Statistica Sinica **28** (2018), 2591-2608
doi:https://doi.org/10.5705/ss.202017.0121

EDGEWORTH EXPANSIONS FOR A CLASS OF SPECTRAL DENSITY ESTIMATORS AND THEIR APPLICATIONS TO INTERVAL ESTIMATION

Arindam Chatterjee and Soumendra N. Lahiri

Indian Statistical Institute and North Carolina State University

Abstract: In this paper we obtain valid Edgeworth expansions (EEs) for a class of spectral density estimators of a stationary time series. The spectral estimators are based on tapered periodograms of overlapping blocks of observations. We give conditions for the validity of a general order EE under an approximate strong mixing condition on the random variables. We use the EE results to study higher order coverage accuracy of confidence intervals (CIs) based on Studentization and on Variance Stabilizing transformation. It is shown that the accuracy of the CIs critically depends on the length of the blocks employed. We use the EE results to determine the optimal orders of the block lengths for one- and two-sided CIs under both methods. Theoretical results are illustrated with a moderately large simulation study.

We dedicate this paper to the memory of Professor Peter Hall who made fundamental contributions to asymptotic theory of Statistics and extensively used EEs to study higher order coverage properties of CIs.

Key words and phrases: Confidence intervals, frequency domain, stationary, studentization, taper, variance stabilizing transformation.

1. Introduction

As many of the seminal works of Peter Hall show, analyses of higher order asymptotic properties of statistical methods are often critical to better understand and improve their finite sample properties. Peter has effectively established this asymptotic point of view to assessing the quality of statistical inference methods through the 600+ papers that he wrote in his exemplary career of nearly four decades, shortened by premature death in 2016. In the spirit of his important work on using EEs to study higher order properties of CIs, in this article, we develop EEs for a class of spectral density estimators of a stationary time series and use it to investigate higher order properties of two most commonly used methods, based on (i) a Variance Stabilizing Transformation and (ii) Studentization, for constructing CIs for the spectral density. We dedicate this article as a tribute to Peter for his vast contributions to Statistics that have influenced more than one generation of statisticians (including both the authors) and will continue to influence the practice of Statistics in years to come.

Spectral densities play an important role in the frequency domain analyses of time series data. Accurate estimation of the spectral density is therefore a central issue for eliciting second order characteristics of a time series from the observed data. This has prompted a large amount of work on consistency and asymptotic normality of spectral estimators. Much less is known about higher order asymptotic properties of different inference methods for the spectral density. A primary reason is the lack of EE results for spectral density estimators in the literature. Indeed, the serial dependence in the time series observations gives rise to some difficult technical challenges and makes the derivation of such expansions very complicated. Here we make use of some recent tools from Lahiri (2007, 2010) on EEs for block-wise functions of time series observations to derive valid EEs for a class of spectral density estimators.

To describe the main results of the paper, let $\{X_t : t \in \mathbb{Z}\}$ be a second order stationary time series with $\mathbf{E}(X_1) = 0$ and spectral density $f(\cdot)$. Let $\mathcal{X}_n = \{X_1, \ldots, X_n\}$ denote the observations from this time series. We construct a class of estimators of $f(\lambda)$ based on blocks of observations from $\mathcal{X}_n$. Let $\mathcal{X}_{i,l} = (X_i, \ldots, X_{i+l-1})$, $i \geqslant 1$ denote overlapping blocks of length l, where $l \equiv l_n \in [1, n]$ is an integer sequence with $l_n \uparrow \infty$ and $n/l \to \infty$ as $n \to \infty$. Note that there are $N = (n - l + 1)$-many blocks $\{\mathcal{X}_{j,l} : j = 1, \ldots, N\}$ contained in $\mathcal{X}_n$. Let $\{h_r, r = 1, \ldots, l\}$ be a data-taper and let

$$d_{j,n}(\lambda) \equiv \left(2\pi \sum_{r=1}^{l} h_r^2\right)^{-1/2} \sum_{r=1}^{l} h_r X_{r+j-1} \exp(\iota\lambda r) \tag{1.1}$$

be the tapered discrete Fourier transform (DFT) for the block $\mathcal{X}_{j,l}$ at frequency λ, $j = 1, \ldots, N$ where $\iota = \sqrt{-1}$. Here we shall consider spectral density estimators of the form

$$\hat{f}_n(\lambda) = \frac{1}{N} \sum_{j=1}^{N} |d_{j,n}(\lambda)|^2, \quad \lambda \in (-\pi, \pi]. \tag{1.2}$$

Estimators of this type have been previously considered by Bartlett (1946, 1950), Welch (1967), Brillinger (1975) and Zhurbenko (1979, 1980). These estimators also have a close connection (cf. Priestley (1981)) with kernel based estimators pioneered by Grenander and Rosenblatt (1957), Blackman and Tukey (1959) and Parzen (1961).

The main results of the paper give valid EEs for the tapered spectral density estimator $\widehat{f}_n(\lambda)$ under a set of regularity conditions on the process $\{X_t\}$. In an important work, Götze and Hipp (1983) derived valid Edgeworth expansions for the sample mean of weakly dependent random variables. Extensions and refinements of their results from the case of a sample mean to normalized sums of block variables are given by Lahiri (2007) and Lahiri (2010), which will be most relevant for deriving EEs for the estimator $\widehat{f}_n(\lambda)$ in (1.2). Many of the commonly used estimators that are employed in time series analysis involving the covariance structure of the process are indeed given by sums of block variables. For example, the normalizing factor in the Studentized sample mean is a block variable. Similarly, Block Bootstrap estimators of population quantities of a time series are functions of sums of block variables. See Lahiri (2007, 2010) for some examples. In the frequency domain, Janas (1994) and Taniguchi, van Garderen and Puri (2003) derived EEs for estimators of weighted integrals of the spectral density and their smooth functions. For the spectral density itself and also for studentized sample mean, Velasco and Robinson (2001) constructed valid EEs for a univariate Gaussian stationary time series. For a detailed account of the higher order theory and results in time series, see Taniguchi and Kakizawa (2000). In this paper, we derive valid EEs for the spectral density estimator (1.2) under the general set-up of Lahiri (2007) that is applicable to non-Gaussian and nonlinear processes. We also derive *simple* sufficient conditions for the validity of the expansions when the time series is a linear process driven by a sequence of independent and identically distributed (iid) random vectors. These sufficient conditions only involve the rate of decay of the co-efficients of the linear process and require a simple smoothness condition, Cramér's condition, on the marginal distribution of the innovations. In particular, the decay condition on the coefficients is satisfied by time series generated by standard time series models, like the autoregressive processes, and the smoothness condition on the innovations hold if the common marginal distribution of the innovation variables has an absolutely continuous distribution with respect to (w.r.t) the Lebesgue measure on the real line.

As applications of the EE results, we investigate higher order properties of the two most commonly used methods for constructing CIs for the spectral density $f(\lambda)$. Let $b \equiv b_n \sim n/l$. It is well-known that under suitable regularity conditions,

$$\sqrt{b}\Big\{\widehat{f}_n(\lambda) - f(\lambda)\Big\} \xrightarrow{d} N(0, c_1^2 f^2(\lambda)),$$

for some *known* constant $c_1 \in (0, \infty)$ (depending only the frequency λ and the taper weights; See Section 3). Thus, a Studentized version of $\widehat{f}_n(\lambda)$ (i.e., the t-statistic for $f(\lambda)$) is given by

$$T_{1,n} = \frac{\sqrt{b}\left\{\widehat{f}_n(\lambda) - f(\lambda)\right\}}{\{c_1\widehat{f}_n(\lambda)\}}$$

which has a $N(0,1)$ limit law. CIs for $f(\lambda)$ can be constructed using the critical points of the limiting Normal distribution. In the literature, a second method is often used for constructing CIs for $f(\lambda)$. Since the limiting variance is a smooth function of $f(\lambda)$, a pivotal quantity can also be constructed using the Variance Stabilizing Transformation (VST) of R.A. Fisher, leading to

$$T_{2,n} = \frac{\sqrt{b}\{\log\widehat{f}_n(\lambda) - \log f(\lambda)\}}{c_1},$$

which also has a $N(0,1)$ limit distribution. Both the Studentized statistic and the VST-based pivot are extensively used for constructing CIs and tests using the critical values from the limiting $N(0,1)$ distribution, but with little information about their accuracy.

In Section 3, we make use of the basic EE results on mean-centered $\widehat{f}_n(\lambda)$ from Section 2 and derive third order EEs for both $T_{1,n}$ and $T_{2,n}$. While the EEs for smooth functions of sample means have terms in powers of $n^{-1/2}$, where n is the sample size, the EEs for $T_{1,n}$ and $T_{2,n}$ are given by superimpositions of two series, one in powers of $b^{-1/2}$ and the other in powers of a scaled version of the bias of $\widehat{f}_n(\lambda)$. The error in the Normal approximation is influenced by the leading terms from both of these series. The EE results allow us to quantify the exact orders of approximations for these traditional approaches to constructing the CIs. The dominant contributor to the error in both cases comes from the scaled bias of the spectral density estimator itself. However, a significant amount of bias also results from the naive approximation in $T_{1,n}$ to the finite sample variance of the spectral density. Although the VST does not involve the Studentization step explicitly, the same variance estimation bias also affects the rate of Normal approximation at a comparable level. See Section 3 for more details.

In Section 4, we derive expansions for the coverage probabilities of one- and two-sided CIs based on $T_{k,n}, k = 1, 2$. Here we also determine the rates of the optimal block sizes that lead to optimal performance of the resulting CIs. Specifically, we show that for the one-sided CIs based on $T_{k,n}, k = 1, 2$, the optimal block sizes are of the order $const. n^{1/2}$ for the non-tapered case, giving an overall accuracy of $O(n^{-1/4})$ in the coverage probabilities. Here and in the following, *const* refers to a constant with values in $(0, \infty)$. For the tapered case, the optimal rate of convergence improves to $O(n^{-1/3})$ with optimal block size of $l \sim$ *const.* $n^{1/3}$. For the two-sided CIs based on both $T_{1,n}$ and $T_{2,n}$, the corresponding optimal block sizes are of the order *const.* $n^{3/5}$ in the non-tapered case and the best possible coverage error of $O(n^{-2/5})$. With tapering, this rate improves to $O(n^{-1/2})$ for $l \sim$ *const.* $n^{1/2}$. Further, the VST and the Studentization-based methods have very similar performance in terms of the best possible error rates, although the optimal block sizes for each method has a different constant multiplier to the rates. Simulation results show that block sizes that are suitable multiples of the optimal order for the error rates give reasonably good performance in finite samples.

The rest of the paper is organized as follows. Section 2 gives an outline of the regularity conditions used for deriving the EEs and provides simple sufficient conditions for linear processes. It also gives the results on the EE for spectral density estimators. In Section 3, EE results for Studentized version of $\widehat{f}_n(\lambda)$ and the VST-based pivots are described. These are then used in Section 4 to study coverage accuracy of one- and two-sided CIs based on these two methods and to determine the optimal block size in each case. Section 5 gives the results from a simulation study. Proofs of the results and the general framework for deriving EEs and exact statements of the regularity conditions that allow nonlinear and non-Gaussian processes are given in the Supplementary Materials (cf. Chatterjee and Lahiri (2018)).

2. Edgeworth Expansion Theory for the Spectral Density Estimator

To derive the EEs for the spectral density estimator for general (non-Gaussian) time series, we adopt a framework similar to Lahiri (2007) for sums of 'block' variables, which is an extension of Götze and Hipp (1983)'s framework for sums of weakly dependent random variables. We give an outline of the regularity conditions in Section 2.1. In Section 2.2, we develop the "formal" EE for the centered and scaled spectral density estimator $\widehat{f}_n(\lambda)$, where we identify the form of the higher order terms. Validity of the EEs (i.e., establishing the order of the error) is proved in Section 2.3.

2.1. Regularity conditions

Note that $\widehat{f}_n(\lambda)$ is a sum of 'block' variables $|d_{j,n}(\lambda)|^2$, $1 \leqslant j \leqslant N$ which are quadratic functions of the X_ts in the respective blocks $\mathcal{X}_{j,l}$, $1 \leqslant j \leqslant N$. Since the blocks are overlapping, the dependence among neighboring block variables can be very strong. This typically destroys the factorization property of the characteristic function of a sum of independent variables. As a result, for establishing validity of the EEs, specially for non-Gaussian time series, the framework of Götze and Hipp (1983) and Lahiri (2007) makes use of regularity conditions in terms of a suitable collection of auxiliary σ-fields to develop some nonstandard factorization arguments. To simplify the exposition, exact statements of the conditions are relegated to the Appendix. Roughly speaking, these conditions require that (i) the process $\{X_t\}$ satisfies an approximate strong mixing condition, (ii) some suitable moment conditions, depending on the order of the EE, hold and (iii) a conditional Cramér's condition holds. See Conditions (C.1) – (C.6) in the Appendix for more details.

We shall now state a set of simple sufficient conditions for these regularity conditions when the X_t-process is given by a linear process with the representation,

$$X_t = \sum_{k \in \mathbb{Z}} a_k \epsilon_{t-k}, \tag{2.1}$$

where $\{\epsilon_k\}_{k\in\mathbb{Z}}$ is a collection of iid random variables with $\mathbf{E}(\epsilon_1) = 0$ and $\mathbf{E}(\epsilon_1^2) = 1$, and where $\{a_k : k \in \mathbb{Z}\} \subset \mathbb{R}$. While the general conditions stated in the Appendix allow for general taper weights $\{h_r\}$, here we shall also restrict our attention to taper weights generated from a smooth taper function $h : [0,1] \to \mathbb{R}$ as

$$h_r = h\left(\frac{r}{l}\right), \quad r = 1, \ldots, l, \quad l \geqslant 1. \tag{2.2}$$

With this, we have the following result on regularity conditions for linear processes.

Proposition 1. *Suppose $\{X_t\}$ is a linear process given by (2.1) and the taper weights $\{h_r : r = 1, \ldots, l\}$ are given by (2.2) for some function h that is continuously differentiable on $(0,1)$, with $\int_0^1 h^2(x)\,dx \in (0,\infty)$. Suppose that, for some $\kappa \in (0,1)$ and an integer $s \geqslant 3$, $\mathbf{E}|\epsilon_1|^{2(s+1)+\kappa} < \infty$, $\kappa^{-1}\log n < l < \kappa^{-1} n^{(1-\kappa)/3}$, and that $(\epsilon_1, \epsilon_1^2)$ satisfies Cramér's condition :*

$$\limsup_{\max\{|t|,|s|\}\to\infty} \left|\mathbf{E}\exp(\iota[s\epsilon_1 + t\epsilon_1^2])\right| < 1. \tag{2.3}$$

If,

$$|a_k| = O\big(c_1^{|k|}\big) \quad \text{as } |k| \to \infty, \tag{2.4}$$

for some constant $0 < c_1 < 1$, and if, $f(\lambda)\int_0^1 \big\{\int_0^{1-x} h(y)h(x+y)\,dy\big\}^2 dx > 0$, then the regularity conditions (C.1) – (C.6) of the Appendix hold.

Thus, for the linear process in (2.1), the EE results of this paper remain valid under some simple sufficient conditions on $\{h_r\}$, $\{a_k\}$ and the joint distribution of $(\epsilon_1, \epsilon_1^2)$. The condition (2.4) on the coefficients of the linear process (2.1) is satisfied for common time series models, such as the autoregressive moving average (ARMA) processes, driven by iid error variables $\{\epsilon_i\}$. Further, Cramér's condition (2.3) holds whenever the marginal distribution of ϵ_1 has an absolutely continuous component (w.r.t. the Lebesgue measure on the real line).

2.2. Development of the "formal" Edgeworth expansion

Let $T_n \equiv \sqrt{b}\big(\hat{f}_n(\lambda) - \mathbf{E}\hat{f}_n(\lambda)\big)$ denote the centered and scaled spectral density estimator $\hat{f}_n(\lambda)$, where $b \equiv b_n = N/l$ and $N = n - l + 1$. The formal EE for T_n is obtained by first deriving a suitable expansion of the characteristic function of T_n in terms of its cumulants, giving a function $\hat{\psi}_{s,n}(t)$, $t \in \mathbb{R}^d$ where s determines the order of the EE. In the next step, the Fourier inversion formula is applied to $\hat{\psi}_{s,n}(t)$ to generate a signed measure on $\mathbb{R}$ with a (Lebesgue) density $\psi_{s,n}$, which yields the *density of the* $(s-2)$*th order EE* of T_n. Let $\iota = \sqrt{-1}$ and let $\chi_{r,n}(t)$ is the r-th cumulant of $t'T_n$,

$$\iota^r \chi_{r,n}(t) = \frac{d^r}{du^r} \log \mathbf{E} \exp(\iota u[tT_n])\bigg|_{u=0}.$$

We define the polynomials $\tilde{p}_{r,n}(t)$ for $t \in \mathbb{R}$ by the identity (in $u \in \mathbb{R}$)

$$\exp\left(\sum_{r=3}^{s} (r!)^{-1} u^{r-2} b^{(r-2)/2} \chi_{r,n}(t)\right) = 1 + \sum_{r=1}^{\infty} u^r \tilde{p}_{r,n}(t). \tag{2.5}$$

Then, the function $\hat{\psi}_{s,n}(t)$ is defined by

$$\hat{\psi}_{s,n}(t) = \exp\left(-\frac{\chi_{2,n}(t)}{2}\right)\left\{1 + \sum_{r=1}^{s-2} b^{-r/2} \tilde{p}_{r,n}(\iota t)\right\}, \quad t \in \mathbb{R}.$$

Then, the density of the $(s-2)$-th order Edgeworth expansion $\psi_{s,n}(x)$ of $T_n(\lambda)$ is defined by inverting its Fourier transform through the relation $\psi_{s,n}(x) = (2\pi)^{-1} \int e^{-\iota t x} \hat{\psi}_{s,n}(t) dt$, $x \in \mathbb{R}$.

It can be shown (cf. Lahiri (2007, Lemma 4.1)) that under the regularity conditions of the paper, the rth order cumulant $\chi_{r,n}(t)$ is $O\big(b^{-(r-2)/2}\big)$ for any fixed $t \in \mathbb{R}$, for all $2 \leqslant r \leqslant s$. Hence, the co-efficients of the polynomials $\tilde{p}_{r,n}(t)$, $2 \leqslant r \leqslant s$ are $O(1)$ as $n \to \infty$. This shows that the density of the $(s-2)$th order EE for T_n is given by adding terms of the order $O(b^{-r/2})$ for $r = 1, \ldots, s-2$ to a normal density function. In contrast to the case of the sample mean, where the successive higher order terms in the EE are of the order $n^{-1/2}, n^{-1}, \ldots$, the EEs for T_n are obtained by adding terms in powers of $b^{-1/2}$. Thus, statements about second, third, ... order properties here refer to those of the terms of orders $b^{-1/2}$, b^{-1}, etc.

To give some insight into the nature of the terms of the EE and for future reference, we state the third order EE explicitly. Let $\mu_{r,n} = \mathbf{E}(T_n^r)$. Since $\mathbf{E}(T_n) = \mu_{1,n} = 0$, it is easy to check that $\chi_{r,n}(t) = t^r \mu_{r,n}$ for all $r = 1, 2, 3$ and $\chi_{4,n}(t) = \left(\mu_{4,n} - \mu_{2,n}^2\right) t^4$, $t \in \mathbb{R}$. From (2.5), we have $\tilde{p}_{1,n}(t) = b^{1/2}\chi_{3,n}(t)/6$ and $\tilde{p}_{2,n}(t) = b\chi_{4,n}(t)/24 + b\chi_{3,n}^2(t)/72$ which yields

$$\hat{\psi}_{4,n}(t) = \exp\left(\frac{-t^2\mu_{2,n}}{2}\right)\left\{1 + \frac{(\iota t)^3}{6}\mu_{3,n} + \frac{(\iota t)^4}{24}(\mu_{4,n} - \mu_{2,n}^2) + \frac{(\iota t)^6}{72}\mu_{3,n}^2\right\}.$$

The density $\psi_{s,n}(x)$ can now be found using the inversion formula and the identity, for $k = 1, 2, \ldots$,

$$(-1)^k\left\{\sigma^{-k} H_k(\sigma^{-1}x)\right\}\phi_\sigma(x) = (2\pi)^{-1}\int_{\mathbb{R}} \exp(-\iota t x)(\iota t)^k \hat{\phi}_\sigma(t)dt, \;\; x \in \mathbb{R},$$

where $\phi_\sigma(x) = \left(2\pi\sigma^2\right)^{-1/2}\exp(-x^2/[2\sigma^2])$, $x \in \mathbb{R}$ and $\hat{\phi}_\sigma(t) = \exp(-t^2\sigma^2/2)$, $t \in \mathbb{R}$ are the probability density function and the characteristic function of the $N(0, \sigma^2)$ distribution, $\sigma \in (0, \infty)$, and where $H_k(x)$ is the kth order Hermite polynomial. Thus,

$$\begin{aligned}&\psi_{4,n}(x)\\ &= \phi_{\sigma_n}(x)\left[1 + \frac{\mu_{3,n}}{6\sigma_n^3}H_3(\sigma_n^{-1}x) + \left\{\frac{(\mu_{4,n} - \mu_{2,n}^2)}{24\sigma_n^4}H_4(\sigma_n^{-1}x) + \frac{\mu_{3,n}^2}{72\sigma_n^6}H_6(\sigma_n^{-1}x)\right\}\right],\end{aligned}$$

where $\sigma_n^2 = \mu_{2,n} = \mathrm{Var}(T_n)$. Under mild conditions, $\mu_{3,n} = O(b^{-1/2})$ and $\mu_{4,n} - \mu_{2,n}^2 = O(b^{-1})$, so that the second term of $\psi_{4,n}(x)$ is $O(b^{-1/2})$ and the third term (within $\{\cdot\}$) is $O(b^{-1})$.

2.3. Main results

We now state the EE results for T_n under the conditions of Section S7.1. To that end, for an integer $s \geqslant 3$, let $s_0 = 2\lfloor s/2 \rfloor$. For a Borel measurable function $f : \mathbb{R} \to \mathbb{R}$ and $\epsilon > 0$, define its integrated modulus of continuity by

$$\omega(f : \epsilon) = \int \sup\Big(|f(x + y) - f(x)| : |y| \leqslant \epsilon\Big)\phi_{\sigma_\infty^2}(x)dx,$$

where σ_∞^2 is as in condition (C.2).

Theorem 1. *Let conditions* (C.1) – (C.6) *hold for some* $a \in ((s-2)/2, \infty)$ *where* $s \geqslant 3$ *is an integer. Let* $f : \mathbb{R} \to \mathbb{R}$ *be a Borel measurable function with* $M_f \equiv \sup\left\{(1 + |x|^{s_0})^{-1}|f(x)| : x \in \mathbb{R}\right\} < \infty$. *Then, there exist constants* $C_1 = C_1(a)$, $C_2 \in (0, \infty)$ *(neither depending on* f*) such that*

$$\left|\mathbf{E}f(T_n) - \int f(x)\psi_{s,n}(x)dx\right| \leqslant C_1\omega(\tilde{f} : b^{-a}) + C_2 M_f b^{-(s-2)/2}(\log n)^{-2}$$

for all $n > C_2$, *where* $\tilde{f}(x) = f(x)(1 + |x|^{s_0})^{-1}$, $x \in \mathbb{R}$.

Corollary 1. *Under the conditions of Theorem* 1,

$$\sup_{u\in\mathbb{R}}\left|\mathbf{P}\Big(T_n \leqslant u\Big) - \int_{-\infty}^{u}\psi_{s,n}(x)dx\right| = O\left(b^{-(s-2)/2}(\log n)^{-2}\right).$$

We use Corollary 1 to derive an EE for the spectral density estimator $\widehat{f}_n(\lambda)$ centered at the true spectral density $f(\lambda)$. To that end, set $B_n = \sqrt{b}\big(\mathbf{E}\widehat{f}_n(\lambda) - f(\lambda)\big)$, and let $g^{(r)}$ denote the rth derivative of a function $g : \mathbb{R} \to \mathbb{R}$.

Corollary 2. *If the conditions of Theorem* 1 *hold and* $B_n = O(b^{-\delta})$ *for some* $\delta \in (0, \infty)$, *then*

$$\sup_{u\in\mathbb{R}}\left|\mathbf{P}\Big(\sqrt{b}\big(\widehat{f}_n(\lambda) - f(\lambda)\big) \leqslant u\Big) - \int_{-\infty}^{u}\tilde{\psi}_{s,n}(x)dx\right| = O\left(b^{-(s-2)/2}(\log n)^{-2}\right), \quad (2.6)$$

where $\tilde{\psi}_{s,n}(x) = \sum_{r=0}^{K-1} B_n^r \psi_{s,n}^{(r)}(x)/r!$ *and* $K = \min\{k \geqslant 1 : k\delta > (s-2)\}$.

Corollary 2 shows that the EE of $\sqrt{b}\big(\widehat{f}_n(\lambda) - f(\lambda)\big)$ is a super-imposition of two series, one in powers of $b^{-1/2}$ (determined by the variance of $\widehat{f}_n(\lambda)$) and the other in powers of B_n. Thus, the overall accuracy of Normal approximation to the distribution of $\widehat{f}_n(\lambda)$ depends on both the bias and the variance of the spectral density estimator. The order of B_n typically depends on the choice of the taper and the smoothness of $f(\lambda)$ (cf. Section 4) while b is determined by the choice of the block size. As a result, the choice of the taper and the block size must be done carefully to improve the performance of the Normal critical points based CIs for the spectral density.

In the next section, we use the results from Corollary 1 and 2 to derive EEs for the Studentized statistic $T_{1,n}$ and the VST-based pivotal quantity $T_{2,n}$.

3. Edgeworth Expansions for the Asymptotic Pivots

Under the regularity conditions of Theorem 1, it can be shown that if $B_n = o(b^{-1/2})$, then $\sqrt{b}\{\widehat{f}_n(\lambda) - f(\lambda)\} \xrightarrow{d} N(0, \sigma_\infty^2(\lambda))$ where $\sigma_\infty^2(\lambda) = c^2 \cdot f^2(\lambda)\{1 + \eta(2\lambda)\}$, with $c^2 = 2\{\int_0^1 h^2(x)\, dx\}^{-2} \times \int_0^1\{\int_0^{1-x} h(y)h(x+y)\, dy\}^2\, dx$, and $\eta(\omega) = 1$ or 0 according as $\omega = 0 \pmod{2\pi}$ or not. Thus, the Studentized version of $\widehat{f}_n(\lambda)$ is

$$T_{1,n} = \frac{\sqrt{b}\left\{\widehat{f}_n(\lambda) - f(\lambda)\right\}}{\widehat{\sigma}_n} \quad (3.1)$$

where, with $c_1^2 \equiv c_1^2(\lambda; h) = c^2\{1 + \eta(2\lambda)\}$, the Studentizing factor $\widehat{\sigma}_n$ is given by $\widehat{\sigma}_n = c_1\widehat{f}(\lambda)$. Under the above conditions, $T_{1,n} \xrightarrow{d} N(0,1)$ as $n \to \infty$ and therefore, the limit distribution of $T_{1,n}$ is free of unknown parameters, making $T_{1,n}$ an (asymptotically) pivotal quantity. Further, given the form of the asymptotic variance of $T_{1,n}$, a different pivotal quantity can also be constructed using the variance Stabilizing Transformation (VST) of R.A. Fisher which, in this case, is given by a logarithmic transformation. The resulting (asymptotically) pivotal quantity is

$$T_{2,n} = \frac{\sqrt{b}\left\{\log \widehat{f}_n(\lambda) - \log f(\lambda)\right\}}{c_1}, \tag{3.2}$$

which also has a $N(0,1)$ limit distribution.

We now refine the asymptotic normality results by deriving EEs for both $T_{1,n}$ and $T_{2,n}$. The first result of this section gives a third-order EE for $T_{1,n}$. To that end, we take

$$\left.\begin{aligned}
&B_{1,n} = b^{1/2}\sigma_n^{-1}\left\{\mathbf{E}\widehat{f}_n(\lambda) - f(\lambda)\right\}, \quad B_{2,n} = c_1\sigma_n^{-1}\mathbf{E}\widehat{f}_n(\lambda) - 1,\\
&a_{0,n} = B_{1,n}\left(1 - B_{2,n}\right), \quad a_{1,n} = 1 - B_{2,n} + B_{2,n}^2 - b^{-1/2}B_{1,n},\\
&a_{2,n} = b^{-1/2}\left(2B_{2,n} - 1\right), \quad a_{3,n} = b^{-1}, \quad \text{and}\\
&\tilde{a}_{j,n} = \frac{a_{j,n}}{a_{1,n}}, \quad \text{for } j = 2, 3.
\end{aligned}\right\} \tag{3.3}$$

Theorem 2. *Under the conditions of Theorem* 1,

$$\mathbf{P}\big(T_{1,n} \leqslant u\big) = \Phi(u_n) + q_{1,n}(u_n)\,\phi(u_n) + q_{2,n}(u_n)\,\phi(u_n) + o\left(b^{-1}\right), \tag{3.4}$$

uniformly in $u \in \mathbb{R}$, *where* $u_n = a_{1,n}^{-1}(u - a_{0,n})$ *and*

$$\left.\begin{aligned}
q_{1,n}(u) &= \frac{\kappa_{3,n}}{6}H_2(u) - \tilde{a}_{2,n}u^2, \quad \textit{and}\\
q_{2,n}(u) &= -\big\{b_{2,n}H_1(u) + b_{4,n}H_3(u) + b_{6,n}H_5(u)\big\}.
\end{aligned}\right\} \tag{3.5}$$

The constants $b_{j,n}$*'s are defined as,* $b_{2,n} = 3\tilde{a}_{3,n} + 3\tilde{a}_{2,n}^2/2 + \tilde{a}_{2,n}\kappa_{3,n}$, $b_{4,n} = \kappa_{4,n}/24 + \tilde{a}_{3,n} + 3\tilde{a}_{2,n}^2 + (7\tilde{a}_{2,n}\kappa_{3,n})/6$ *and* $b_{6,n} = \kappa_{3,n}^2/72 + \tilde{a}_{2,n}^2/2 + (\tilde{a}_{2,n}\kappa_{3,n})/6$, *with* $\kappa_{r,n} = r^{th}$ *cumulant of* T_n/σ_n.

The second and third order terms in the EE for $T_{1,n}$ depend on $B_{1,n}$, $B_{2,n}$ as well as on powers of $b^{-1/2}$. Here $B_{1,n}$ is the bias of $\widehat{f}_n(\lambda)$ as an estimator of the true spectral density $f(\lambda)$ scaled by $\sqrt{b}$, and is often the dominant term in the EE. The second factor $B_{2,n}$ is the bias resulting from approximating the true standard deviation σ_n of $\widehat{f}_n(\lambda)$ with the plug-in estimator based on the asymptotic standard deviation $c_1 f(\lambda)$. The leading terms in both $B_{1,n}$ and $B_{2,n}$ are decreasing functions of the block size. On the other hand, a choice of a larger block size necessarily makes $b^{-1/2}$ larger. Thus, the overall accuracy of the

Normal approximation depends on the choice of l through an intricate interaction among these three quantities.

A similar scenario results when using the VST-based pivotal quantity, as follows from the theorem below. To state the EE result for $T_{2,n}$, let $B_{3,n} = \sigma_n^{-1} c_1 f(\lambda) - 1$, and define the constants

$$a_{0,n}^{\dagger} = \frac{B_{1,n}}{1+B_{3,n}} - \frac{c_1 B_{1,n}^2}{2\sqrt{b}(1+B_{3,n})^2},\ a_{1,n}^{\dagger} = \frac{1}{1+B_{3,n}} - \frac{c_1 B_{1,n}}{\sqrt{b}(1+B_{3,n})^2},$$
$$a_{2,n}^{\dagger} = -\frac{c_1}{2\sqrt{b}(1+B_{3,n})^2},\ a_{3,n}^{\dagger} = \frac{c_1^2}{3b}.$$

Theorem 3. *Under the conditions of Theorem* 1*,*

$$\mathbf{P}\big(T_{2,n} \leqslant v\big) = \Phi(v_n) + q_{1,n}^{\dagger}(v_n)\phi(v_n) + q_{2,n}^{\dagger}(v_n)\phi(v_n) + o(b^{-1})$$

uniformly in $v \in \mathbb{R}$*, where* $v_n = (v - a_{0,n}^{\dagger})/a_{1,n}^{\dagger}$ *and where* $q_{k,n}^{\dagger}$ *is defined by replacing* $\{a_{j,n} : j = 0, \ldots, 3\}$ *respectively with* $\{a_{j,n}^{\dagger} : j = 0, \ldots, 3\}$*, for* $k = 1, 2$*.*

4. Accuracy of One- and Two-sided CIs

We now use the EEs from the last section to develop expansions for the coverage probabilities of one- and two-sided CIs for $f(\lambda)$. Let z_α denote the α-quantile of the $N(0,1)$ distribution, $\alpha \in (0,1)$. Then, a large sample $(1-\alpha)100\%$ one-sided lower CI for $f(\lambda)$ based on $T_{k,n}$ is given by

$$I_{k,n} = [L_{k,n}, \infty), \tag{4.1}$$

$k = 1, 2$, where $L_{1,n} = \hat{f}_n(\lambda) - z_{1-\alpha}(c_1 \hat{f}_n(\lambda))/\sqrt{b}$ and $L_{2,n} = \exp(\log \hat{f}_n(\lambda) - z_{1-\alpha}(c_1/\sqrt{b}))$. It is clear that

$$\mathbf{P}\Big(f(\lambda) \in I_{k,n}\Big) = \mathbf{P}(T_{k,n} \leqslant z_{1-\alpha}) \to 1-\alpha \quad \text{as } n \to \infty, \quad k = 1, 2,$$

thus attaining the nominal level $(1-\alpha)$ in the limit. Next we define the two sided CIs. A large sample $(1-\alpha)100\%$ two-sided CI for $f(\lambda)$ based on $T_{1,n}$ is given by

$$J_{1,n} = \left\{ \hat{f}_n(\lambda) - z_{1-\alpha/2}\frac{c_1 \hat{f}_n(\lambda)}{\sqrt{b}},\ \hat{f}_n(\lambda) - z_{\alpha/2}\frac{c_1 \hat{f}_n(\lambda)}{\sqrt{b}} \right\}. \tag{4.2}$$

Similarly, a large sample $(1-\alpha)100\%$ two-sided CI for $f(\lambda)$ based on the VST is given by

$$J_{2,n} = \left\{ \exp\left(\hat{f}_n(\lambda) - z_{1-\alpha/2}\frac{c_1}{\sqrt{b}}\right),\ \exp\left(\hat{f}_n(\lambda) - z_{\alpha/2}\frac{c_1}{\sqrt{b}}\right) \right\}. \tag{4.3}$$

Expansions for the coverage probabilities of $I_{k,n}$ and $J_{k,n}$ are given below.

4.1. Coverage accuracy of one-sided CIs

Let $q_{j,n}^{[1]}(x) = q_{j,n}(x)$ and $q_{j,n}^{[2]}(x) = q_{j,n}^{\dagger}(x)$, $x \in \mathbb{R}$ for $j \in \{1, 2\}$.

Theorem 4. *Under the conditions of Theorem* 1*, for* $k = 1, 2$*,*

$$\mathbf{P}\Big(f(\lambda) \in I_{k,n}\Big) = (1-\alpha) + \Big[\Big\{B_{k+1,n} z_{1-\alpha} - B_{1,n} + q_{1,n}^{[k]}(z_{1-\alpha})\Big\}\phi(z_{1-\alpha})\Big]\{1 + o(1)\}.$$

The errors in coverage accuracy of both the Studentization-based and the VST-based one-sided CIs depend on the magnitude of all three terms in the expansion. It can be shown that under some mild conditions on $h(\cdot)$ (cf. Dahlhaus (1985)) and under conditions (C.1) – (C.5),

$$B_{1,n} = A_1 \sqrt{b} l^{-r}\{1 + o(1)\}, \tag{4.4}$$

where $A_1 = A_1(h)$ is a constant depending on the taper function $h(\cdot)$ and where $r = 1$ for the untapered case and $r = 2$ for $h \in \mathcal{H}_2$. Here $\mathcal{H}_2$ denotes the class of taper functions $g(\cdot)$ that are symmetric around 1/2 with $g(0+) = 0$ and that have a non-vanishing continuous second derivative $g''(\cdot)$. Similarly, if the fourth order cumulant density of the X_t-process satisfies $f_4(\lambda) \neq 0$, then

$$B_{2,n} = A_2 l^{-1}\{1 + o(1)\}. \tag{4.5}$$

Using (4.4) and (4.5), we can determine the exact error rates for the CIs $I_{k,n}$. Here and in the following, let C to denote a generic constant in $(0, \infty)$.

First, we consider the untapered one-sided CI based on $T_{1,n}$. Using the approximations in (4.4) and (4.5), it is not difficult to verify that for the CI $I_{1,n}$ based on Studentization, the optimal order of the error term is obtained when the second and the third terms are of the same order. Equating the two terms, *the optimal order of the block size* l is given by $l \sim Cn^{1/2}$ and the error rate corresponding to this choice of l is given by

$$O(b^{-1/2}) = O(n^{-1/4}).$$

Next, we consider the case of one-sided CIs with a smooth taper $h \in \mathcal{H}_2$. In this case, the orders of $B_{1,n}$, $B_{2,n}$ and of $q_{1,n}(z_{1-\alpha})$ are identical for $n^{1/3}$, while for $l \ll n^{1/3}$, $B_{2,n}$ tends to be of a larger order than $O(n^{1/3})$ and for $l \gg n^{1/3}$, $q_{1,n}(z_{1-\alpha})$ grows at a faster rate than $n^{1/3}$. Thus, *the optimal order of the block size* l is given by $l \sim Cn^{1/3}$ and the optimal error rate is given by

$$O(b^{-1/2}) = O(n^{-1/3}).$$

Thus, the use of taper reduces the error approximation for the one-sided CIs based on Studentization, but the block size must be of a much smaller size compared to the nontapered case. It can be shown that the VST-based one-sided CIs also have the same rates of the optimal errors of approximations for the tapered and the nontapered cases, with respective block sizes $l \sim Cn^{1/3}$ and $l \sim Cn^{1/2}$ for some different constant(s) $C \in (0, \infty)$, determined by the terms of the EE given by Theorem 4. Thus the performance of the Studentization-based and the VST-based CIs are expected to be "similar" when these are based on block sizes of the optimal orders.

4.2. Coverage accuracy of two-sided CIs

Next consider the two-sided CIs. Using Theorems 2 and 3, one can derive expansions for the coverage probabilities of the two CIs $J_{1,n}$ and $J_{2,n}$. Essentially, this requires considering the difference between the expansions for the probabilities $\mathbf{P}(T_{k,n} \leqslant -z_{1-\alpha/2}) - \mathbf{P}(T_{k,n} \leqslant z_{\alpha/2})$. Since $z_{\alpha/2} = -z_{1-\alpha/2}$, the odd/even function property of the terms leads to cancellation of some of the terms. We have expansions in the two-sided case.

Theorem 5. *Under the conditions of Theorem* 1*, for* $k = 1, 2$*,*

$$\begin{aligned}\mathbf{P}(f(\lambda) \in J_{k,n}) = (1-\alpha) + \phi(z_{1-\alpha/2})\Big[&2B_{k+1,n} z_{1-\alpha/2} \\ &- 2B_{1,n}\Big\{(q_{1,n}^{[k]})'(z_{1-\alpha/2}) + q_{1,n}^{[k]}(z_{1-\alpha/2}) z_{1-\alpha/2}\Big\} + q_{2,n}^{[k]}(z_{1-\alpha/2})\Big] \\ &\times \{1 + o(1)\}.\end{aligned}$$

We can use Theorem 5 to determine the optimal orders of the block sizes that minimize the coverage errors of the two-sided CIs. From the definitions of the higher order terms in the EEs, it follows again that the optimal orders are the same for the Studentization-based and the VST-based two-sided CIs. Hence, we only describe the details for the former. We also suppose that (4.4) and (4.5) hold. Then, in the case of no-tapering, the order of $B_{1,n}$ dominates $B_{2,n}$ and therefore, the optimal order is obtained when the order of $B_{1,n}$ matches that of $q_{2,n}(z_{1-\alpha/2})$. Since $q_{2,n}(z_{1-\alpha/2}) = O(b^{-1})$, using (4.4), it follows that *the optimal order of the block size* l is given by $l \sim Cn^{3/5}$ for some $C \in (0, \infty)$, leading to an

$$\text{optimal error rate} = O(n^{-2/5}).$$

When tapering is used with a $h \in \mathcal{H}_2$ and (4.4) holds with $r = 2$, $|B_{2,n}| \gg |B_{1,n}|$ for $l \gg n^{1/3}$. Thus, for large values of l, the $B_{1,n}$ term no longer determines the error rate. One must balance the orders of the $B_{2,n}$-term and the last term $q_{2,n}(z_{1-\alpha/2})$. As a result, *the optimal order of the block size* l is given by $l = Cn^{1/2}$, and with this choice one has

$$\text{optimal error rate} = O(n^{-1/2}).$$

Thus, the two-sided case has a better accuracy compared to the one-sided case. It is a relatively easy task to determine the constants in the optimal block sizes for each of the CIs using the explicit expansions given by Theorems 4 and 5. We omit the routine details to save space.

5. Simulation Results

In this simulation study, we studied the empirical coverage accuracy of one-sided and two-sided CI's obtained by using the studentized pivot $T_{1,n}$ (cf. (3.1)) and the VST-based pivot $T_{2,n}$ (cf. (3.2)). The target parameter was $f(0.5)$, where f is the underlying spectral density. One-sided CI's are then provided in (4.1) by using $k = 1, 2$, and two-sided CI's are provided in (4.2) and (4.3). For the tapered case, we used the Tukey-Hanning taper with $h(u) = \{1 - \cos(2\pi u)\}/2$, $u \in [0, 1]$. Three different stationary time series models were used for generating the samples:

(i) ARMA(3,2) model : $X_t - 0.3740X_{t-1} + 0.0143X_{t-2} + 0.0833X_{t-3} = \epsilon_t - 0.25\epsilon_{t-2}$.

(ii) AR(2) model: $X_t - 0.924X_{t-1} + 1/4X_{t-2} = \epsilon_t$.

(iii) A linear process model: $X_t = \sum_{j\in\mathbb{Z}} a_j\epsilon_{t-j}$, where $a_j = 0$, if $|j| > 30$, $a_j = 5|j|^{-2.1}$, if $-30 \leqslant j < 0$, $a_j = 1$, if $j = 0$ and $a_j = 10|j|^{-2.1}$, if $0 < j \leqslant 30$.

For each of these models, the innovations $\{\epsilon_t : t \in \mathbb{Z}\}$ were i.i.d. observations from a distribution F, with $\mathbf{E}(\epsilon_1) = 0$ and $\mathbf{E}(\epsilon_1^2) = 1$, leading to the spectral densities plotted in Figure 1. Two different choices of F were used: (a) $F_1 \stackrel{d}{=} (\chi_1^2 - 1)/\sqrt{2}$, to study the effect of highly skewed errors, and (b) $F_2 \stackrel{d}{=} t_3/\sqrt{3}$, to study the effect of a relatively fat-tailed errors. Here, χ_ν^2 and t_ν denote chi-square and Students-t distributions with ν degrees of freedom. The sample size was fixed at $n = 500$. Block lengths (l) for tapered or non-tapered cases and one or two-sided CI's were chosen according the optimum rates provided in Sections 4.1 and 4.2. Accordingly, we set $l = an^\beta$, where $a > 0$ and β is the optimum rate for that particular type of CI. The results for one and two-sided CI's are presented in Tables 1 and 2, respectively.

From Table 1, we observe that tapered DFT-based one-sided CI's are comparatively better than the the non-tapered case, although a direct comparison is not feasible due to different optimum block lengths used. For both cases, there exist an optimum choice of a, such that the coverage at $l = an^\beta$, is close to the nominal accuracy of 90%. However, the performance of the CI's become worse

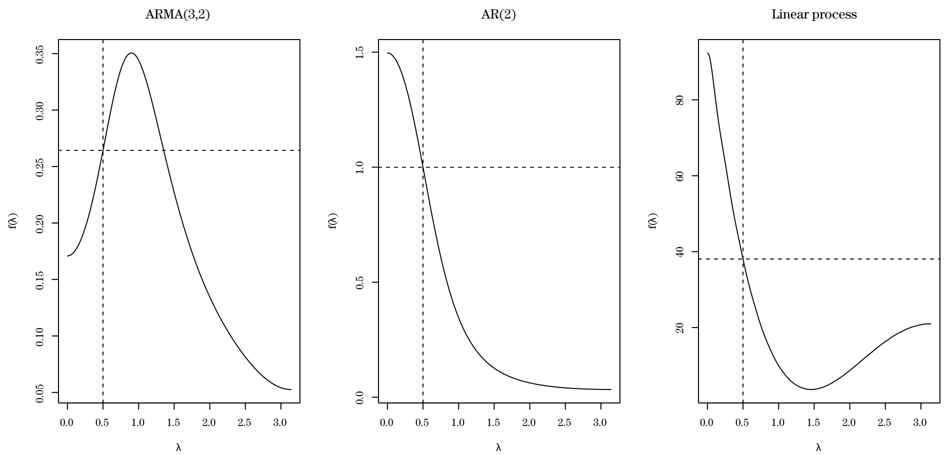

Figure 1. Spectral densities of the ARMA(3,2) (left), AR(2) (middle) and the Linear process model (right), used in Section 5. Vertical dotted line is plotted at $\lambda = 0.5$ and horizontal dotted line shows the value $f(0.5)$.

Table 1. Empirical coverages rates for 90% **lower CI's** for $f(0.5)$, for the ARMA(3,2), AR(2) and Linear process models. CI's based on studentization (Stud.) and variance stabilizing transformation (VST) are studied for two different choices of innovation distributions (F_1 and F_2). Here, $n = 500$ and empirical coverages are based on 1,500 simulations.

	Empirical coverages of lower CI's with varying block lengths[a] (l)											
	Innovation distribution= F_1						Innovation distribution= F_2					
	ARMA(3,2)		AR(2)		Lin. Proc.		ARMA(3,2)		AR(2)		Lin. Proc.	
	Stud.	VST	Stud.	VST	Stud.	VST	Stud.	VST	Stud.	VST	Stud.	VST
$a_1^\dagger$	Tapered Case: $l = a_1 \cdot n^{1/3}$											
0.75	0.837	0.831	0.963	0.963	0.842	0.831	0.865	0.859	0.936	0.934	0.857	0.853
1	0.843	0.833	0.917	0.913	0.728	0.711	0.871	0.864	0.899	0.892	0.774	0.765
2	0.878	0.861	0.905	0.884	0.789	0.757	0.894	0.886	0.896	0.877	0.826	0.810
3.5	0.919	0.894	0.929	0.904	0.903	0.867	0.919	0.905	0.925	0.907	0.896	0.873
5	0.948	0.920	0.946	0.917	0.931	0.894	0.937	0.915	0.936	0.917	0.921	0.899
7	0.965	0.925	0.961	0.931	0.958	0.911	0.957	0.925	0.950	0.927	0.941	0.916
$a_2^\dagger$	Non-tapered Case: $l = a_2 \cdot n^{1/2}$											
0.25	0.897	0.885	0.948	0.941	0.783	0.764	0.893	0.888	0.925	0.919	0.819	0.808
0.5	0.920	0.895	0.945	0.933	0.868	0.847	0.914	0.904	0.929	0.917	0.881	0.859
1	0.959	0.937	0.963	0.941	0.945	0.908	0.944	0.928	0.948	0.935	0.924	0.906
1.5	0.975	0.946	0.975	0.947	0.971	0.935	0.960	0.935	0.965	0.939	0.946	0.918
2	0.984	0.953	0.984	0.953	0.982	0.947	0.973	0.945	0.971	0.945	0.961	0.929
3	0.993	0.966	0.993	0.963	0.996	0.957	0.985	0.958	0.981	0.949	0.976	0.942

[a]Optimum block lengths rates were used, depending on the type of CI's and the use of tapering.
†Different constants were used due to the different optimum block length rates for tapered and non-tapered cases.

Table 2. Empirical coverages rates and average lengths (shown below in parenthesis) for 90% **two-sided CI's** for $f(0.5)$, for the ARMA(3,2), AR(2) and Linear process models. CI's based on studentization (Stud.) and variance stabilizing transformation (VST) are studied for two different choices of innovation distributions (F_1 and F_2). Here, $n = 500$ and empirical coverages are based on 1,500 simulations.

	Empirical coverages and (average lengths) of two-sided CI's with varying block lengths[a] (l)											
	Innovation distribution= F_1						Innovation distribution= F_2					
	ARMA(3,2)		AR(2)		Lin. Proc.		ARMA(3,2)		AR(2)		Lin. Proc.	
	Stud.	VST	Stud.	VST	Stud.	VST	Stud.	VST	Stud.	VST	Stud.	VST
a_1[†]	Tapered Case: $l = a_1 \cdot n^{1/2}$											
0.25	0.557	0.568	0.266	0.287	0.483	0.488	0.336	0.352	0.169	0.181	0.333	0.348
	(0.0731)	(0.0733)	(0.231)	(0.232)	(10.6)	(10.7)	(0.070)	(0.071)	(0.226)	(0.227)	(11.9)	(12.0)
0.5	0.737	0.753	0.691	0.709	0.692	0.679	0.531	0.553	0.525	0.548	0.622	0.616
	(0.112)	(0.113)	(0.410)	(0.413)	(17.6)	(17.7)	(0.108)	(0.109)	(0.402)	(0.405)	(19.8)	(19.9)
1	0.824	0.834	0.824	0.841	0.839	0.821	0.667	0.708	0.701	0.734	0.717	0.735
	(0.159)	(0.162)	(0.597)	(0.606)	(24.2)	(24.6)	(0.153)	(0.155)	(0.584)	(0.593)	(27.4)	(27.8)
1.5	0.865	0.881	0.865	0.877	0.863	0.861	0.719	0.775	0.756	0.788	0.763	0.783
	(0.194)	(0.199)	(0.734)	(0.751)	(29.1)	(29.8)	(0.186)	(0.190)	(0.718)	(0.734)	(33.0)	(33.8)
2	0.871	0.891	0.873	0.887	0.866	0.873	0.762	0.809	0.780	0.814	0.779	0.815
	(0.224)	(0.231)	(0.850)	(0.876)	(33.4)	(34.5)	(0.214)	(0.221)	(0.829)	(0.855)	(38.1)	(39.3)
3	0.880	0.903	0.878	0.899	0.868	0.891	0.792	0.846	0.803	0.848	0.797	0.834
	(0.277)	(0.290)	(1.05)	(1.10)	(40.8)	(42.8)	(0.262)	(0.275)	(1.03)	(1.08)	(46.9)	(49.2)
a_2[†]	Non-tapered Case: $l = a_2 \cdot n^{3/5}$											
0.125	0.668	0.687	0.489	0.507	0.677	0.663	0.431	0.459	0.324	0.339	0.539	0.550
	(0.095)	(0.096)	(0.329)	(0.331)	(15.4)	(15.5)	(0.092)	(0.093)	(0.322)	(0.324)	(17.2)	(17.4)
0.25	0.815	0.820	0.746	0.776	0.819	0.807	0.610	0.641	0.585	0.611	0.667	0.695
	(0.139)	(0.141)	(0.501)	(0.507)	(21.5)	(21.8)	(0.134)	(0.136)	(0.490)	(0.496)	(24.2)	(24.5)
0.5	0.888	0.904	0.872	0.890	0.893	0.896	0.744	0.797	0.751	0.806	0.781	0.805
	(0.198)	(0.203)	(0.734)	(0.752)	(30.1)	(30.9)	(0.191)	(0.195)	(0.717)	(0.735)	(34.1)	(35.0)
1	0.915	0.937	0.898	0.929	0.914	0.931	0.828	0.884	0.832	0.881	0.831	0.878
	(0.286)	(0.300)	(1.07)	(1.13)	(42.4)	(44.6)	(0.272)	(0.285)	(1.05)	(1.10)	(48.4)	(50.9)
1.5	0.909	0.945	0.887	0.945	0.895	0.948	0.844	0.897	0.856	0.907	0.847	0.893
	(0.352)	(0.379)	(1.33)	(1.43)	(51.8)	(55.9)	(0.332)	(0.357)	(1.30)	(1.40)	(59.8)	(64.5)
2	0.900	0.945	0.883	0.952	0.891	0.953	0.835	0.914	0.851	0.911	0.858	0.907
	(0.409)	(0.451)	(1.54)	(1.70)	(59.6)	(65.9)	(0.382)	(0.422)	(1.51)	(1.66)	(69.8)	(77.2)

[a] Optimum block lengths rates were used, depending on the type of CI's and the use of tapering.

[†] Different constants were used due to the different optimum block length rates for tapered and non-tapered cases.

with the increase of block lengths in all cases, leading to over-coverage. The performance of the Studentized and VST-based intervals are comparable, although the best choice of block length for each method is different. The empirical coverages are dependent on the choice of innovation distribution and the underlying model.

From Table 2, we see that the average lengths and coverages of the intervals increase with the increase of block lengths. In case of the linear process, the CI's are wider, but that is due to the nature of its underlying spectral density (cf. right panel in Figure 1). However, in all cases, irrespective of the choice of innovation distribution or the underlying model, there is a block length, where the empirical coverages are very close to the nominal coverage. It seems that the VST-based intervals always have higher coverage than the Studentized intervals, at the expense of slight increase in average length.

The empirical coverage rates under F_2 are comparatively better, than those under F_1 (cf. Table 1). The most likely reason is the symmetry of F_2, even though F_2 has restricted moments. However, in the two-sided case, CI's based on F_1 achieve greater accuracy than those under F_2. This is true for both the tapered and non-tapered cases. The empirical coverages obviously depend on the choice of block length (cf. Table 2). In the one-sided case, for the AR(2) model, there is over-coverage throughout the range of selected block lengths (cf. Table 1). This issue is likely to be resolved with a larger sample size. But, for the two-sided case this is not the situation where over-coverage happens less frequently. Specially, for the AR(2) model over-coverage is drastically reduced.

Overall, both Studentized and VST-based methods are applicable in practice with appropriate choice of block lengths, for a range of time series models and varying choices of innovation distributions.

Supplementary Materials

Proofs of the results and the general framework for deriving EEs and exact statements of the regularity conditions that allow nonlinear and non-Gaussian processes are given in the Supplementary Materials.

Acknowledgment

We wish to thank an associate editor and a referee for their careful reading and helpful comments.

References

Bartlett, M. S. (1946). On the theoretical specification and sampling properties of autocorrelated time-series. *Suppl J. Roy. Statist. Soc.* **8**, 27–41.

Bartlett, M. S. (1950). Periodogram analysis and continuous spectra. *Biometrika* **37**, 1–16.

Blackman, R. B. and Tukey, J. W. (1959). *The Measurement of Power Spectra: From the Point of View of Communications Engineering.* Dover Publications Inc., New York.

Brillinger, D. R. (1975). *Time Series.* Data analysis and theory, International Series in Decision Processes. Holt, Rinehart and Winston, Inc., New York.

Chatterjee, A. and Lahiri, S. N. (2018). Supplementary materials for 'Edgeworth expansions for a class of spectral density estimators and their applications to interval estimation'. Available online. `http://www3.stat.sinica.edu.tw/statistica/`.

Dahlhaus, R. (1985). On a spectral density estimate obtained by averaging periodograms. *J. Appl. Probab.* **22**, 598–610.

Götze, F. and Hipp, C. (1983). Asymptotic expansions for sums of weakly dependent random vectors. *Z. Wahrsch. Verw. Gebiete* **64**, 211–239.

Grenander, U. and Rosenblatt, M. (1957). *Statistical Analysis of Stationary Time Series.* John

Wiley & Sons, New York.

Janas, D. (1994). Edgeworth expansions for spectral mean estimates with applications to Whittle estimates. *Ann. Inst. Statist. Math.* **46**, 667–682.

Lahiri, S. N. (2007). Asymptotic expansions for sums of block-variables under weak dependence. *Ann. Statist.* **35**, 1324–1350.

Lahiri, S. N. (2010). Edgeworth expansions for Studentized statistics under weak dependence. *Ann. Statist.* **38**, 388–434.

Parzen, E. (1961). Mathematical considerations in the estimation of spectra. *Technometrics* **3**, 167–190.

Priestley, M. B. (1981). *Spectral Analysis and Time Series. Vol. 2.* Multivariate series, prediction and control, Probability and Mathematical Statistics. Academic Press Inc. (Harcourt Brace Jovanovich Publishers), London.

Taniguchi, M. and Kakizawa, Y. (2000). *Asymptotic Theory of Statistical Inference for Time Series*. Springer, New York.

Taniguchi, M., van Garderen, K. J. and Puri, M. (2003). Higher order asymptotic theory for minimum contrast estimators of spectral parameters of stationary processes. *Econometric Theory* **19**, 984–1007.

Velasco, C. and Robinson, P. M. (2001). Edgeworth expansions for spectral density estimates and Studentized sample mean. *Econometric Theory* **17**, 497–539.

Welch, P. D. (1967). The use of Fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Trans. Audio. Electro. Acous.* **15**, 70–73.

Zhurbenko, I. G. (1979). Spectral density estimates obtained by a time shift. *Dokl. Akad. Nauk SSSR* **246**, 801–805.

Zhurbenko, I. G. (1980). On a limit theorem for statistics of spectral density with a time shift. *Ukrain. Mat. Zh.* **32**, 463–476.

Theoretical Statistics & Mathematics Unit, Indian Statistical Institute, New Delhi 110016, India.

E-mail: cha@isid.ac.in

Department of Statistics, North Carolina State University, Raleigh, NC 27607, USA.

E-mail: snlahiri@ncsu.edu

(Received March 2017; accepted January 2018)