) + \cos(b+c-a) + \cos(c+a-b) + \cos(a+b+c))$. $\square$

LEMMA 7.3. *Assume* (A1) *and* (A2). *Then the periodogram $I_n^+(\lambda_j)$ defined in Step* 3 *of the bootstrap algorithm satisfies*

$$I_n^+(\lambda_j) = \hat{\sigma}^2(p)\left|1 + \sum_{\nu=1}^{p}\hat{\alpha}_\nu(p)e^{-i\lambda_j\nu}\right|^2 I_{n,\varepsilon^+}(\lambda_j) + R_n^+(\lambda_j), \tag{7.17}$$

*where* $\max_{\lambda_j\in[0,\pi]} E^+(R_n^+(\lambda_j))^2 = O_P(n^{-1})$.

PROOF. Let $\lambda = \lambda_j$. Following the derivation of Brockwell and Davis [(1991), page 346], we get for the discrete Fourier transform $J_n^+(\lambda)$ of $\{X_t^+\}$ that

$$J_n^+(\lambda) = \sigma(p)\left(1 + \sum_{\nu=1}^{\infty}\hat{\alpha}_\nu(p)e^{-i\lambda_j\nu}\right)J_{\varepsilon^+}(\lambda) + \sigma(p)Y_n^+(\lambda), \tag{7.18}$$

where

$$J_{\varepsilon^+}(\lambda) = n^{-1/2}\sum_{t=1}^{n}\varepsilon_t^+ e^{-i\lambda t},$$

$$Y_n^+(\lambda) = n^{-1/2}\sum_{\nu=0}^{\infty}\hat{\alpha}_\nu(p)e^{-\lambda\nu}U_{n,\nu}^+(\lambda)$$

and

$$U_{n,\nu}^{+}(\lambda) = \sum_{t=1-\nu}^{n-\nu} \varepsilon_t^{+} e^{-\lambda t} - \sum_{t=1}^{n} \varepsilon_t^{+} e^{-\lambda t}.$$

Since $I_n^{+}(\lambda) = (2\pi)^{-1}|J_n^{+}(\lambda)|^2$, we have obtained expression (7.17) and the remainder $R_n^{+}(\lambda)$ is given by

$$\begin{aligned} R_n^{+}(\lambda) &= \hat{\sigma}^2(p)\left(1 + \sum_{\nu=1}^{\infty} \hat{\alpha}_\nu(p) e^{-i\lambda_j \nu}\right) J_{\varepsilon^+}(\lambda) Y_n^{+}(-\lambda) \\ &\quad + \hat{\sigma}^2(p)\left(1 + \sum_{\nu=1}^{\infty} \hat{\alpha}_\nu(p) e^{i\lambda_j \nu}\right) J_{\varepsilon^+}(-\lambda) Y_n^{+}(\lambda) + \hat{\sigma}^2(p)|Y_n^{+}(\lambda)|^2. \end{aligned} \tag{7.19}$$

Since, for $p$ fixed, $\hat{\mathbf{a}}(p)$ is an $\sqrt{n}$-consistent estimator of $\mathbf{a}(p)$ and $E^{+}(\varepsilon_1^{+})^2 = 1$, we get, using Lemma 7.2 and the bound $E^{+}|U_{n,\nu}^{+}|^4 \le 2|\nu| E^*(\varepsilon_1^{+})^4 + 12|\nu|^2$, that

$$\begin{aligned} E^{+}|Y_n^{+}(\lambda)|^4 &\le n^{-2}\left(\sum_{\nu=0}^{\infty} |\hat{\alpha}_\nu(p)| \big(2|\nu| E^*(\varepsilon_1^{+})^4 + 12|\nu|^2\big)^{1/4}\right)^4 \\ &= O_P(n^{-2}). \end{aligned} \tag{7.20}$$

Using expression (7.19), the assertion that $\max_{\lambda_j \in [0,\pi]} E^{+}(R_n^{+}(\lambda_j))^2 = O_P(n^{-1})$ follows then by the Cauchy–Schwarz inequality and taking into account Lemma 7.2, the fact that $E^{+}|J_{\varepsilon^+}(\lambda)|^2 = (2\pi)^{-1}$ and the bound (7.20). $\square$

To prove Theorem 5.1, we use the decomposition

$$\begin{aligned} &\sqrt{nb}\big(\hat{f}^*(\lambda) - \tilde{f}(\lambda)\big) \\ &\quad = \sqrt{\frac{b}{n}} \sum_{j=-N}^{N} K_b(\lambda - \lambda_j) \hat{q}(\lambda_j) \big(I_n^{+}(\lambda_j) - E^{+}\big(I_n^{+}(\lambda_j)\big)\big) \\ &\qquad + \sqrt{nb}\left(\frac{1}{n} \sum_{j=-N}^{N} K_b(\lambda - \lambda_j) \hat{q}(\lambda_j) E^{+}\big(I_n^{+}(\lambda_j)\big) - \hat{q}(\lambda) \hat{f}_{\mathrm{AR}}(\lambda)\right) \\ &\quad = L_n^{+}(\lambda) + B_n^{+}(\lambda) \end{aligned} \tag{7.21}$$

and an obvious notation for $L_n^{+}(\lambda)$ and $B_n^{+}(\lambda)$. The following two lemmas can then be established.

LEMMA 7.4. *Assume* (A1), (A2) *and* (A4)–(A6) *and let* $p \in \mathbb{N}$ *be fixed and* $\lambda \in [0,\pi]$. *If* $n \to \infty$, *then*

$$d_2\Big(\mathcal{L}\big(\sqrt{nb}\big(\hat{f}(\lambda) - E\big(\hat{f}(\lambda)\big)\big)\big), \mathcal{L}\big(L_n^{+}(\lambda) \mid X_1, X_2, \ldots, X_n\big)\Big) \to 0$$

*in probability.*

PROOF. Since convergence in the $d_2$ metric is equivalent to weak convergence and convergence of the first two moments [cf. Bickel and Freedman (1981), Lemma 8.3], it suffices to show that

$$E^+\big(L_n^+(\lambda)\big)^2 \to \tau^2(\lambda) := (1+\delta_{0,\pi}) f^2(\lambda)\frac{1}{2\pi}\int K^2(x)\,dx \tag{7.22}$$

and

$$\mathcal{L}\big(L_n^+(\lambda) \mid X_1, X_2, \dots, X_n\big) \Rightarrow N\big(0, \tau^2(\lambda)\big) \tag{7.23}$$

in probability, where $\delta_{0,\pi} = 1$ if $\lambda = 0$ or $\pi$ and $\delta_{0,\pi} = 0$ otherwise. Recall that $nb\operatorname{Var}(\hat f(\lambda)) \to \tau^2(\lambda)$ and $\sqrt{nb}(\hat f(\lambda) - E\hat f(\lambda)) \Rightarrow N(0, \tau^2(\lambda))$; compare Anderson (1971) for a different but asymptotically equivalent estimator.

Consider first (7.22). We have

$$\begin{aligned} E^+\big(L_n^+(\lambda)\big)^2 &= 4\pi^2\frac{b}{n}\sum_{j=-N}^{N} K_b^2(\lambda-\lambda_j)\hat q^2(\lambda_j)\hat f_{\mathrm{AR}}^2(\lambda_j)E^+\big(I_{n,\varepsilon^+}(\lambda_j)-1\big)^2 \\ &\qquad + O_P(b) \\ &= \frac{b}{2\pi}\int K_b^2(\lambda-x)\hat q^2(x)\hat f_{\mathrm{AR}}^2(x)\,dx + O_P(b) \\ &\to q^2(\lambda) f_{\mathrm{AR}}^2(\lambda)\frac{1}{2\pi}\int K^2(x)\,dx \end{aligned}$$

in probability, by the continuity of the functions $q$ and $f_{\mathrm{AR}}$ and the uniform convergences

$$\sup_{\lambda\in[0,\pi]} |\hat f_{\mathrm{AR}}(\lambda) - f_{\mathrm{AR}}(\lambda)| \to 0 \tag{7.24}$$

and

$$\sup_{\lambda\in[0,\pi]} |\hat q(\lambda) - q(\lambda)| \to 0 \tag{7.25}$$

in probability. Equation (7.24) follows using a standard Taylor series argument, the continuity of the derivative and the fact that, for $p$ fixed, $\hat{\mathbf{a}}(p)$ is an $\sqrt{n}$-consistent estimator of $\mathbf{a}(p)$. Equation (7.25) follows because by the arguments used in the proof of Lemma 7.1 we have $\hat q(\lambda) = n^{-1}\sum_{j=-N}^{N} K_h(\lambda-\lambda_j)I_n(\lambda_j)/f_{\mathrm{AR}}(\lambda_j) + O_P(n^{-1/2})$, where the $O_P(n^{-1/2})$ term is uniform in $\lambda \in [0,\pi]$. The uniform consistency of the first term on the right-hand side of the last equation as an estimator of $q$ follows then by standard arguments; compare, for instance, the proof of Theorem A1 of Franke and Härdle (1992).

We next show (7.23). For this, note first that, by Lemma 7.3,

$$\sqrt{\frac{b}{n}}\sum_{j=-N}^{N} K_b(\lambda-\lambda_j)\hat q(\lambda_j)R_n^+(\lambda_j) = O_P\big(\sqrt{b}\big)$$

and, therefore,

$$L_n^+(\lambda) = \sqrt{\frac{b}{n}} \sum_{j=-N}^{N} K_b(\lambda - \lambda_j)\tilde{f}(\lambda_j)\big(I_{n,\varepsilon^+}(\lambda_j) - 1\big) + o_P(1). \tag{7.26}$$

Now, instead of the first term on the right-hand side of the above equality, we consider the asymptotically equivalent statistic

$$\sqrt{nb}\,\frac{1}{2\pi}\int_{-\pi}^{\pi} K_b(\lambda - x)\tilde{f}(x)\big(I_{\varepsilon^+}(x) - 1\big)\,dx, \tag{7.27}$$

which appears by approximating the Riemann sum in (7.26) by the corresponding integral; compare Brillinger (1981), Theorem 5.9.1. We then have

$$\begin{aligned}
&\sqrt{nb}\frac{1}{2\pi}\int_{-\pi}^{\pi} K_b(\lambda - x)\tilde{f}(x)\big(I_{\varepsilon^+}(x) - 1\big)\,dx \\
&\quad = \sqrt{nb}\frac{1}{2\pi}\int_{-\pi}^{\pi} K(u)\tilde{f}(\lambda - ub)\big(I_{\varepsilon^+}(\lambda - ub) - 1\big)\,du \\
&\quad = \tilde{f}(\lambda)\sqrt{nb}\frac{1}{2\pi}\int_{-\pi}^{\pi} K(u)\big(I_{\varepsilon^+}(\lambda - ub) - 1\big)\,du + D_n^+(\lambda),
\end{aligned}$$

where

$$D_n^+(\lambda) = \sqrt{nb}\,\frac{1}{2\pi}\int_{-\pi}^{\pi} K(u)\big(\tilde{f}(\lambda - ub) - \tilde{f}(\lambda)\big)\big(I_{\varepsilon^+}(\lambda - ub) - 1\big)\,du.$$

Straightforward calculations yield that $E^+(D_n^+(\lambda)) = 0$ and $E^+(D_n^+(\lambda))^2 = O_P(|\tilde{f}(\lambda - ub) - \tilde{f}(\lambda)|^2) = O_P(b^2)$, where the last assertion follows by the uniform convergence of $\tilde{f}$ and the Lipschitz continuity of $f$. Thus, it suffices to consider the distribution of the asymptotically equivalent statistic

$$\tilde{f}(\lambda)\sqrt{nb}\frac{1}{2\pi}\int_{-\pi}^{\pi} K(u)\big(I_{\varepsilon^+}(\lambda - ub) - 1\big)\,du = \tilde{f}(\lambda)L_{1,n}^+(\lambda), \tag{7.28}$$

with an obvious notation for $L_{1,n}^+(\lambda)$. After substituting in $L_{1,n}^+(\lambda)$ the expression $I_{\varepsilon^+}(\lambda) = (2\pi)^{-1}\sum_{s=-n+1}^{n-1}\hat{\gamma}_{\varepsilon^+}(s)\cos(s\lambda)$, where $\hat{\gamma}_{\varepsilon^+}(s) = n^{-1}\sum_{t=1}^{n-s}\varepsilon_t^+\varepsilon_{t+s}^+$, we get

$$\begin{aligned}
L_{1,n}^+(\lambda) &= \frac{1}{\pi}\sqrt{nb}\sum_{s=1}^{n-1}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} K(u)\cos\big(s(\lambda - ub)\big)\,du\right)\hat{\gamma}_{\varepsilon^+}(s) \\
&\quad + \sqrt{nb}\big(\hat{\gamma}_{\varepsilon^+}(0) - 1\big)\frac{1}{2\pi}\int_{-\pi}^{\pi} K(u)\,du \\
&= \frac{1}{\pi}\sqrt{nb}\sum_{s=1}^{n-1}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} K(u)\cos(sub)\,du\right)\cos(s\lambda)\hat{\gamma}_{\varepsilon^+}(s) + O_P\big(\sqrt{b}\big) \\
&= \frac{1}{2\pi^2}\sqrt{nb}\sum_{s=1}^{n-1} k(sb)\cos(s\lambda)\hat{\gamma}_{\varepsilon^+}(s) + o_P(1),
\end{aligned}$$

where the last equality follows by (A4). Ignoring the factor $(2\pi^2)^{-1}$, using the definition $m_n = [1/b]$ and substituting for $\hat{\gamma}_{\varepsilon^+}(s)$, the first term on the right-hand side of the last equality above becomes

$$\begin{aligned}
&\sqrt{\frac{n}{m_n}} \sum_{s=1}^{n-1} k\left(\frac{s}{m_n}\right) \cos(s\lambda) \hat{\gamma}_{\varepsilon^+}(s) \\
&\quad = \frac{1}{\sqrt{n m_n}} \sum_{s=1}^{m_n(1+C)} \sum_{t=1}^{n-s} k\left(\frac{s}{m_n}\right) \cos(s\lambda) \varepsilon_t^+ \varepsilon_{t+s}^+ \\
&\qquad + \frac{1}{\sqrt{n m_n}} \sum_{s=m_n(1+C)+1}^{n-1} \sum_{t=1}^{n-s} k\left(\frac{s}{m_n}\right) \cos(s\lambda) \varepsilon_t^+ \varepsilon_{t+s}^+ \\
&\quad = \widetilde{L}_n^+ + \widetilde{R}_n^+,
\end{aligned}$$

with an obvious notation for $\widetilde{L}_n^+$ and $\widetilde{R}_n^+$. Here $C > 0$ is a constant to be specified later. Using the independence of the $\varepsilon_t^+$'s and $E^+(\varepsilon_t^+)^2 = 1$, we get for any $\delta > 0$ that

$$\begin{aligned}
&E^+(\widetilde{R}_n^+)^2 \\
&\quad = \frac{1}{n m_n} \sum_{s_1, s_2 = m_n(1+C)+1}^{n-1} \sum_{t_1, t_2 = 1}^{n-s} \prod_{i=1}^{2} k\left(\frac{s_i}{m_n}\right) \cos(s_i\lambda) E^+\left(\varepsilon_{t_1}^+ \varepsilon_{t_1+s_1}^+ \varepsilon_{t_2}^+ \varepsilon_{t_2+s_2}^+\right) \\
&\quad = \frac{1}{n m_n} \sum_{s=m_n(1+C)+1}^{n-1} \sum_{t=1}^{n-s} k^2\left(\frac{s}{m_n}\right) \cos^2(s\lambda) \\
&\quad \le \frac{1}{m_n} \sum_{j=1}^{n-m_n(1+C)-1} k^2\left(\frac{m_n(1+C)+j}{m_n}\right) \\
&\quad \to \int_{1+C}^{\infty} k^2(x)\, dx < \delta
\end{aligned}$$

for $C$ sufficiently large and because of (A4). Consider next the term $\widetilde{L}_n^+$ and verify by straightforward calculations that we can replace $\widetilde{L}_n^+$ by

$$\overline{L}_n^+ = \sum_{s=1}^{n-1} W_{t,n}^+,$$

where

$$W_{t,n}^+ = \frac{1}{\sqrt{n m_n}} \sum_{s=1}^{m_n(1+C)} k\left(\frac{s}{m_n}\right) \cos(s\lambda) \varepsilon_t^+ \varepsilon_{t+s}^+.$$

To see why, note that for $C$ sufficiently large,

$$\begin{aligned}
&E^{+}\Biggl(\frac{1}{\sqrt{nm_n}}\sum_{s=m_n(1+C)+1}^{n-1} k\Bigl(\frac{s}{m_n}\Bigr)\cos(s\lambda)\sum_{t=n-s+1}^{n-1}\varepsilon_t^{+}\varepsilon_{t+s}^{+}\Biggr)^2\\
&\quad\le \frac{1}{m_n}\sum_{s=m_n(1+C)+1}^{n-1} k^2\Bigl(\frac{s}{m_n}\Bigr)\\
&\quad\to \int_{1+C}^{\infty} k^2(x)\,dx < \delta.
\end{aligned}$$

Note that $E^{+}(W_{t,n}^{+}) = 0$ and that $m_n(1+C)/n \to 0$ as $n\to\infty$. By Theorem 4 of Rosenblatt [(1985), page 63], to establish asymptotic normality of $\sum_{t=1}^{n-1} W_{t,n}^{+}$ it suffices to show that

$$E^{+}\Biggl|\sum_{t=1}^{n-1} W_{t,n}^{+}\Biggr|^4 = O(1) \tag{7.29}$$

as $n\to\infty$. For this we have

$$\begin{aligned}
&E^{+}\Biggl|\sum_{t=1}^{n-1} W_{t,n}^{+}\Biggr|^4\\
&\quad= \frac{1}{n^2m_n^2}\sum_{t_1,t_2,t_3,t_4}\sum_{s_1,s_2,s_3,s_4}\prod_{i=1}^{4} k\Bigl(\frac{s_i}{m_n}\Bigr)\cos(s_i\lambda)\\
&\qquad\qquad \times E^{+}(\varepsilon_{t_1}^{+}\varepsilon_{t_1+s_1}^{+}\varepsilon_{t_2}^{+}\varepsilon_{t_2+s_2}^{+}\varepsilon_{t_3}^{+}\varepsilon_{t_3+s_3}^{+}\varepsilon_{t_4}^{+}\varepsilon_{t_4+s_4}^{+}).
\end{aligned}\tag{7.30}$$

Evaluation of the above expectation shows that (7.29) is valid. Note that in evaluating (7.30) the cases with the largest contributions are those consisting of pairs of the $\varepsilon_t^{+}$'s. Such pairs occur, for instance, if $t_i = t_j$ and $s_i = s_j$ for $i \ne j$ and $i, j \in \{1,2,3,4\}$. In this case and because $E^{+}(\varepsilon_t^{+})^2 = 1$, we get the sum

$$\begin{aligned}
&(E^{+}(\varepsilon_t^{+})^2)^4\frac{1}{n^2m_n^2}\sum_{t_1=1}^{n-1}\sum_{t_2=1}^{n-1}\sum_{s_1=1}^{m_n(1+C)}\sum_{s_2=1}^{m_n(1+C)} k^2\Bigl(\frac{s_1}{m_n}\Bigr)k^2\Bigl(\frac{s_2}{m_n}\Bigr)\cos^2(s_1\lambda)\cos^2(s_2\lambda)\\
&\quad\le \frac{1}{m_n^2}\sum_{s_1=1}^{n}\sum_{s_2=1}^{n} k^2\Bigl(\frac{s_1}{m_n}\Bigr)k^2\Bigl(\frac{s_2}{m_n}\Bigr)\\
&\quad\to \Bigl(\int_0^{\infty} k^2(x)\,dx\Bigr)^2 = O(1).
\end{aligned}$$

□

LEMMA 7.5. *Assume* (A1), (A2) *and* (A4)–(A8) *and let* $p \in \mathbb{N}$ *be fixed. If* $b \sim n^{-1/5}$ *and* $nh^3 \to \infty$ *as* $n \to \infty$, *then*

$$B_n^+(\lambda) \to \frac{1}{4\pi} f''(\lambda) \int u^2 K(u)\, du$$

*in probability.*

PROOF. Using Lemma 7.3 and the fact that $|n^{-1} \sum_{j=1}^n K_b(\lambda - \lambda_j) - 1| = O(n^{-1}b^{-1})$, we get

$$\begin{aligned}
B_n^+(\lambda) &= \sqrt{\frac{b}{n}} \sum_{j=-N}^{N} K_b(\lambda - \lambda_j)\big(\hat{q}(\lambda_j)\hat{f}_{\mathrm{AR}}(\lambda_j) - \hat{q}(\lambda)\hat{f}_{\mathrm{AR}}(\lambda)\big) \\
&\quad + O_P(n^{-1/2}b^{-1/2}) + O_P\big(\sqrt{b}\big) \\
&= \hat{f}_{\mathrm{AR}}(\lambda)\sqrt{\frac{b}{n}} \sum_{j=-N}^{N} K_b(\lambda - \lambda_j)\big(\hat{q}(\lambda_j) - \hat{q}(\lambda)\big) \\
&\quad + \sqrt{\frac{b}{n}} \sum_{j=-N}^{N} K_b(\lambda - \lambda_j)\big(\hat{q}(\lambda_j) - \hat{q}(\lambda)\big)\big(\hat{f}_{\mathrm{AR}}(\lambda_j) - \hat{f}_{\mathrm{AR}}(\lambda)\big) \qquad (7.31) \\
&\quad + \hat{q}(\lambda)\sqrt{\frac{b}{n}} \sum_{j=-N}^{N} K_b(\lambda - \lambda_j)\big(\hat{f}_{\mathrm{AR}}(\lambda_j) - \hat{f}_{\mathrm{AR}}(\lambda)\big) \\
&\quad + O_P(n^{-1/2}b^{-1/2}) + O_P\big(\sqrt{b}\big) \\
&= B_{1,n}^+(\lambda) + B_{2,n}^+(\lambda) + B_{3,n}^+(\lambda) + O_P(n^{-1/2}b^{-1/2}) + O_P\big(\sqrt{b}\big),
\end{aligned}$$

with an obvious notation for $B_{i,n}^+(\lambda)$, $i = 1, 2, 3$. To establish the desired convergence, it suffices to show that

$$B_{1,n}^+(\lambda) \to \frac{1}{4\pi} f_{\mathrm{AR}}(\lambda) q''(\lambda) \int_{-\pi}^{\pi} u^2 K(u)\, du, \tag{7.32}$$

$$B_{2,n}^+(\lambda) \to \frac{1}{2\pi} f'_{\mathrm{AR}}(\lambda) q'(\lambda) \int_{-\pi}^{\pi} u^2 K(u)\, du \tag{7.33}$$

and

$$B_{3,n}^+(\lambda) \to \frac{1}{4\pi} f''_{\mathrm{AR}}(\lambda) q(\lambda) \int_{-\pi}^{\pi} u^2 K(u)\, du \tag{7.34}$$

in probability since $f''(\lambda) = f_{\mathrm{AR}}(\lambda)q''(\lambda) + 2f'_{\mathrm{AR}}(\lambda)q'(\lambda) + f''_{\mathrm{AR}}(\lambda)q(\lambda)$. We proceed by showing that (7.32)–(7.34) are true.

To prove (7.32), note that a Taylor series expansion yields

$$
\begin{aligned}
&\sqrt{\frac{b}{n}}\sum_{j=-N}^{N} K_b(\lambda-\lambda_j)\big(\hat{q}(\lambda_j)-\hat{q}(\lambda)\big)\\
&\quad = n^{-3/2}b^{1/2}\sum_{j=-N}^{N}\sum_{s=-N}^{N} K_b(\lambda-\lambda_j)\\
&\qquad\qquad \times \big[K_h(\lambda_j-\lambda_s)-K_h(\lambda-\lambda_s)\big]\frac{I_n(\lambda_s)}{\hat{f}_{\mathrm{AR}}(\lambda_s)}\\
&\quad = \frac{1}{n^{3/2}b^{1/2}h}\sum_{j=-N}^{N}\sum_{s=-N}^{N} K\Big(\frac{\lambda-\lambda_j}{b}\Big)\frac{\lambda_j-\lambda}{h}\\
&\qquad\qquad \times K'\Big(\frac{\lambda-\lambda_s}{h}\Big)\frac{I_n(\lambda_s)}{\hat{f}_{\mathrm{AR}}(\lambda_s)}\\
&\qquad + \frac{1}{2n^{3/2}b^{1/2}h}\sum_{j=-N}^{N}\sum_{s=-N}^{N} K\Big(\frac{\lambda-\lambda_j}{b}\Big)\Big(\frac{\lambda_j-\lambda}{h}\Big)^2\\
&\qquad\qquad \times K''\Big(\frac{\lambda-\lambda_s}{h}\Big)\frac{I_n(\lambda_s)}{\hat{f}_{\mathrm{AR}}(\lambda_s)}\\
&\qquad + O_P(n^{1/2}b^{7/2}).
\end{aligned}
\tag{7.35}
$$

Now,

$$
\begin{aligned}
&\frac{1}{n^{3/2}b^{1/2}h}\sum_{j=-N}^{N}\sum_{s=-N}^{N} K\Big(\frac{\lambda-\lambda_j}{b}\Big)\frac{\lambda_j-\lambda}{h}K'\Big(\frac{\lambda-\lambda_s}{h}\Big)\frac{I_n(\lambda_s)}{f_{\mathrm{AR}}(\lambda_s)}\\
&\quad = \frac{\sqrt{n}}{2\pi\sqrt{b}}\int_{-\pi}^{\pi}(x-\lambda)K\Big(\frac{\lambda-x}{b}\Big)\,dx\\
&\qquad \times \frac{1}{nh^2}\sum_{s=-N}^{N} K'\Big(\frac{\lambda-\lambda_s}{h}\Big)\frac{I_n(\lambda_s)}{f_{\mathrm{AR}}(\lambda_s)} + o_P(1)\\
&\quad = o_P(1),
\end{aligned}
\tag{7.36}
$$

using $\int uK(u)\,du=0$ and the fact that the second multiplicative term on the right-hand side of the equality before the last one is $O_P(1)$ because it converges to $q'(\lambda)$.

Similarly,

$$\frac{1}{2n^{3/2}b^{1/2}h}\sum_{j=-N}^{N}\sum_{s=-N}^{N}K\Big(\frac{\lambda-\lambda_j}{b}\Big)\Big(\frac{\lambda_j-\lambda}{h}\Big)^2 \times K''\Big(\frac{\lambda-\lambda_s}{h}\Big)\frac{I_n(\lambda_s)}{f_{\mathrm{AR}}(\lambda_s)}$$

$$\begin{aligned}&=\frac{1}{2}\frac{1}{\sqrt{nb}}\sum_{j=-N}^{N}(\lambda_j-\lambda)^2K\Big(\frac{\lambda-\lambda_j}{b}\Big)\\ &\qquad\times\frac{1}{nh^3}\sum_{s=-N}^{N}K''\Big(\frac{\lambda-\lambda_s}{h}\Big)\frac{I_n(\lambda_s)}{f_{\mathrm{AR}}(\lambda_s)}\\ &\to\frac{1}{4\pi}\int_{\pi}^{\pi}u^2K(u)\,du\,q''(\lambda)\end{aligned}\tag{7.37}$$

in probability. Note that the last assertion follows using $b\sim n^{-1/5}$ and because for $nh^3\to\infty$ we have $n^{-1}h^{-3}\sum_{s=-N}^{N}K''((\lambda-\lambda_s)/h)I_n(\lambda_s)/f_{\mathrm{AR}}(\lambda_s)\to q''(\lambda)$ in probability.

Now, by (7.35)–(7.37) and because $\hat{f}_{\mathrm{AR}}(\lambda)=f_{\mathrm{AR}}(\lambda)+O_P(n^{-1/2})$ uniformly in $\lambda$, we obtain (7.32).

Since (7.33) and (7.34) follow using similar arguments, we stress only the essentials.

For $B^+_{2,n}(\lambda)$ we have using the differentiability of $\hat{f}_{\mathrm{AR}}(\lambda)$ with respect to $\lambda$ and similar arguments as in obtaining (7.35) that

$$\begin{aligned}B^+_{2,n}(\lambda)&=\hat{f}'_{\mathrm{AR}}(\lambda)n^{-1/2}b^{1/2}\sum_{j=-N}^{N}K_b(\lambda-\lambda_j)(\lambda_j-\lambda)\big(\hat{q}(\lambda_j)-q(\lambda)\big)\\ &\qquad+o_P(1)\\ &=\hat{f}'_{\mathrm{AR}}(\lambda)\frac{1}{\sqrt{nb}}\sum_{j=-N}^{N}K\Big(\frac{\lambda-\lambda_j}{b}\Big)(\lambda_j-\lambda)^2\\ &\qquad\times\frac{1}{nh^2}\sum_{s=-N}^{N}K'\Big(\frac{\lambda-\lambda_s}{h}\Big)\frac{I_n(\lambda_s)}{\hat{f}_{\mathrm{AR}}(\lambda_s)}+o_P(1)\\ &\to f'_{\mathrm{AR}}(\lambda)q'(\lambda)\frac{1}{2\pi}\int_{-\pi}^{\pi}u^2K(u)\,du\end{aligned}$$

in probability. To show (7.34), we use a Taylor series expansion of $\hat{f}_{\mathrm{AR}}(\lambda_j)$ around

$\hat{f}_{\mathrm{AR}}(\lambda)$ and obtain

$$\begin{aligned}
B_{3,n}^{+}(\lambda) &= \hat{q}(\lambda)\hat{f}_{\mathrm{AR}}'(\lambda)n^{-1/2}b^{1/2}\sum_{j=-N}^{N}K_b(\lambda-\lambda_j)(\lambda_j-\lambda) \\
&\qquad + \frac{1}{2}\hat{q}(\lambda)\hat{f}_{\mathrm{AR}}''(\lambda)n^{-1/2}b^{1/2}\sum_{j=-N}^{N}K_b(\lambda-\lambda_j)(\lambda_j-\lambda)^2 + o_P(1) \\
&\to q(\lambda)f_{\mathrm{AR}}''(\lambda)\frac{1}{4\pi}\int_{-\pi}^{\pi}u^2K(u)\,du. \qquad \square
\end{aligned}$$

PROOF OF THEOREM 5.1. Consider part (ii) of the theorem. By Lemma 8.8 of Bickel and Freedman (1981), we can split the squared Mallows' metric into a variance part $V_n^2(\lambda)$ and a squared bias part $b_n^2(\lambda)$, where

$$\begin{aligned}
V_n^2(\lambda) = d_2^2\Big(&\mathcal{L}\big(\sqrt{nb}\big(\hat{f}(\lambda)-E\hat{f}(\lambda)\big)\big), \\
&\mathcal{L}\big(\sqrt{nb}\big(\hat{f}^*(\lambda)-E^*\hat{f}^*(\lambda)\big) \mid X_1, X_2, \dots, X_n\big)\Big)
\end{aligned}$$

and

$$b_n^2(\lambda) = nb\big|\big(E\hat{f}(\lambda)-f(\lambda)\big)-\big(E^*\hat{f}^*(\lambda)\big)-\tilde{f}(\lambda)\big|^2.$$

By Lemmas 7.4 and 7.5, we then have that $V_n^2(\lambda)\to 0$ and $b_n^2(\lambda)\to 0$ in probability. Part (i) of the theorem follows by the same arguments but by ignoring the bias term. $\square$

**Acknowledgments.** The authors are grateful to an Associate Editor and two referees for constructive comments and suggestions.

## REFERENCES


ANDERSON, T. W. (1971). *The Statistical Analysis of Time Series*. Wiley, New York.

BAUER, H. (1974). *Wahrscheinlichkeitstheorie und Grundzüge der Masstheorie*. de Gruyter, Berlin.

BAXTER, G. (1962). An asymptotic result for the finite predictor. *Math. Scand.* **10** 137–144.

BELTRÃO, K. I. and BLOOMFIELD, P. (1987). Determining the bandwidth of a kernel spectrum estimate. *J. Time Ser. Anal.* **8** 21–38.

BERK, K. N. (1974). Consistent autoregressive spectral estimates. *Ann. Statist.* **2** 489–502.

BICKEL, P. J. and FREEDMAN, D. A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* **9** 1196–1217.

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

BROCKWELL, P. J. and DAVIS, R. A. (1991). *Time Series*: *Theory and Methods*, 2nd ed. Springer, New York.

DAHLHAUS, R. (1985). Asymptotic normality of spectral estimates. *J. Multivariate Anal.* **16** 412–431.

DAHLHAUS, R. and JANAS, D. (1996). A frequency domain bootstrap for ratio statistics in time series analysis. *Ann. Statist.* **24** 1934–1963.

FRANKE, J. and HÄRDLE, W. (1992). On bootstrapping kernel spectral estimates. *Ann. Statist.* **20** 121–145.

HANNAN, E. J. and KAVALIERIS, L. (1986). Regression, autoregression models. *J. Time Ser. Anal.* **7** 27–49.

HURVICH, C. M. and ZEGER, S. L. (1987). Frequency domain bootstrap methods for time series. Technical Report 87–115, Graduate School of Business Administration, New York Univ.

JANAS, D. and DAHLHAUS, R. (1994). A frequency domain bootstrap for time series. In *Computationally Intensive Statistical Methods. Proc. 26th Symposium on the Interface* (J. Sall and A. Lehman, eds.) 423–425. Interface Foundation of North America, Fairfax Station, VA.

KREISS, J.-P. (1999). Residual and wild bootstrap for infinite order autoregression. Unpublished manuscript.

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

NORDGAARD, A. (1992). Resampling stochastic process es using a bootstrap approach. In *Lecture Notes in Econom. and Math. Systems* **376** 181–185.

PAPARODITIS, E. (2002). Frequency domain bootstrap for time series. In *Empirical Process Techniques for Dependent Data* (H. Dehling, T. Mikosch and M. Sørensen, eds.) 365–381. Birkhäuser, Boston.

PAPARODITIS, E. and POLITIS, D. N. (1999). The local bootstrap for the periodogram. *J. Time Ser. Anal.* **20** 193–222.

PRESS, H. and TUKEY, J. W. (1956). Power spectral methods of analysis and their application to problems in airplane dynamics. Bell System Monographs 2606.

PRIESTLEY, M. B. (1981). *Spectral Analysis and Time Series*. Academic Press, New York.

ROBINSON, P. (1991). Automatic frequency domain inference on semiparametric and nonparametric models. *Econometrica* **59** 1329–1363.

ROMANO, J. P. (1988). Bootstrapping the mode. *Ann. Inst. Statist. Math.* **40** 565–586.

ROSENBLATT, M. (1985). *Stationary Sequences and Random Fields*. Birkhäuser, Boston.

SHIBATA, R. (1981). An optimal autoregressive spectral estimate. *Ann. Statist.* **9** 300–306.

SHIRYAYEV, A. N. (1984). *Probability*. Springer, New York.

THEILER, J., PAUL, L. S. and RUBIN, D. M. (1994). Detecting nonlinearity in data with long coherence times. In *Time Series Prediction* (A. Weigend and N. Gershenfeld, eds.). Addison–Wesley, Reading, MA.

INSTITUT FÜR MATHEMATISCHE STOCHASTIK
TECHNISCHE UNIVERSITÄT BRAUNSCHWEIG
POCKELSTRASSE 14
D-38106 BRAUNSCHWEIG
GERMANY

DEPARTMENT OF MATHEMATICS
AND STATISTICS
UNIVERSITY OF CYPRUS
P.O.BOX 20537
CY-1678 NICOSIA
CYPRUS
E-MAIL: stathisp@ucy.ac.cy