Article:

Martin, G.M., Nadarajah, K. and Poskitt, D.S. (2020) Issues in the estimation of mis-specified models of fractionally integrated processes. Journal of Econometrics, 215 (2). pp. 559-573. ISSN 0304-4076

https://doi.org/10.1016/j.jeconom.2019.09.007

Supplementary Appendix: Issues in the Estimation of Mis-Specified Models of Fractionally Integrated Processes

Gael M. Martin, K. Nadarajah & D. S. Poskitt

Department of Econometrics and Business Statistics, Monash University

Abstract

This material contains the proofs of Lemmas 1, 2, 3 and 4, Proposition 1, and Theorems 1, 2 and 3. Additional Lemmas used to prove the main theorems in the paper are also provided with proofs. These technical details are given in Appendix A. Appendix B contains the expressions used to calculate the bias term associated with the four estimators.

Appendix A: Proofs

A.1 Proof of Lemmas 1, 2, 3 and 4

A.1.1 Proof of Lemma 1:

The proof of the lemma uses a method that parallels that employed by Fox and Taqqu in the proof of their Lemma 1 (see Fox and Taqqu, 1986, pages 523-524), which in turn employs an argument first developed by Hannan in the proof of his Lemma 1 (see Hannan, 1973, pages 133-134). To describe the approach, set

$$c_n(\tau) = c_n(-\tau) = \frac{1}{n}\sum_{t=1}^{n-\tau} y_t y_{t+\tau}\,, \quad \tau \geq 0\,,$$

and let

$$k_M(\boldsymbol{\eta}, \lambda) = \sum_{r=-M}^{M} \kappa(r)\left(1 - \frac{|r|}{M}\right)\exp(i\lambda r)$$

denote the Cesaro sum of the first M terms of the Fourier series of $(f_1(\boldsymbol{\eta}, \lambda) + \nu_f)^{-1}$ where M is chosen such that $|(f_1(\boldsymbol{\eta}, \lambda) + \nu_f)^{-1} - k_M(\boldsymbol{\eta}, \lambda))| < \epsilon$ uniformly in $\boldsymbol{\eta} \in \mathbb{E}_\delta^0$. Then following the same steps as in the derivation presented in Hannan (1973, pages 133-134) we have

$$\left| \frac{4\pi}{n}\sum_{j=1}^{[n/2]} I(\lambda_j)\left\{(f_1(\boldsymbol{\eta}, \lambda_j) + \nu_f)^{-1} - k_M(\boldsymbol{\eta}, \lambda_j)\right\} \right| < \epsilon c_n(0)\,,$$

and
$$\lim_{n\to\infty}\left|\frac{4\pi}{n}\sum_{j=1}^{[n/2]} I(\lambda_j)k_M(\boldsymbol{\eta},\lambda_j)-\sum_{r=-M}^{M}\kappa(r)\left(1-\frac{|r|}{M}\right)\gamma_0(r)\right|=0$$
almost surely, the latter result since $I(\lambda)=(2\pi)^{-1}\sum_{r=1-n}^{n-1}c_n(r)\exp(-i\lambda r)$ and $c_n(r)$ converges to $\gamma_0(r)$ almost surely by ergodicity. Moreover,
$$\sum_{r=-M}^{M}\kappa(r)\left(1-\frac{|r|}{M}\right)\gamma_0(r)=\frac{\sigma_0^2}{2\pi}\int_{-\pi}^{\pi}f_0(\lambda)k_M(\boldsymbol{\eta},\lambda)d\lambda$$
differs from the required limiting value by a quantity bounded by $\epsilon\gamma_0(0)$, from which the desired result follows because ϵ is arbitrary.

An alternative proof of this lemma can be obtained by extending the arguments adopted by Brockwell and Davis (1991, §10.8.2, page 378-379), in the proof of their Proposition 10.8.2, to the stationary fractional case, as suggested in Brockwell and Davis (1991, page 528).

A.1.2 Proof of Lemma 2:

The proof parallels the proof of Lemma 1, only now we use the Cesaro sum of M terms of the Fourier series of $h_1(\boldsymbol{\eta},\lambda)^{-1}$. Denote this sum by $c_M(\boldsymbol{\eta},\lambda)>0$. Since by construction $h_1(\boldsymbol{\eta},\lambda)>0$, M can be chosen so that $|h_1(\boldsymbol{\eta},\lambda)^{-1}-c_M(\boldsymbol{\eta},\lambda)|<\varepsilon$ uniformly on $\mathbb{E}_\delta^0$ since the Cesaro sum converges uniformly in $(\boldsymbol{\eta},\lambda)$ for $\boldsymbol{\eta}\in\mathbb{E}_\delta^0$. Once again the detailed steps follow Hannan (1973, page 133-134), as above, or Brockwell and Davis (1991, §10.8.2, page 378-379).

A.1.3 Proof of Lemma 3:

Observe that $f_1(\boldsymbol{\eta},\lambda)>0$ when $d\geq 0$ and hence for δ sufficiently small we have $h_1(\boldsymbol{\eta},\lambda)=f_1(\boldsymbol{\eta},\lambda)$ for all $\lambda\in[-\pi,\pi]$. It follows immediately from Lemma 2 that $\lim_{n\to\infty}|Q_n^{(1)}(\boldsymbol{\eta})-Q(\boldsymbol{\eta})|=0$ almost surely and uniformly in $\boldsymbol{\eta}$ on $\mathbb{E}_\delta^0$ when $d\geq 0$. We have thus established Lemma 3 in the case where $d\geq 0$, (*cf.* Chen and Deo, 2006, Lemma 2). To establish that Lemma 3 also holds on $\mathbb{E}_\delta^0$ when $d<0$, observe that Lemma 1 implies that $Q(\boldsymbol{\eta})$ provides a limit inferior for $Q_n^{(1)}(\boldsymbol{\eta})$ and it therefore only remains for us to establish that $Q(\boldsymbol{\eta})$ also provides a limit superior for $Q_n^{(1)}(\boldsymbol{\eta})$ on $\boldsymbol{\eta}\in\mathbb{E}_\delta^0$ when $d<0$.

In the latter case $f_1(\boldsymbol{\eta},\lambda)=|\lambda|^{2|d|}L(\lambda)$ where $L(\lambda)$ is slowly varying and bounded as $\lambda\to 0$ and there exists an $\epsilon\in(0,2|d|)$ and a $K>0$, that may depend on ϵ, such that $f_1(\boldsymbol{\eta},\lambda)=$

$|\lambda|^{2|d|}K|\lambda|^{-\epsilon}$. We therefore have that $f_1(\boldsymbol{\eta},\lambda) > K|\lambda|^{2|d|}$ when $|\lambda| < 1$ and $h_1(\boldsymbol{\eta},\lambda) \neq f_1(\boldsymbol{\eta},\lambda)$ whenever $\lambda < (K^{-1}\delta)^{1/(2|d|-\epsilon)}$, from which it follows that

$$Q_n^{(1)}(\boldsymbol{\eta}) \leq \frac{2\pi}{n}\sum_{j=k_\delta+1}^{\lfloor n/2\rfloor}\frac{I(\lambda_j)}{h_1(\boldsymbol{\eta},\lambda_j)} + \frac{1}{K}\left(\frac{2\pi}{n}\right)^{1-2|d|}\sum_{j=1}^{k_\delta} I(\lambda_j) \tag{1}$$

where $k_\delta = \lfloor (K^{-1}\delta)^{1/(2|d|-\epsilon)}(2\pi/n)\rfloor + 1$. The inequality in (1) follows because for all $\lambda_j < (K^{-1}\delta)^{1/(2|d|-\epsilon)} < 2\pi k_\delta/n$ we have

$$\left(\frac{h_1(\boldsymbol{\eta},\lambda_j)}{f_1(\boldsymbol{\eta},\lambda_j)} - 1\right) \leq \left(\frac{\delta}{K}\left(\frac{n}{2\pi}\right)^{2|d|} - 1\right)$$

and

$$\begin{aligned} Q_n^{(1)}(\boldsymbol{\eta}) &= \frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor}\frac{I(\lambda_j)}{h_1(\boldsymbol{\eta},\lambda_j)} + \frac{2\pi}{n}\sum_{j=1}^{k_\delta} I(\lambda_j)\left(\frac{1}{f_1(\boldsymbol{\eta},\lambda_j)} - \frac{1}{h_1(\boldsymbol{\eta},\lambda_j)}\right) \\ &\leq \frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor}\frac{I(\lambda_j)}{h_1(\boldsymbol{\eta}_1,\lambda_j)} + \frac{2\pi}{n}\sum_{j=1}^{k_\delta}\frac{I(\lambda_j)}{h_1(\boldsymbol{\eta},\lambda_j)}\left(\frac{\delta}{K}\left(\frac{n}{2\pi}\right)^{2|d|} - 1\right) \\ &= \frac{2\pi}{n}\sum_{j=k_\delta+1}^{\lfloor n/2\rfloor}\frac{I(\lambda_j)}{h_1(\boldsymbol{\eta},\lambda_j)} + \frac{1}{K}\left(\frac{2\pi}{n}\right)^{1-2|d|}\sum_{j=1}^{k_\delta} I(\lambda_j)\,. \end{aligned}$$

Applying Lemma 2 to the first term on the right hand side in (1) gives a limit of

$$\frac{\sigma_0^2}{2\pi}\int_{(K^{-1}\delta)^{1/(2|d|-\epsilon)}}^{\pi}\frac{f_0(\lambda)}{f_1(\boldsymbol{\eta}_1,\lambda)}d\lambda\,.$$

Similarly

$$\lim_{n\to\infty}\frac{2\pi}{n}\sum_{j=1}^{k_\delta} I(\lambda_j) = \frac{\sigma_0^2}{2\pi}\int_0^{(K^{-1}\delta)^{1/(2|d|-\epsilon)}} f_0(\lambda)d\lambda = \frac{\sigma_0^2}{2\pi}f_0(\lambda')(K^{-1}\delta)^{1/(2|d|-\epsilon)}$$

for some $\lambda' \in [0,(K^{-1}\delta)^{1/(2|d|-\epsilon)}]$ by the first mean value theorem for integrals. Setting $\delta = (2\pi)^{2|d|-\epsilon}/n^p$ where $p > 2|d|-\epsilon$, we find that

$$\begin{aligned} \frac{1}{K}\left(\frac{2\pi}{n}\right)^{1-2|d|}\sum_{j=1}^{k_\delta} I(\lambda_j) &\sim \frac{1}{K}\left(\frac{n}{2\pi}\right)^{2|d|}\frac{\sigma_0^2}{2\pi}f_0(\lambda')\frac{2\pi k_\delta}{n} \\ &\sim \frac{1}{K}\left(\frac{2\pi}{n}\right)^{1-2|d|}\frac{\sigma_0^2}{2\pi}f_0(\lambda')\left(\frac{1}{n}\right)^{\frac{p-2|d|+\epsilon}{(2|d|-\epsilon)}} \end{aligned}$$

and hence we can conclude that

$$\limsup_{n\to\infty} Q_n^{(1)}(\boldsymbol{\eta}) \leq Q(\boldsymbol{\eta})$$

uniformly in $\boldsymbol{\eta} \in \mathbb{E}_\delta^0$, as required.

A.1.4 Proof of Lemma 4:

Let $L_1(\boldsymbol{\eta}, \lambda) = \lambda^{2d} f_1(\boldsymbol{\eta}, \lambda)$ and suppose that $\boldsymbol{\eta} \in \overline{\mathbb{E}}_{\delta 1}^0 \cup \overline{\mathbb{E}}_{\delta 2}^0 \neq \emptyset$. Then

$$\begin{aligned}
\liminf_{n\to\infty} Q_n^{(1)}(\boldsymbol{\eta}) &= \liminf_{n\to\infty} \frac{2\pi}{n} \sum_{j=1}^{\lfloor n/2 \rfloor} \frac{I(\lambda_j)}{f_1(\boldsymbol{\eta}, \lambda_j)} \\
&= \liminf_{n\to\infty} \frac{2\pi}{n} \sum_{j=1}^{\lfloor n/2 \rfloor} \frac{I(\lambda_j)\lambda_j^{2d}}{L_1(\boldsymbol{\eta}, \lambda_j)} \\
&\geq \liminf_{n\to\infty} \frac{(2\pi)^{-2\delta}}{n} \sum_{j=1}^{\lfloor n/2 \rfloor} \frac{I(\lambda_j)}{L_1(\boldsymbol{\eta}, \lambda_j)\lambda_j^{1-2(d_0+\delta)}}\,, \qquad (2)
\end{aligned}$$

where the inequality in (2) arises because for all $\boldsymbol{\eta} \in \overline{\mathbb{E}}_{\delta 1}^0 \cup \overline{\mathbb{E}}_{\delta 2}^0$ we have $(d_0 - d) > 0.5 - \delta$ and it follows that $\lambda_j^{-2(d_0-d)} \geq (2\pi)^{-(2\delta+1)}\lambda_j^{2\delta-1}$ for all $\lambda_j = 2\pi j/n$, $j = 1, \ldots, \lfloor n/2 \rfloor$.

Applying Lemma 1 and Lemma 2 to (2) by replacing $f_1(\boldsymbol{\eta}, \lambda_j)$ by $L_1(\boldsymbol{\eta}, \lambda)\lambda^{1-2(d_0+\delta)}$, and then letting the constant $\nu_f > 0$in the lemmas approach zero, it follows from Fatou's theorem that

$$\lim_{n\to\infty} \left| \frac{(2\pi)^{-2\delta}}{n} \sum_{j=1}^{\lfloor n/2 \rfloor} \frac{I(\lambda_j)}{L_1(\boldsymbol{\eta}, \lambda_j)\lambda_j^{1-2(d_0+\delta)}} - \frac{1}{(2\pi)^{2\delta+1}} \int_0^\pi \frac{(\sigma_0^2/2\pi) f_0(\lambda)\lambda^{2d_0}}{L_1(\boldsymbol{\eta}, \lambda)\lambda^{1-2\delta}} d\lambda \right| = 0\,,$$

wherein we recognise that $0 \leq 1-2(d_0+\delta) \leq 2(1-2\delta)$ and that $L_1(\boldsymbol{\eta}, \lambda) = (\sigma_1^2/2\pi) g_1(\boldsymbol{\beta}, \lambda)\,\mathrm{sinc}(\lambda/2)^{-2d}$ and $(\sigma_0^2/2\pi) f_0(\lambda)\lambda^{2d_0} = (\sigma_0^2/2\pi) g_0(\lambda)\,\mathrm{sinc}(\lambda/2)^{-2d_0}$ where $\mathrm{sinc}(x) = \sin(x)/x$, the cardinal sine function. Since $2/\pi \leq \mathrm{sinc}(\lambda/2) \leq 1$ for $0 \leq \lambda \leq \pi$, it follows from Assumption $(A.3)$ and

Conditions A that there exists a finite positive constant R such that

$$\begin{aligned}\frac{1}{(2\pi)^{2\delta+1}}\int_0^\pi \frac{(\sigma_0^2/2\pi)f_0(\lambda)\lambda^{2d_0}}{L_1(\boldsymbol{\eta},\lambda)\lambda^{1-2\delta}}d\lambda &\geq \frac{R}{(2\pi)^{2\delta+1}}.\int_0^\pi \lambda^{2\delta-1}d\lambda \\ &= \frac{R}{(2\pi)^{2\delta+1}}.\frac{\pi^{2\delta}}{2\delta} \\ &\geq \frac{R}{8\pi}.\frac{1}{\delta}. \qquad (3)\end{aligned}$$

The statements in Lemma 4 now follow from (3), directly in the case of $\boldsymbol{\eta} \in \overline{\mathbb{E}}^0_{\delta 1}$, and for $\boldsymbol{\eta} \in \overline{\mathbb{E}}^0_{\delta 2}$ on setting $\delta < R/(8\pi C)$ and letting $\delta \to 0$ as $C \to \infty$.

A.2 Proof of Proposition 1

Let $\boldsymbol{\eta}_n$ denote a sequence in $\mathbb{E}^0_\delta$ that converges to $\boldsymbol{\eta}$. For any $\nu_f > 0$ we have

$$\begin{aligned}\left|\frac{1}{f_1(\boldsymbol{\eta}_n,\lambda)+\nu_f} - \frac{1}{f_1(\boldsymbol{\eta},\lambda)+\nu_f}\right| &= \left|\frac{|f_1(\boldsymbol{\eta}_n,\lambda)-f_1(\boldsymbol{\eta},\lambda)|}{(f_1(\boldsymbol{\eta}_n,\lambda)+\nu_f)(f_1(\boldsymbol{\eta},\lambda)+\nu_f)}\right| \\ &\leq \frac{|f_1(\boldsymbol{\eta}_n,\lambda)-f_1(\boldsymbol{\eta},\lambda)|}{\nu_f^2}.\end{aligned}$$

Moreover, by assumption $f_1(\boldsymbol{\eta},\lambda)$ is continuous for all $\lambda \neq 0$ and hence uniformly continuous for λ in any closed interval of the form $[\varepsilon,\pi]$, $\varepsilon > 0$. Consequently we can determine a value n' such that for $n \geq n'$ there exists an ε sufficiently small that $|f_1(\boldsymbol{\eta}_n,\lambda) - f_1(\boldsymbol{\eta},\lambda)| < \nu_f^3$ and

$$\left|\frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor}\frac{I(\lambda_j)}{f_1(\boldsymbol{\eta}_n,\lambda_j)+\nu_f} - \frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor}\frac{I(\lambda_j)}{f_1(\boldsymbol{\eta},\lambda_j)+\nu_f}\right| \leq \frac{2\nu_f\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor} I(\lambda_j)\,. \qquad (4)$$

Using Lemma 1 in conjunction with (4), it follows that

$$\begin{aligned}\liminf_{n\to\infty} Q_n^{(1)}(\boldsymbol{\eta}_n) &\geq \liminf_{n\to\infty}\frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor}\frac{I(\lambda_j)}{f_1(\boldsymbol{\eta}_n,\lambda_j)+\nu_f} \\ &\geq \lim_{n\to\infty}\left\{\frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor}\frac{I(\lambda_j)}{f_1(\boldsymbol{\eta},\lambda_j)+\nu_f} - \frac{2\nu_f\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor} I(\lambda_j)\right\} \\ &= \frac{\sigma_0^2}{2\pi}\int_0^\pi \frac{f_0(\lambda)}{f_1(\boldsymbol{\eta},\lambda)+\nu_f}d\lambda - \nu_f\pi\gamma_0(0)\,,\end{aligned}$$

where $\gamma_0(0)$ is the variance of the TDGP. Letting $\nu_f \to 0$ and applying Lebegue's monotone convergence theorem gives

$$\liminf_{n\to\infty} Q_n^{(1)}(\boldsymbol{\eta}_n) \geq \frac{\sigma_0^2}{2\pi}\int_0^{\pi} \frac{f_0(\lambda)}{f_1(\boldsymbol{\eta},\lambda)} d\lambda = Q(\boldsymbol{\eta})\,.$$

Since by definition $\boldsymbol{\eta}_1$ minimises $Q(\boldsymbol{\eta})$ it follows that $Q(\boldsymbol{\eta}_1)$ provides a lower bound to the limit inferior of $Q_n^{(1)}(\boldsymbol{\eta}_n)$ for any sequence in $\mathbb{E}_\delta^0$.

Now let $\boldsymbol{\eta}_n$ denote a sequence in $\overline{\mathbb{E}}_{\delta 1}^0 \cup \overline{\mathbb{E}}_{\delta 2}^0$ that converges to $\boldsymbol{\eta}$. Setting

$$\delta \ll \min\left\{\frac{\sigma_0^2}{4(2\pi)^2}\frac{C_l}{C_u}\frac{1}{(Q(\boldsymbol{\eta}_1)+q)}, 0.25 - 0.5(d_0 - d_1)\right\} \quad \text{where} \quad q \gg 0\,,$$

and applying Lemma 4 in conjunction with (4) implies that

$$\liminf_{n\to\infty} Q_n^{(1)}(\boldsymbol{\eta}_n) \gg Q(\boldsymbol{\eta}_1) + q\,.$$

Hence we can conclude that for any sequence $\boldsymbol{\eta}_n \in \overline{\mathbb{E}}_{\delta 1}^0 \cup \overline{\mathbb{E}}_{\delta 2}^0$ the criterion value $Q_n^{(1)}(\boldsymbol{\eta}_n)$ will, for all n sufficiently large, exceed $Q(\boldsymbol{\eta}_1)$, which equals $\lim_{n\to\infty} Q_n^{(1)}(\boldsymbol{\eta}_1)$ by Lemma 3.

By definition of $\widehat{\boldsymbol{\eta}}_1^{(1)}$, however, $Q_n^{(1)}(\widehat{\boldsymbol{\eta}}_1^{(1)}) \leq Q_n^{(1)}(\boldsymbol{\eta}_1)$ and it follows from Lemma 3 that

$$\limsup_{n\to\infty} Q_n^{(1)}(\widehat{\boldsymbol{\eta}}_1^{(1)}) \leq \limsup_{n\to\infty} Q_n^{(1)}(\boldsymbol{\eta}_1) = Q(\boldsymbol{\eta}_1)\,.$$

We can therefore conclude that $|Q_n^{(1)}(\widehat{\boldsymbol{\eta}}_1^{(1)}) - Q(\boldsymbol{\eta}_1)| \to 0$ almost surely and an argument by contradiction then shows that $\widehat{\boldsymbol{\eta}}_1^{(1)} \to \boldsymbol{\eta}_1$ with probability one.

A.3 Proof of Theorem 1:

In what follows we assume that the mean is known, and without loss of generality set $\mu = 0$ and suppose that the data is mean corrected.

A.3.1 The Whittle estimator:

Concentrating $Q_n^{(2)}(\sigma^2, \boldsymbol{\eta})$ with respect to σ^2 and setting $n \cdot \lfloor n/2 \rfloor = 0.5$ yields the profile (negative) log-likelihood

$$Q_n^{(2)}(\boldsymbol{\eta}) = \frac{2\pi}{2} \log\left(\frac{\widehat{\sigma}^2(\boldsymbol{\eta})}{2\pi}\right) + \frac{2\pi}{n} \sum_{j=1}^{\lfloor n/2 \rfloor} \log f_1(\boldsymbol{\eta}, \lambda_j) + \pi$$

where $\widehat{\sigma}^2(\boldsymbol{\eta}) = 2Q_n^{(1)}(\boldsymbol{\eta})$ and $Q_n^{(1)}(\boldsymbol{\eta}) = \sum_{j=1}^{\lfloor n/2 \rfloor} I(\lambda_j)/f_1(\boldsymbol{\eta}, \lambda_j)$. Now, following Beran (1994, page 116), we have

$$\frac{2\pi}{n} \sum_{j=1}^{\lfloor n/2 \rfloor} \log f_1(\boldsymbol{\eta}, \lambda_j) = \frac{1}{2} \sum_{r=-\infty}^{\infty} \rho_1(\boldsymbol{\eta}, rn) \to \frac{1}{2} \int_{-\pi}^{\pi} \log f_1(\boldsymbol{\eta}, \lambda) d\lambda \,,$$

where the Fourier coefficients $\rho_1(\boldsymbol{\eta}, r) = \int_{-\pi}^{\pi} \log f_1(\boldsymbol{\eta}, \lambda) \exp(\imath\lambda r) d\lambda$ form a convergent series and

$$\begin{aligned} \int_{-\pi}^{\pi} \log f_1(\boldsymbol{\eta}, \lambda) d\lambda &= \int_{-\pi}^{\pi} \log\left(g_1(\boldsymbol{\beta},\lambda)|2\sin(\lambda/2)|^{-2d}\right) d\lambda \\ &= \int_{-\pi}^{\pi} \log g_1(\boldsymbol{\beta},\lambda) d\lambda - 2d \int_{-\pi}^{\pi} \log|2\sin(\lambda/2)| d\lambda \,. \end{aligned}$$

By Assumption $(A.2)$ $\int_{-\pi}^{\pi} \log g_1(\boldsymbol{\beta},\lambda) d\lambda = 0$, and from standard results for trigonometric integrals Gradshtein and Ryzhik (2007, page 583)

$$\int_{-\pi}^{\pi} \log|2\sin(\lambda/2)| d\lambda = 2\int_{0}^{\pi} \log|2\sin(\lambda/2)| d\lambda = 0 \,.$$

Furthermore, since $\log f_1(\boldsymbol{\eta}, \lambda)$ is integrable, and continuously differentiable for all $\lambda \neq 0$ by Assumption A.3, $\rho_1(\boldsymbol{\eta}, n) = o(1/n)$, which implies that

$$\frac{2\pi}{n} \sum_{j=1}^{\lfloor n/2 \rfloor} \log f_1(\boldsymbol{\eta}, \lambda_j) = \sum_{r=1}^{\infty} \rho_1(\boldsymbol{\eta}, rn) = O(n^{-1} \log n) \,.$$

Hence it follows that

$$\left| Q_n^{(2)}(\boldsymbol{\eta}) - \pi \log Q_n^{(1)}(\boldsymbol{\eta}) - \pi(\log \pi + 1) \right| = O(n^{-1} \log n) \,, \tag{5}$$

almost surely and uniformly in $\boldsymbol{\eta}$. From this we can deduce that

$$\lim_{n\to\infty} \left| Q_n^{(2)}(\widehat{\boldsymbol{\eta}}_1^{(2)}) - \pi \log Q_n^{(1)}(\widehat{\boldsymbol{\eta}}_1^{(1)}) - \pi(\log \pi + 1) \right| = 0 \quad a.s. \,,$$

where $\widehat{\boldsymbol{\eta}}_1^{(1)}$ is the value of $\boldsymbol{\eta}$ that minimises the profile log-likelihood, *having first deleted the term* $2\pi \sum_{j=1}^{\lfloor n/2 \rfloor} \log f_1(\boldsymbol{\eta}, \lambda_j)/n$, namely $\widehat{\boldsymbol{\eta}}_1^{(1)} = \arg\min_{\boldsymbol{\eta}} Q_n^{(1)}(\boldsymbol{\eta})$. We are thereby lead directly to the conclusion that $\widehat{\boldsymbol{\eta}}_1^{(2)}$ and $\widehat{\boldsymbol{\eta}}_1^{(1)}$ converge, *i.e.* $\lim_{n\to\infty} \|\widehat{\boldsymbol{\eta}}_1^{(2)} - \widehat{\boldsymbol{\eta}}_1^{(1)}\| = 0$.

A.3.2 The TML estimator:

Using the argument employed by Hannan (1973, page 134-135) in the proof of his Lemma 4, following the detailed steps given by Brockwell and Davis (1991, §10.8.2, page 380-382) in their proof of their Proposition 10.8.3, shows that

$$\lim_{n\to\infty} \left| \frac{1}{n} \mathbf{Y}^\top \boldsymbol{\Sigma}_\eta^{-1} \mathbf{Y} - \frac{4\pi}{n} \sum_{j=1}^{\lfloor n/2 \rfloor} \frac{I(\lambda_j)}{f_1(\boldsymbol{\eta}, \lambda_j)} \right| = 0 \quad a.s. \,, \tag{6}$$

and the convergence is uniform in $\boldsymbol{\eta}$ on $\mathbb{E}_\delta^0$. From a theorem due to Grenander and Szego (1958, Chapter 5) we know that

$$\frac{1}{n} \log |\boldsymbol{\Sigma}_\eta| = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f_1(\boldsymbol{\eta}, \lambda) d\lambda + O(n^{-1}). \tag{7}$$

That the convergence in (7) is uniform in $\boldsymbol{\eta}$ is not stated in Grenander and Szegö, although it follows from the uniformity of the order relations used in their proof. Their proof depends on approximating $f_1(\boldsymbol{\eta}, \lambda)$ by trigonometric polynomials, and since $f_1(\boldsymbol{\eta}, \lambda)$ is a continuous function of $\boldsymbol{\eta}$ and λ for all $\lambda \neq 0$ by Assumption A.3 the Stone-Weierstrass Theorem implies that $f_1(\boldsymbol{\eta}, \lambda)$ can be so approximated uniformly. It follows that

$$\lim_{n\to\to\infty} \left| Q_n^{(3)}(\sigma^2, \boldsymbol{\eta}) - \log \sigma^2 - \frac{2Q_n^{(1)}(\boldsymbol{\eta})}{\sigma^2} \right| = 0$$

almost surely, and the convergence is uniform in $\boldsymbol{\eta}$ on $\mathbb{E}_\delta^0$.

The almost sure limit of the criterion function $Q_n^{(3)}(\sigma^2, \boldsymbol{\eta})$ is therefore

$$\mathcal{Q}^{(3)}(\sigma^2, Q(\boldsymbol{\eta})) = \log \sigma^2 + \frac{2Q(\boldsymbol{\eta})}{\sigma^2},$$

uniformly in $\boldsymbol{\eta}$ on $\mathbb{E}_\delta^0$ by Lemma 3, whereas $Q_n^{(3)}(\sigma^2, \boldsymbol{\eta})$ is either arbitrarily large for δ sufficiently small or divergent on $\overline{\mathbb{E}}_{\delta 1}^0 \cup \overline{\mathbb{E}}_{\delta 2}^0$ by Lemma 4. Concentrating $\mathcal{Q}^{(3)}(\sigma^2, Q(\boldsymbol{\eta}))$ with respect to σ^2 we find that the minimum of the asymptotic criterion function is given by $\log(2Q(\boldsymbol{\eta}_1))+1$. Once again $\boldsymbol{\eta}_1 = \arg\min_{\boldsymbol{\eta}} Q(\boldsymbol{\eta})$ is the pseudo-true parameter for the estimator under misspecification and we can conclude that $\lim_{n\to\infty} \widehat{\boldsymbol{\eta}}_1^{(3)} = \boldsymbol{\eta}_1$ and $\lim_{n\to\infty} \|\widehat{\boldsymbol{\eta}}_1^{(3)} - \widehat{\boldsymbol{\eta}}_1^{(1)}\| = 0$.

A.3.3 The CSS estimator:

Recall that the objective function of the CSS estimation method is

$$Q_n^{(4)}(\boldsymbol{\eta}) = \frac{1}{n}\sum_{t=1}^{n} e_t(\boldsymbol{\eta})^2, \tag{8}$$

where

$$e_t(\boldsymbol{\eta}) = \sum_{i=0}^{t-1} \tau_i(\boldsymbol{\eta})\,(y_{t-i} - \mu)\,, \quad t = 1, \ldots, n\,, \tag{9}$$

and the coefficients $\tau_j(\boldsymbol{\eta})$, $j = 0, 1, 2, \ldots$, are given by $\tau_0(\boldsymbol{\eta}) = 1$ and

$$\tau_j(\boldsymbol{\eta}) = \sum_{s=0}^{j} \frac{\alpha_{j-s}(\boldsymbol{\beta})\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}\,, \quad j = 1, 2, \ldots\,. \tag{10}$$

Let $\mathbf{T}_\eta$ and $\mathbf{H}_\eta$ denote the $n \times n$ upper triangular Toeplitz matrix with non-zero elements $\tau_{|i-j|}(\boldsymbol{\eta})$, $i, j = 1, \ldots, n$, and the $n \times \infty$ reverse Hankel matrix with typical element $\tau_{n-i+j}(\boldsymbol{\eta})$, $i = 1, \ldots, n$, $j = 1, \ldots, \infty$, respectively. Let $\mathbf{A}_\eta = [a_{s-r}(\eta)]$ where

$$a_{s-r}(\eta) = \int_{-\pi}^{\pi} \frac{1}{f_1(\boldsymbol{\eta},\lambda)} \exp(i(s-l)\lambda)d\lambda\,, \quad r, s = 1, \ldots, n\,. \tag{11}$$

Then from (11) we can deduce that $\mathbf{A}_\eta = \mathbf{T}_\eta\mathbf{T}_\eta^\top + \mathbf{H}_\eta\mathbf{H}_\eta^\top$ and from (8) and (9) it follows that $Q_n^{(4)}(\boldsymbol{\eta}) = \frac{1}{n}\mathbf{Y}^\top\mathbf{T}_\eta\mathbf{T}_\eta^\top\mathbf{Y}$. Replacing $\boldsymbol{\Sigma}_\eta^{-1}$ by $\mathbf{A}_\eta$ in (6) and adapting the argument used to establish (6) accordingly, in a manner similar to the proof of Lemma 1, shows that

$$\lim_{n\to\infty}\left|\frac{1}{n}\mathbf{Y}^\top\mathbf{A}_\eta\mathbf{Y}-\frac{4\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor}\frac{I(\lambda_j)}{f_1(\boldsymbol{\eta},\lambda_j)}\right|=0 \quad a.s. \tag{12}$$

and that the convergence is uniform in $\boldsymbol{\eta}$ on $\mathbb{E}_\delta^0$. It is also shown below that $\frac{1}{n}\mathbf{Y}^\top\mathbf{H}_\eta\mathbf{H}_\eta^\top\mathbf{Y}=o(1)$ when $|d|<0.5$, $|d_0|<0.5$ and $d_0-d<0.5$.

We can therefore conclude that $\left|Q_n^{(4)}(\boldsymbol{\eta})-2Q_n^{(1)}(\boldsymbol{\eta})\right|$ converges to zero almost surely when $\boldsymbol{\eta}\in\mathbb{E}_\delta$, and hence that the limiting value of the criterion function $Q_n^{(4)}(\boldsymbol{\eta})$ is $2Q(\boldsymbol{\eta})$ by Lemma 3. When $\boldsymbol{\eta}\in\overline{\mathbb{E}}_{\delta 1}^0\cup\overline{\mathbb{E}}_{\delta 2}^0$, expression (12) and Lemma 4, together with the equality $Q_n^{(4)}(\boldsymbol{\eta})=\frac{1}{n}\mathbf{Y}^\top\mathbf{A}_\eta\mathbf{Y}-\frac{1}{n}\mathbf{Y}^\top\mathbf{H}_\eta\mathbf{H}_\eta^\top\mathbf{Y}$, imply that $\liminf_{n\to\infty}Q_n^{(4)}(\boldsymbol{\eta})\geq\lim_{n\to\infty}\frac{1}{n}\mathbf{Y}^\top\mathbf{A}_\eta\mathbf{Y}$ and the CSS criterion function is either arbitrarily large for δ sufficiently small or divergent. That the pseudo-true parameter for the CSS estimator under mis-specification is $\boldsymbol{\eta}_1=\arg\min_{\boldsymbol{\eta}}Q(\boldsymbol{\eta})$ and $\lim\widehat{\boldsymbol{\eta}}_1^{(4)}=\boldsymbol{\eta}_1$ and $\lim_{n\to\infty}\|\widehat{\boldsymbol{\eta}}_1^{(4)}-\widehat{\boldsymbol{\eta}}_1^{(1)}\|=0$ follows directly.

It remains for us to establish that $\frac{1}{n}\mathbf{Y}^\top\mathbf{H}_\eta\mathbf{H}_\eta^\top\mathbf{Y}=o(1)$ in regions of the parameter space where $d_0-d<0.5$. Suppressing the dependence on the parameter $\boldsymbol{\eta}$ for notational simplicity, set $\mathbf{M}=\mathbf{H}\mathbf{H}^\top$. Then $\mathbf{M}=[m_{ij}]_{i,j=1,\ldots,n}$ where $m_{ij}=\sum_{u=0}^\infty\tau_{u+n-i}\tau_{u+n-j}$, and

$$E_0[\mathbf{Y}^\top\mathbf{M}\mathbf{Y}]=\mathrm{tr}\,(\mathbf{M}\boldsymbol{\Sigma}_0)=\sum_{i=1}^n\sum_{j=1}^n m_{ij}\gamma_0(j-i)$$

where $\gamma_0(\tau)$, $\tau=0,\pm1,\pm2,\ldots$, denotes the autocovariance function of the TDGP. Since $|\tau_k|\sim k^{-(1+d)}\mathcal{C}_\tau$, $\mathcal{C}_\tau<\infty$, the series $\sum_{k=0}^\infty|\tau_k|^2\sim\mathcal{C}_\tau^2\zeta(2(d+1))$ for all $d>-1/2$, where $\zeta(\cdot)$ denotes the Riemann zeta function, from which we can deduce that $|m_{ij}|\sim\{(n-i+1)(n-j+1)\}^{-(1+d)}\mathcal{C}'_m$ for some $\mathcal{C}'_m<\infty$. Hence on setting $r=n-i+1$ and $s=n-j+1$ we have that

$$0\leq\sum_{i=1}^n\sum_{j=1}^n m_{ij}\gamma_0(j-i)\sim\mathcal{C}_m n^{-2(d+1)}\sum_{r=1}^n\sum_{s=1}^n|\gamma_0(r-s)|\,, \tag{13}$$

where $\mathcal{C}_m < \infty$. But $|\gamma_0(\tau)| \leq \mathcal{C}_\varrho \gamma_0(0)|\tau|^{2d_0-1}$, $\mathcal{C}_\varrho < \infty$, for all $\tau \neq 0$, and

$$\begin{aligned} n^{-2(d+1)} \sum_{r=1}^{n}\sum_{s=1}^{n} |\gamma_0(r-s)| &\leq n^{-2(d+1)}\gamma_0(0)(n + 2\mathcal{C}_\varrho \sum_{k=1}^{n-1}(n-k)k^{2d_0-1}) \\ &\leq n^{-(2d+1)}\gamma_0(0)(1 + 2\mathcal{C}_\varrho \sum_{k=1}^{n-1} k^{2d_0-1}) \\ &\sim \frac{\gamma_0(0)}{n^{(2d+1)}} \times \begin{cases} 1 + 2\mathcal{C}_\varrho \zeta(1-2d_0)\,, & d_0 < 0; \\ 1 + 2\mathcal{C}_\varrho \log n\,, & d_0 = 0; \\ 1 + 2\mathcal{C}_\varrho n^{2d_0}/2d_0\,, & d_0 > 0. \end{cases} \end{aligned}$$

It follows that for all d where $|d| < 0.5$

$$E_0[\mathbf{Y}^\top \mathbf{M}\mathbf{Y}] \leq \frac{\mathcal{C}_m \gamma_0(0)}{n^{1-2(d_0-d)}} \times \begin{cases} 1 + 2\mathcal{C}_\varrho \zeta(1-2d_0)/n^{2d_0}\,, & d_0 < 0; \\ 1 + 2\mathcal{C}_\varrho \log n\,, & d_0 = 0; \\ 1 + \mathcal{C}_\varrho / d_0\,, & d_0 > 0; \end{cases}$$

We can therefore conclude that

$$Pr\left(n^{-1}\mathbf{Y}^\top \mathbf{M}\mathbf{Y} > \epsilon\right) = \begin{cases} O(n^{-2(d+1)})\,, & 0.5 < d_0 < 0; \\ O(\log n / n^{2(d+1)})\,, & d_0 = 0; \\ O(n^{2(d_0-d)-2})\,, & 0 < d_0 < 0.5; \end{cases} \tag{14}$$

for all $\epsilon > 0$ by Markov's inequality. Since ϵ is arbitrary it follows that when $|d| < 0.5$ and $|d_0| < 0.5$ the almost sure limit of $n^{-1}\mathbf{Y}^\top \mathbf{M}\mathbf{Y}$ is zero whenever $d_0 - d < 0.5$, by the Borell-Cantelli lemma, giving the desired result.

A.4 Proof of Theorem 2:

First note that

$$Q_N(\boldsymbol{\eta}) = \left\{ \frac{\sigma_0^2 \Gamma(1-2(d_0-d))}{2\Gamma^2(1-(d_0-d))} \right\} K_N(\boldsymbol{\eta}) \tag{15}$$

by the same argument that gives **(15)**. Now let $\Delta C_N(z) = \sum_{j=N+1}^{\infty} c_j z^j = C(z) - C_N(z)$. Then

$$|C(e^{i\lambda})|^2 = |C_N(e^{i\lambda})|^2 + C_N(e^{i\lambda})\Delta C_N(e^{-i\lambda}) + \Delta C_N(e^{i\lambda})C_N(e^{-i\lambda}) + |\Delta C_N(e^{i\lambda})|^2$$

and the remainder term can be decomposed as $R_N = R_{1N} + R_{2N}$ where

$$R_{1N} = \left(\frac{\sigma_{\varepsilon 0}^2}{2\pi}\right)\int_0^{\pi} |\Delta C_N(e^{i\lambda})|^2 |2\sin(\lambda/2)|^{-2(d_0-d)} d\lambda \tag{16}$$

and

$$R_{2N} = \left(\frac{\sigma_{\varepsilon 0}^2}{2\pi}\right)\int_{-\pi}^{\pi} \Delta C_N(e^{i\lambda})C_N(e^{-i\lambda}) |2\sin(\lambda/2)|^{-2(d_0-d)} d\lambda\,. \tag{17}$$

The first integral in (16) equals

$$\left\{\frac{\sigma_0^2\Gamma(1-2(d_0-d))}{2\Gamma^2(1-(d_0-d))}\right\}\left(\sum_{j=N+1}^{\infty} c_j^2 + 2\sum_{k=N+1}^{\infty}\sum_{j=k+1}^{\infty} c_j c_k \rho(j-k)\right).$$

Because $B(z) \neq 0$, $|z| \leq 1$, it follows that $|c_j| < \mathcal{C}\zeta^j$, $j = 1, 2, \ldots$, for some $\mathcal{C} < \infty$ and $\zeta \in (0,1)$, and hence that

$$\sum_{j=N+1}^{\infty} c_j^2 < \zeta^{2(N+1)}\frac{\mathcal{C}^2}{(1-\zeta^2)}\,.$$

Furthermore, since $|d_0 - d| < 0.5$ Sterling's approximation can be used to show that $|\rho(h)| < \mathcal{C}'^{2(d_0-d)-1}$, $h = 1, 2, \ldots$, for some $\mathcal{C}' < \infty$. This implies that

$$\begin{aligned}\left|\sum_{k=N+1}^{\infty}\sum_{j=k+1}^{\infty} c_j c_k \rho(j-k)\right| &< \sum_{r=0}^{\infty}\sum_{s=r+1}^{\infty} \mathcal{C}^2\mathcal{C}'^{2(N+1)}\zeta^r\zeta^s(s-r)^{2(d_0-d)-1} \\ &< \zeta^{2(N+1)}\frac{\mathcal{C}^2\mathcal{C}'}{(1-\zeta)^2}\,.\end{aligned}$$

Thus we can conclude that $R_{1N} < \mathit{const.}\,\zeta^{2(N+1)}$ where $0 < \zeta < 1$. Applying the Cauchy-Schwarz inequality to the second integral in (17) enables us to bound $|R_{2N}|$ by $2(\sigma_{\varepsilon 0}/\sigma)\sqrt{I_N \cdot R_{1N}}$. It therefore follows from the preceding analysis that $|R_{2N}| < \mathit{const.}\,\zeta^{(N+1)}$. Since $|R_N| \leq R_{1N} + |R_{2N}|$ and $(N+1)/\exp(-(N+1)\log\zeta) \to 0$ as $N \to \infty$ it follows that $R_N = o(N^{-1})$, as stated. The gradient vector of $Q(\boldsymbol{\eta})$ with respect to $\boldsymbol{\eta}$ is

$$\frac{\partial Q(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} = \left(\frac{\sigma_0^2}{2\pi}\right) \int_{-\pi}^{\pi} \frac{C(e^{i\lambda})}{|2\sin(\lambda/2)|^{(d_0-d)}} \frac{\partial}{\partial \boldsymbol{\eta}} \{\frac{C(e^{-i\lambda})}{|2\sin(\lambda/2)|^{(d_0-d)}}\} d\lambda$$

and substituting $C(z) = C_N(z) + \Delta C_N(z)$ gives $\partial Q(\boldsymbol{\psi})/\partial \eta_j = \partial Q_N(\boldsymbol{\eta})/\partial \eta_j + R_{3N} + R_{4N}$ for the typical j'th element where

$$R_{3N} = \left(\frac{\sigma_0^2}{2\pi}\right) \int_{-\pi}^{\pi} \frac{C_N(e^{i\lambda})}{|2\sin(\lambda/2)|^{(d_0-d)}} \frac{\partial}{\partial \eta_j} \{\frac{\Delta C_N(e^{-i\lambda})}{|2\sin(\lambda/2)|^{(d_0-d)}}\} d\lambda$$

and

$$R_{4N} = \left(\frac{\sigma_0^2}{2\pi}\right) \int_{-\pi}^{\pi} \frac{\Delta C_N(e^{i\lambda})}{|2\sin(\lambda/2)|^{(d_0-d)}} \frac{\partial}{\partial \eta_j} \{\frac{C(e^{-i\lambda})}{|2\sin(\lambda/2)|^{(d_0-d)}}\} d\lambda \, .$$

The Cauchy-Schwarz inequality now yields the inequalities

$$|R_{3N}|^2 \leq R_{1N} \left(\frac{\sigma_0^2}{2\pi}\right) \int_{-\pi}^{\pi} \frac{|C_N(e^{i\lambda})|^2}{|2\sin(\lambda/2)|^{2(d_0-d)}} \left| \frac{\partial}{\partial \eta_j} \{\log \frac{\Delta C_N(e^{-i\lambda})}{|2\sin(\lambda/2)|^{(d_0-d)}}\} \right|^2 d\lambda$$

and

$$|R_{4N}|^2 \leq R_{1N} \left(\frac{\sigma_0^2}{2\pi}\right) \int_{-\pi}^{\pi} \left| \frac{\partial}{\partial \eta_j} \{\frac{C(e^{-i\lambda})}{|2\sin(\lambda/2)|^{(d_0-d)}}\} \right|^2 d\lambda \, ,$$

from which we can infer that $|R_{3N} + R_{4N}| \leq const. \zeta^{(N+1)} = o(N^{-1})$, thus completing the proof.

A.5 Proof of Theorem 3:

The distributions exhibited in the three cases presented in Theorem 3 correspond to those given in Theorems 1, 3 and 2 of Chen and Deo (2006), and in the following lemmas we state the properties necessary to generalise the applicability of these distributions and establish their validity under the current scenario and assumptions. Although the distributions are non-standard, the proof proceeds standardly via the use of the mean value theorem and convergence in probability of a Hessian in a neighbourhood of $\boldsymbol{\eta}_1$, plus the application to the criterion differential function of an appropriate central limit theorem.

Lemma 1 (A.1) *Let*

$$\frac{1}{\sqrt{2\pi n}}\sum_{t=1}^{n} y_t \exp(-i\lambda t) = \xi(\lambda) = \xi_c(\lambda) - \imath\xi_s(\lambda)$$

and set $\mathbf{X}^\top = (\xi_c(\lambda_1), \xi_s(\lambda_1), \ldots, \xi_c(\lambda_{\lfloor n/2 \rfloor}), \xi_s(\lambda_{\lfloor n/2 \rfloor})\mathbf{F}_0^{-1/2}$ *where*

$$\mathbf{F}_0 = diag(f_0(\lambda_1), f_0(\lambda_1), \ldots, f_0(\lambda_{\lfloor n/2 \rfloor}), f_0(\lambda_{\lfloor n/2 \rfloor}))\,.$$

Assume that Conditions A hold. Then under Assumption A.$1'$ *the vector* $\mathbf{X}^\top$ *converges in distribution to a Gaussian random variable with zero mean and variance-covariance matrix* $\mathbf{\Omega} = \frac{1}{2}(\mathbf{I}+\mathbf{\Delta})$, $\mathbf{X}^\top \xrightarrow{D} \boldsymbol{\xi} \sim N(\mathbf{0}, \mathbf{\Omega})$, *where* $\mathbf{\Delta} = [\triangle_{rc}]$, $\triangle_{rc} = O(j^{-d_0}k^{d_0-1}\log k)$ *for* $r = 2j-1$ *or* $r = 2j$, *and* $c = 2k-1$ *or* $2k$, $1 \leq j \leq k \leq \lfloor n/2 \rfloor$.

Proof of Lemma A.1.

Assumption $(A.1')$ implies that Assumption $(A.1)$ of Lahiri (2003) holds. Since *Conditions A* imply that Assumption $(A.3)$ of Lahiri (2003) also holds, the asymptotic normality of $\mathbf{X}^T$ follows from Theorem 2.1 of Lahiri (2003). The stated covariance structure follows from Lemmas 1 and 4 of Moulines and Soulier (1999) in which the moment properties of $\xi_c(\lambda_j)$ and $\xi_s(\lambda_j)$ are derived supposing that *exact* Gaussianity holds for the sine and cosine transforms for all n, with bounds that are uniform with respect to n for each $j = 1, \ldots, \lfloor n/2 \rfloor$. See also Corollary 5.2 of Lahiri (2003) and the discussion in Lahiri (2003, page 624).

Since the limiting joint distribution of the sine and cosine transforms is Gaussian, and the sine and cosine transforms are uniformly integrable, the form of the asymptotic distribution and covariance properties of the corresponding periodogram ordinates are determined by the limit law of $\xi_c(\lambda_j)$ and $\xi_s(\lambda_j)$, $j = 1, \ldots, \lfloor n/2 \rfloor$.

Corollary 1 (A.1) *Assume that the conditions of Lemma A.1 hold, and for each* $j = 1, \ldots \lfloor n/2 \rfloor$ *set* $Z_j = I(\lambda_j)/f_0(\lambda_j) = |\xi(\lambda_j)|^2/f_0(\lambda_j)$ *and let* $\rho_j = Cov_0[\xi_c(\lambda_j)\xi_s(\lambda_j)]/f_0(\lambda_j)$. *Then* $Z_j - \rho_j\xi_c(\lambda_j)\xi_s(\lambda_j)/f_0(\lambda_j)$ *converges in distribution to* $\frac{1}{2}\chi^2(2)(1+\triangle_{2j2j})(1-\rho_j^2)$ *where* $\chi^2(2)$ *denotes a Chi-squared random variable with two degrees of freedom. Furthermore,* $E_0[Z_j] = 1 + O(\log j/j)$, $Var_0[Z_j] = 1 + O(\log j/j)$ *and* $Cov_0[Z_j Z_k] = O(j^{-2|d_0|}k^{2|d_0|-2}log^2 k)$ *for* $1 \leq j < k \leq \lfloor n/2 \rfloor$.

Proof of Corollary A.1.

For $j = 1, \dots \lfloor n/2 \rfloor$ set

$$U_j = \frac{\xi_c(\lambda_j) - \xi_s(\lambda_j)]}{\sqrt{f_0(\lambda_j)(1 + \triangle_{2j2j})(1 - \rho_j)}} \quad \text{and} \quad V_j = \frac{\xi_c(\lambda_j) + \xi_s(\lambda_j)]}{\sqrt{f_0(\lambda_j)(1 + \triangle_{2j2j})(1 + \rho_j)}} .$$

Then the Continuous Mapping Theorem implies that

$$\frac{Z_j - \rho_j \xi_c(\lambda_j)\xi_s(\lambda_j)/f_0(\lambda_j)}{(1 + \triangle_{2j2j})(1 - \rho_j^2)} = U_j^2 + V_j^2 \xrightarrow{D} \tfrac{1}{2}\chi^2(2)$$

since by Lemma A.1 $\mathbf{X}^T \xrightarrow{D} \boldsymbol{\xi} \sim N(\mathbf{0}, \boldsymbol{\Omega})$. Let $\mathbf{A}$ and $\mathbf{B}$ be any $\lfloor n/2 \rfloor \times \lfloor n/2 \rfloor$ symmetric selection matrices. Then $E[\boldsymbol{\xi}^\top \mathbf{A} \boldsymbol{\xi}] = \mathrm{tr}\boldsymbol{\Omega}\mathbf{A}$ and $E[(\boldsymbol{\xi}^\top \mathbf{A}\boldsymbol{\xi})(\boldsymbol{\xi}^\top \mathbf{B}\boldsymbol{\xi})] = \mathrm{tr}\boldsymbol{\Omega}\mathbf{A}\mathrm{tr}\boldsymbol{\Omega}\mathbf{B} + \mathrm{tr}\boldsymbol{\Omega}\mathbf{A}\boldsymbol{\Omega}\mathbf{B}$, from which the stated moments can be derived via appropriate choice of $\mathbf{A}$ and $\mathbf{B}$. Note, in particular, that $\rho_j = Cov_0[\xi_c(\lambda_j)\xi_s(\lambda_j)]/f_0(\lambda_j) = \frac{1}{2}\triangle_{(2j-1)2j} = O(\log j/j)$ and $Cov[\xi_j^2 \xi_k^2] = (E[\xi_j \xi_k])^2 = \frac{1}{4}\triangle_{2j2k}^2 = O(j^{-2|d_0|}k^{2|d_0|-2}log^2 k)$ for $1 \le j < k \le \lfloor n/2 \rfloor$.

The remaining steps in the proof of Theorem 3 are based on Taylor expansions of the gradient vector (or score function) of the criterion functions. For the FML estimator we have

$$\mathbf{0} = \frac{\partial Q_n^{(1)}(\boldsymbol{\eta}_1)}{\partial \boldsymbol{\eta}} + \frac{\partial^2 Q_n^{(1)}(\overline{\boldsymbol{\eta}}_1)}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}'}(\widehat{\boldsymbol{\eta}}_1 - \boldsymbol{\eta}_1)$$

where

$$\begin{aligned} \frac{\partial Q_n^{(1)}(\boldsymbol{\eta}_1)}{\partial \boldsymbol{\eta}} &= -\frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2 \rfloor} \frac{I(\lambda_j)}{f_1(\boldsymbol{\eta}, \lambda_j)^2}\frac{\partial f_1(\boldsymbol{\eta}, \lambda_j)}{\partial \boldsymbol{\eta}} = -\frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2 \rfloor} \frac{I(\lambda_j)}{f_0(\lambda_j)}\mathbf{w}(\boldsymbol{\eta}, \lambda_j) \\ \mathbf{w}(\boldsymbol{\eta}, \lambda_j) &= \frac{f_0(\lambda_j)}{f_1(\boldsymbol{\eta}, \lambda_j)}\frac{\partial \log f_1(\boldsymbol{\eta}, \lambda_j)}{\partial \boldsymbol{\eta}} , \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2 Q_n^{(1)}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}'} &= \frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2 \rfloor} \frac{I(\lambda_j)}{f_1(\boldsymbol{\eta}, \lambda_j)}\mathbf{H}(\boldsymbol{\eta}, \lambda_j) , \\ \mathbf{H}(\boldsymbol{\eta}, \lambda_j) &= 2\frac{\partial \log(f_1(\boldsymbol{\eta}, \lambda_j))}{\partial \boldsymbol{\eta}}\frac{\partial \log(f_1(\boldsymbol{\eta}, \lambda_j))}{\partial \boldsymbol{\eta}'} - \frac{1}{f_1(\boldsymbol{\eta}, \lambda_j)}\frac{\partial^2 f_1(\boldsymbol{\eta}, \lambda_j)}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}'} , \end{aligned}$$

and the components of $\overline{\boldsymbol{\eta}}_1$ lie on the line segment between $\widehat{\boldsymbol{\eta}}_1$ and $\boldsymbol{\eta}_1$. Existence of the Taylor expansion is justified by convexity and Assumptions $(A.3)$ and $(A.5)$.

Lemma 2 (A.2) *Let* $dQ_n^{(1)}(\boldsymbol{\eta};\mathbf{t})$, *where* $\mathbf{t}=(t_1,\ldots,t_{l+1})^\top$, *denote the differential of* $Q_n^{(1)}(\boldsymbol{\eta})$. *Then under the assumptions of Theorem 3*

$$\frac{n}{2\pi}\left(\sum_{j=1}^{\lfloor n/2\rfloor}\left(\mathbf{t}^\top\mathbf{w}(\boldsymbol{\eta}_1,\lambda_j)\right)^2\right)^{-\frac{1}{2}}\left(dQ_n^{(1)}(\boldsymbol{\eta}_1;\mathbf{t})-E[dQ_n^{(1)}(\boldsymbol{\eta}_1;\mathbf{t})]\right)\xrightarrow{D} Z\sim N(0,1)$$

for all $\mathbf{t}\in\mathbb{R}^{l+1}$, $0<\|\mathbf{t}\|<\infty$.

Proof of Lemma A.2. By Assumption A.3 the differential of $Q_n^{(1)}(\boldsymbol{\eta})$ exists and is given by $\partial Q_n^{(1)}(\boldsymbol{\eta}_1)/\partial\boldsymbol{\eta}^\top\mathbf{t}$, from which it follows that

$$dQ_n^{(1)}(\boldsymbol{\eta}_1;\mathbf{t})-E[dQ_n^{(1)}(\boldsymbol{\eta}_1;\mathbf{t})]=-\frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor}(Z_j-E[Z_j])\mathbf{w}(\boldsymbol{\eta}_1,\lambda_j)^\top\mathbf{t}\,.$$

Theorem 2 of Moulines and Soulier (1999) provides a generalisation of central limit theorems for triangular arrays of martingale differences and weakly dependent sequences to similarly weighted sums of correlated variables. Replacing Moulines and Soulier's η_{nj} by $Z_j-E[Z_j]$ and their $b_{n,j}$ by $\mathbf{w}(\boldsymbol{\eta},\lambda_j)^\top\mathbf{t}$, recognising from Corollary A.1 that $Z_j-E[Z_j]$, $j=1,\ldots,\lfloor n/2\rfloor$, share the same moment structure and order of correlation as Moulines and Soulier's η_{nj}, the proof of the lemma follows Moulines and Soulier's proof of their Theorem 2 presented in Moulines and Soulier (1999, Appendix B). Conditions (i) and (ii) of Theorem 2 of Moulines and Soulier (1999) are satisfied because $C_1\lambda_j^{-2d^*}\log\lambda_j\le\|\mathbf{w}(\boldsymbol{\eta},\lambda_j)\|\le C_2\lambda_j^{-2d^*}\log\lambda_j$ for some constants C_1 and C_2 (see Chen and Deo, 2006, expression (21) pg. 276) and

$$\lim_{n\to\infty}\sup_{j=1,\ldots,\lfloor n/2\rfloor}\left(\sum_{j=1}^{\lfloor n/2\rfloor}\left(\mathbf{t}^\top\mathbf{w}(\boldsymbol{\eta}_1,\lambda_j)\right)^2\right)^{-1}(\mathbf{w}(\boldsymbol{\eta}_1,\lambda_j)^\top\mathbf{t})^2=0\,.\ \blacksquare$$

■

The following lemma parallels Lemma 3 of Chen and Deo (2006) and is derived in a similar fashion. The lemma and its proof are presented here for completeness.

Lemma 3 (A.3) *Let* $\mathbb{E}_c$ *denote a compact convex subset of* $\mathbb{E}_\delta^0$ *and denote the second order differential of the FML criterion function by* $d^2Q_n^{(1)}(\boldsymbol{\eta};\mathbf{t})=\mathbf{t}^\top\left(\partial^2Q_n^{(1)}(\boldsymbol{\eta})/\partial\boldsymbol{\eta}\partial\boldsymbol{\eta}^\top\right)\mathbf{t}$. *Then for all* $\mathbf{t}$, $\|\mathbf{t}\|<\infty$,

$$plim_{n\to\infty}\sup_{\boldsymbol{\eta}\in\mathbb{E}_c}\left|d^2Q_n^{(1)}(\boldsymbol{\eta};\mathbf{t})-d^2Q(\boldsymbol{\eta};\mathbf{t})\right|=0\,.$$

under Assumptions $(A.1')$ *and* $(A.2)-(A.5)$.

Proof of Lemma A.2.

By definition of the second order differential we have

$$\begin{aligned}E_0\left[d^2Q_n^{(1)}(\boldsymbol{\eta};\mathbf{t})\right]=&E_0\left[\frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor}\frac{I(\lambda_j)}{f_1(\boldsymbol{\eta},\lambda_j)}\mathbf{t}^\top\mathbf{H}(\boldsymbol{\eta},\lambda_j)\mathbf{t}\right]\\ =&\frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor}\frac{f_0(\lambda_j)}{f_1(\boldsymbol{\eta},\lambda_j)}\mathbf{t}^\top\mathbf{H}(\boldsymbol{\eta},\lambda_j)\mathbf{t}+\left(\frac{E_0[I(\lambda_j)]}{f_0(\lambda_j)}-1\right)\frac{f_0(\lambda_j)}{f_1(\boldsymbol{\eta},\lambda_j)}\mathbf{t}^\top\mathbf{H}(\boldsymbol{\eta},\lambda_j)\mathbf{t}\,,\end{aligned}$$

where $E_0[I(\lambda_j)]/f_0(\lambda_j)-1=O(\log j/j)$, by Corollary A.1, and $\mathbf{t}^\top\mathbf{H}(\boldsymbol{\eta},\lambda_j)\mathbf{t}=O(\log^2\lambda_j)$ since $\sup_{\boldsymbol{\eta}}\partial\log f_1(\boldsymbol{\eta},\lambda_j)/\partial\boldsymbol{\eta}$ is of order $O(\log\lambda_j)$ by Assumptions $(A.2)$ and $(A.3)$ and $\|\mathbf{t}\|<\infty$. Thus we can conclude that

$$\begin{aligned}\frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor}\left(\frac{E_0[I(\lambda_j)]}{f_0(\lambda_j)}-1\right)\frac{f_0(\lambda_j)}{f_1(\boldsymbol{\eta},\lambda_j)}\mathbf{t}^\top\mathbf{H}(\boldsymbol{\eta},\lambda_j)\mathbf{t}=&O\left(\frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor}\frac{\log j}{j}\lambda_j^{-2d^*}\log^2\lambda_j\right)\\ =&O\left(n^{2d^*-1}\sum_{j=1}^{\lfloor n/2\rfloor}\frac{\log j}{j}j^{-2d^*}\log^2(j/n)\right)\\ =&\begin{cases}O(n^{2d^*-1}\log^2 n), & 0<d^*<0.5\,;\\ O(n^{-1}\log^4 n), & -1.0<d^*\le 0\,,\end{cases}\end{aligned}$$

and hence that $E_0\left[d^2Q_n^{(1)}(\boldsymbol{\eta};\mathbf{t})\right]\to\mathbf{t}^\top\frac{\partial^2 Q(\boldsymbol{\eta})}{\partial\boldsymbol{\eta}\partial\boldsymbol{\eta}'}\mathbf{t}=d^2Q(\boldsymbol{\eta};\mathbf{t})$. Similarly, setting $h(\boldsymbol{\eta};\mathbf{t},\lambda_j,\lambda_k)=\mathbf{t}^\top\mathbf{H}(\boldsymbol{\eta},\lambda_j)\mathbf{t}\cdot\mathbf{t}^\top\mathbf{H}(\boldsymbol{\eta},\lambda_k)\mathbf{t}$ and invoking Corollary A.1 once again we have

$$\begin{aligned}Var_0\left[d^2Q_n^{(1)}(\boldsymbol{\eta};\mathbf{t})\right]=&\left(\frac{2\pi}{n}\right)^2\sum_{j=1}^{\lfloor n/2\rfloor}\sum_{k=1}^{\lfloor n/2\rfloor}\frac{f_0(\lambda_j)}{f_1(\boldsymbol{\eta},\lambda_j)}\frac{f_0(\lambda_k)}{f_1(\boldsymbol{\eta},\lambda_k)}h(\boldsymbol{\eta};\mathbf{t},\lambda_j,\lambda_k)Cov_0\left[\frac{I(\lambda_j)}{f_0(\lambda_j)}\frac{I(\lambda_k)}{f_0(\lambda_k)}\right]\\ =&O\left(\frac{1}{n^2}\sum_{j=1}^{\lfloor n/2\rfloor}\sum_{k\ge j}^{\lfloor n/2\rfloor}\lambda_j^{-2d^*}\lambda_k^{-2d^*}\log^2\lambda_j\log^2\lambda_k j^{-2|d_0|}k^{2|d_0|-2}\log^2 k\right)\\ =&O\left(n^{4d^*-2}\sum_{j=1}^{\lfloor n/2\rfloor}j^{-2(d^*+|d_0|)}\log^2(j/n)\sum_{k=1}^{\lfloor n/2\rfloor}k^{-2(d^*-|d_0|)-2}\log^2 k\log^2(k/n)\right)\\ =&\begin{cases}O(n^{4d^*-2}\log^4 n), & d^*+|d_0|>0.5\quad 0<d^*<0.5\,;\\ O(n^{-(1+2(|d_0|-d^*))}\log^5 n), & d^*+|d_0|\le 0.5\quad 0<d^*<0.5\,;\\ O(n^{-(1+2|d_0|)}\log^5 n), & d^*+|d_0|\le 0.5\quad -1<d^*\le 0\,.\end{cases}\end{aligned}$$

It therefore follows from Markov's inequality that $d^2Q_n^{(1)}(\boldsymbol{\eta};\mathbf{t})$ converges in probability to $\mathrm{d}^2Q(\boldsymbol{\eta};\mathbf{t})$.

Now, by the Mean Value Theorem, for any $\boldsymbol{\eta}_1$ and $\boldsymbol{\eta}_2$ in $\mathbb{E}_c$

$$\left|d^2Q_n^{(1)}(\boldsymbol{\eta}_1;\mathbf{t}) - d^2Q_n^{(1)}(\boldsymbol{\eta}_2;\mathbf{t})\right| \leq \left\| \frac{\partial\left\{d^2Q_n^{(1)}(\overline{\boldsymbol{\eta}};\mathbf{t})\right\}}{\partial\boldsymbol{\eta}} \right\| \cdot \|\boldsymbol{\eta}_1 - \boldsymbol{\eta}_2\|$$

for some $\overline{\boldsymbol{\eta}}$ between $\boldsymbol{\eta}_1$ and $\boldsymbol{\eta}_2$. Moreover,

$$\begin{aligned} E_0\left[\frac{\partial\left\{d^2Q_n^{(1)}(\boldsymbol{\eta};\mathbf{t})\right\}}{\partial\boldsymbol{\eta}}\right] =& \frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor}\left\{\frac{f_0(\lambda_j)}{f_1(\boldsymbol{\eta},\lambda_j)} + \left(\frac{E_0[I(\lambda_j)]}{f_0(\lambda_j)} - 1\right)\frac{f_0(\lambda_j)}{f_1(\boldsymbol{\eta},\lambda_j)}\right\}\cdot\mathbf{k}(\boldsymbol{\eta};\mathbf{t},\lambda_j) \\ =& \frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor}\frac{f_0(\lambda_j)}{f_1(\boldsymbol{\eta},\lambda_j)}\mathbf{k}(\boldsymbol{\eta};\mathbf{t},\lambda_j) + \mathbf{r}_n \end{aligned} \tag{18}$$

where

$$\mathbf{k}(\boldsymbol{\eta};\mathbf{t},\lambda_j) = \frac{\partial\mathbf{t}^\top\mathbf{H}(\boldsymbol{\eta},\lambda_j)\mathbf{t}}{\partial\boldsymbol{\eta}} - \mathbf{t}^\top\mathbf{H}(\boldsymbol{\eta},\lambda_j)\mathbf{t}\frac{\partial\log f_1(\boldsymbol{\eta},\lambda_j)}{\partial\boldsymbol{\eta}} = O(\log^3\lambda_j)$$

and the remainder

$$\begin{aligned} \mathbf{r}_n =& \frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor}\left(\frac{E_0[I(\lambda_j)]}{f_0(\lambda_j)} - 1\right)\frac{f_0(\lambda_j)}{f_1(\boldsymbol{\eta},\lambda_j)}\mathbf{k}(\boldsymbol{\eta};\mathbf{t},\lambda_j) \\ =& O\left(\frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor}\frac{\log j}{j}\lambda_j^{-2d^*}\log^3\lambda_j\right) \\ =& \begin{cases} O(n^{2d^*-1}\log^3 n), & 0 < d^* < 0.5\,; \\ O(n^{-1}\log^5 n), & -1 < d^* \leq 0\,, \end{cases} \end{aligned}$$

From Assumption $(A.3)$ and $(A.5)$ it follows that the components of the first term on the right hand side of (18) converge to finite constants, and hence that

$$\left|d^2Q_n^{(1)}(\boldsymbol{\eta}_1;\mathbf{t}) - d^2Q_n^{(1)}(\boldsymbol{\eta}_2;\mathbf{t})\right| \leq C_n\|\boldsymbol{\eta}_1 - \boldsymbol{\eta}_2\|$$

where

$$C_n = \sup_{\boldsymbol{\eta}\in\mathbb{E}_c}\left\|\frac{\partial\left\{d^2Q_n^{(1)}(\boldsymbol{\eta};\mathbf{t})\right\}}{\partial\boldsymbol{\eta}}\right\| = O_p(1)$$

since $\sup_n E_0\left[\partial\left\{d^2Q_n^{(1)}(\boldsymbol{\eta};\mathbf{t})\right\}/\partial\boldsymbol{\eta}\right] < \infty$ for all $\boldsymbol{\eta} \in \mathbb{E}_c$. We can therefore conclude that $d^2Q_n^{(1)}(\boldsymbol{\eta}_1;\mathbf{t})$ is stochastically equicontinuous, and hence that

$$\text{plim}_{n\to\infty} \sup_{\boldsymbol{\eta}\in\mathbb{E}_c} \left|d^2Q_n^{(1)}(\boldsymbol{\eta};\mathbf{t}) - d^2Q(\boldsymbol{\eta};\mathbf{t})\right| = 0\,,$$

for all $\mathbf{t}$, $\|\mathbf{t}\| < \infty$, as required.

That the FML estimator possesses the asymptotic distributions as specified in Theorem 3 now follows by replacing Lemma 5 of Chen and Deo (2006) by Lemma and Corollary A.1, Lemmas 8 and 9 by Lemma A.2, and Lemma 3 of Chen and Deo (2006) by Lemma A.3. Having made these replacements we then find that the convergence rates and asymptotic approximations given in Chen and Deo's Lemma 4 and for their Cases 1, 2 and 3 in their lemmas 6, 7, 10, 11 and 12 remain valid, thus establishing Theorem 3 for the FML estimator.

For the Whittle estimator we have, via definition of the differential and application of the chain rule, that

$$\begin{aligned}\left|dQ_n^{(2)}(\boldsymbol{\eta};\mathbf{t}) - \frac{dQ_n^{(1)}(\boldsymbol{\eta};\mathbf{t})}{Q_n^{(1)}(\boldsymbol{\eta})}\right| \leq& |\nabla Q_n^{(2)}(\boldsymbol{\eta};\mathbf{t}) - dQ_n^{(2)}(\boldsymbol{\eta};\mathbf{t})| + \\ & |\nabla Q_n^{(2)}(\boldsymbol{\eta};\mathbf{t}) - \nabla\log Q_n^{(1)}(\boldsymbol{\eta};\mathbf{t})| + \\ & \left|\nabla\log Q_n^{(1)}(\boldsymbol{\eta};\mathbf{t}) - \frac{dQ_n^{(1)}(\boldsymbol{\eta};\mathbf{t})}{Q_n^{(1)}(\boldsymbol{\eta})}\right| \\ \leq& 2\epsilon\|\mathbf{t}\| + |\nabla Q_n^{(2)}(\boldsymbol{\eta};\mathbf{t}) - \nabla\log Q_n^{(1)}(\boldsymbol{\eta};\mathbf{t})|\end{aligned}$$

where

$$\nabla\log Q_n^{(1)}(\boldsymbol{\eta};\mathbf{t}) = \log Q_n^{(1)}(\boldsymbol{\eta}+\mathbf{t}) - \log Q_n^{(1)}(\boldsymbol{\eta}) \quad \textit{and} \quad \nabla Q_n^{(2)}(\boldsymbol{\eta};\mathbf{t}) = Q_n^{(2)}(\boldsymbol{\eta}+\mathbf{t}) - Q_n^{(2)}(\boldsymbol{\eta})$$

and $\epsilon \to 0$ as $\|\mathbf{t}\| \to 0$. Setting $\|\mathbf{t}\| = O(n^{-1}\log n)$, noting that (5) implies that the difference in differences $|\nabla Q_n^{(2)}(\boldsymbol{\eta};\mathbf{t}) - \nabla\log Q_n^{(1)}(\boldsymbol{\eta};\mathbf{t})| = O(n^{-1}\log n)$, we find that

$$\left|dQ_n^{(2)}(\boldsymbol{\eta};\mathbf{t}) - \frac{dQ_n^{(1)}(\boldsymbol{\eta};\mathbf{t})}{Q_n^{(1)}(\boldsymbol{\eta})}\right| \leq O(n^{-1}\log n)\,. \tag{19}$$

Equation (19) leads, in turn, to the conclusion that

$$\left|d^2Q_n^{(2)}(\boldsymbol{\eta};\mathbf{t}) - \frac{d^2Q_n^{(1)}(\boldsymbol{\eta};\mathbf{t})}{Q_n^{(1)}(\boldsymbol{\eta})}\right| \leq \left\{\frac{dQ_n^{(1)}(\boldsymbol{\eta};\mathbf{t})}{Q_n^{(1)}(\boldsymbol{\eta})}\right\}^2 + O(n^{-1}\log n). \tag{20}$$

But by Lemma 4 of Chen and Deo (2006)

$$\begin{aligned} E\left[dQ_n^{(1)}(\boldsymbol{\eta}_1;\mathbf{t})\right] &= -\frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor} E\left[\frac{I(\lambda_j)}{f_0(\lambda_j)}\right]\mathbf{w}(\boldsymbol{\eta}_1,\lambda_j)^\top\mathbf{t} \\ &= \begin{cases} O(n^{2d^*-1}\log n), & 0 < d^* < 0.5\,; \\ O(n^{-1}\log^3 n), & -1.0 < d^* \leq 0\,. \end{cases} \end{aligned} \tag{21}$$

In addition,

$$\begin{aligned} Var_0\left[dQ_n^{(1)}(\boldsymbol{\eta};\mathbf{t})\right] =& O\left(\frac{1}{n^2}\sum_{j=1}^{\lfloor n/2\rfloor}\sum_{k\geq j}^{\lfloor n/2\rfloor} \lambda_j^{-2d^*}\lambda_k^{-2d^*}\log\lambda_j\log\lambda_k j^{-2|d_0|}k^{2|d_0|-2}\log^2 k\right) \\ =& O\left(n^{4d^*-2}\sum_{j=1}^{\lfloor n/2\rfloor} j^{-2(d^*+|d_0|)}\log(j/n)\sum_{k=1}^{\lfloor n/2\rfloor} k^{-2(d^*-|d_0|)-2}\log^2 k\log(k/n)\right) \\ =& \begin{cases} O(n^{4d^*-2}\log^2 n), & d^*+|d_0| > 0.5 \quad 0 < d^* < 0.5\,; \\ O(n^{-(1-2(d^*-|d_0|))}\log^3 n), & d^*+|d_0| \leq 0.5 \quad 0 < d^* < 0.5\,; \\ O(n^{-(1+2|d_0|)}\log^3 n), & d^*+|d_0| \leq 0.5 \quad -1.0 < d^* \leq 0\,. \end{cases} \end{aligned} \tag{22}$$

The asymptotic equivalence of the FML and Whittle estimators now follows since: by Lemma 3 $Q_n^{(1)}(\boldsymbol{\eta}_1)$ converges almost surely to $Q(\boldsymbol{\eta}_1) \geq \sigma_0^2 > 0$; equations (20), (21) and (22) imply that $|d^2Q_n^{(2)}(\boldsymbol{\eta}_1;\mathbf{t}) - d^2Q_n^{(1)}(\boldsymbol{\eta}_1;\mathbf{t})/Q_n^{(1)}(\boldsymbol{\eta}_1)| = o_p(1)$; and equation (19) implies that

$$\begin{aligned} \frac{n}{2\pi}\left(\sum_{j=1}^{\lfloor n/2\rfloor}\left(\mathbf{t}^\top\mathbf{w}(\boldsymbol{\eta}_1,\lambda_j)\right)^2\right)^{-\frac{1}{2}} & \left|dQ_n^{(2)}(\boldsymbol{\eta}_1;\mathbf{t}) - \frac{dQ_n^{(1)}(\boldsymbol{\eta}_1;\mathbf{t})}{Q_n^{(1)}(\boldsymbol{\eta}_1)}\right| \\ &= \begin{cases} O(n^{-2(d_0-d_1)}), & 0.25 < d_0-d_1 < 0.5\,; \\ O((n\log n)^{-\frac{1}{2}}), & -1.0 < d_0-d_1 \leq 0.25\,. \end{cases} \end{aligned}$$

since

$$\sum_{j=1}^{\lfloor n/2 \rfloor} \left\{ \mathbf{t}^\top \mathbf{w}(\boldsymbol{\eta}, \lambda_j) \right\}^2 = O(\sum_{j=1}^{\lfloor n/2 \rfloor} \lambda_j^{-4d^*} \log^2 \lambda_j) = \begin{cases} O(n^{4d^*} \log^2 n), & 0.25 < d^* < 0.5\,; \\ O(n \log^3 n), & -1.0 < d^* \leq 0.25\,. \end{cases}$$

This establishes that Lemma A.2 also holds with $\mathrm{d}Q_n^{(1)}(\boldsymbol{\eta}_1; \mathbf{t})$ replaced by $Q(\boldsymbol{\eta}_1)\mathrm{d}Q_n^{(2)}(\boldsymbol{\eta}_1; \mathbf{t})$.

For the TML estimator we begin by noting that

$$\left| \frac{2\pi}{n} \sum_{j=1}^{\lfloor n/2 \rfloor} \log f_1(\boldsymbol{\eta}, \lambda_j) - \frac{\pi}{n} \log |\boldsymbol{\Sigma}_\eta| \right| = O(n^{-1} \log n)\,,$$

and concentrating $Q_n^{(2)}(\sigma^2, \boldsymbol{\eta})$ and $Q_n^{(3)}(\sigma^2, \boldsymbol{\eta})$ with respect to σ^2 yields the inequality

$$|Q_n^{(2)}(\boldsymbol{\eta}) - \pi Q_n^{(3)}(\boldsymbol{\eta})| \leq O(n^{-1} \log n) + |\log 2Q_n^{(1)}(\boldsymbol{\eta}) - \log(2\pi/n)\mathbf{Y}^\top \boldsymbol{\Sigma}_\eta^{-1} \mathbf{Y}|\,. \tag{23}$$

If we let $\mathbf{U}$ denote the $n \times n$ unitary matrix with entries $n^{-\frac{1}{2}} \exp(\imath 2\pi(r-1)(c-1)/n)$ in row r and column c, $r, c = 1, \ldots, n$, then the off diagonal entries in $\mathbf{U}\boldsymbol{\Sigma}_\eta\mathbf{U}^*$ are of order $O(n^{-1})$, and the diagonal entries are

$$\sum_{s=-(n-1)}^{n-1} \left(1 - \frac{|s|}{n}\right) \frac{\gamma_1(s)}{\sigma^2} \exp(\imath 2\pi(j-1)s/n) \quad j = 1, \ldots, n\,.$$

Since $f_1(\boldsymbol{\eta}, \lambda)$ is absolutely integrable on $[-\pi, \pi]$, and by Assumptions 3 and 5 $f_1(\boldsymbol{\eta}, \lambda)$ is continuously differentiable for all $\lambda \neq 0$, from Fejer's Theorem it follows that $\mathbf{U}\boldsymbol{\Sigma}_\eta\mathbf{U}^* - \mathbf{F}_1 = O(n^{-1})$ where $\mathbf{F}_1$ equals

$$\begin{cases} \mathrm{diag}(Csf_1, f_1(\boldsymbol{\eta}, \lambda_1) \ldots, f_1(\boldsymbol{\eta}, \lambda_{\lfloor n/2 \rfloor}), f_1(\boldsymbol{\eta}, \lambda_{\lfloor n/2 \rfloor}), \ldots, f_1(\boldsymbol{\eta}, \lambda_1)), & \text{for } n \text{ odd;} \\ \mathrm{diag}(Csf_1, f_1(\boldsymbol{\eta}, \lambda_1) \ldots, f_1(\boldsymbol{\eta}, \lambda_{(n-2)/2}), f_1(\boldsymbol{\eta}, \lambda_{\lfloor n/2 \rfloor}), f_1(\boldsymbol{\eta}, \lambda_{(n-2)/2}), \ldots, f_1(\boldsymbol{\eta}, \lambda_1)), & \text{for } n \text{ even,} \end{cases}$$

and the Cèsaro sum

$$Csf_1 = \sum_{s=-(n-1)}^{n-1} \left(1 - \frac{|s|}{n}\right) \frac{\gamma_1(s)}{\sigma^2} = \begin{cases} O(n^{2d} \log n), & 0 < d < 0.5 \\ O(1), & -0.5 < d \leq 0\,. \end{cases}$$

Conditions A and Assumption A.3 imply that $\boldsymbol{\Sigma}_\eta$ and $\mathbf{F}_1$ are positive definite and it therefore follows, upon application of the Rayleigh-Ritz theorem, that

$$\begin{aligned}(2\pi/n)\left|\mathbf{Y}^\top\boldsymbol{\Sigma}_\eta^{-1}\mathbf{Y}-\mathbf{Y}^\top\mathbf{U}\mathbf{F}_1^{-1}\mathbf{U}^*\mathbf{Y}\right| &= \left|(2\pi/n)\mathbf{Y}^\top\boldsymbol{\Sigma}_\eta^{-1}\mathbf{Y}-2Q_n^{(1)}(\boldsymbol{\eta})\right|\\ &= n^{-1}|\mathbf{Y}^\top\mathbf{R}_\eta\mathbf{Y}|\\ &\leq n^{-1}\max_{i=1,\dots,n}\{|\mu_i(\mathbf{R}_\eta)|\}\|\mathbf{Y}\|^2\end{aligned}$$

where $\mu_i(\mathbf{R}_\eta)$, $i=1,\dots,n$, are the eigenvalues of the residual $\mathbf{R}_\eta=\boldsymbol{\Sigma}_\eta^{-1}-\mathbf{U}\mathbf{F}_1^{-1}\mathbf{U}^*=O(n^{-1})$. Evaluating the characteristic polynomial of $\mathbf{R}_\eta$ via the leading principle minors, or using the Faddeev-Leverrier method, then indicates that $|\mu_i(\mathbf{R}_\eta)|^n \leq |\mu_i(\mathbf{R}_\eta)|^{n-1}O(n^{-1})$ and the spectral radius of $\mathbf{R}_\eta$ is $O(n^{-1})$.

We can therefore use the method leading to (19) and (20) to deduce from the inequality in (23) that the first and second differentials satisfy $|dQ_n^{(2)}(\boldsymbol{\eta}_1;\mathbf{t})-\pi dQ_n^{(3)}(\boldsymbol{\eta}_1;\mathbf{t})|=O(n^{-1}\log n)$ and $|d^2Q_n^{(2)}(\boldsymbol{\eta}_1;\mathbf{t})-\pi d^2Q_n^{(3)}(\boldsymbol{\eta}_1;\mathbf{t})|=o_p(1)$. It therefore follows that the Whittle estimator and the TML estimator converge in distribution as

$$\frac{n}{2\pi}\left(\sum_{j=1}^{\lfloor n/2\rfloor}\left(\mathbf{t}^\top\mathbf{w}(\boldsymbol{\eta}_1,\lambda_j)\right)^2\right)^{-\frac{1}{2}}\left|dQ_n^{(2)}(\boldsymbol{\eta}_1;\mathbf{t})-\pi dQ_n^{(3)}(\boldsymbol{\eta}_1;\mathbf{t})\right| = \begin{cases} O(n^{-2(d_0-d_1)}), & 0.25<d_0-d_1<0.5\,;\\ O((n\log n)^{-\frac{1}{2}}), & -1.0<d_0-d_1\leq 0.25\,.\end{cases}$$

For the CSS estimator we have $Q_n^{(4)}(\boldsymbol{\eta}_1)=\left\{\mathbf{Y}^\top\mathbf{A}_\eta\mathbf{Y}-\mathbf{Y}^\top\mathbf{M}_\eta\mathbf{Y}\right\}/n$. Replacing $\boldsymbol{\Sigma}_\eta$ by $\mathbf{A}_\eta$ and adapting the argument used previously shows that $\mathbf{U}\mathbf{A}_\eta\mathbf{U}^*=2\pi\mathbf{F}_1^{-1}+O(n^{-1})$ and hence, using (14), that

$$|Q_n^{(4)}(\boldsymbol{\eta}_1)-2Q_n^{(1)}(\boldsymbol{\eta}_1)|\leq O(n^{-1})+o_p(n^{-\frac{1}{2}})\,.$$

Apart from notational changes, the remaining steps in showing that the CSS and FML estimators converge in distribution are the same as those used in establishing the equivalence of the FML, Whittle and TML estimators, and are therefore omitted.

The preceding derivations imply that Lemma A.2 also holds with $dQ_n^{(1)}(\boldsymbol{\eta}_1;\mathbf{t})$ replaced by $Q(\boldsymbol{\eta}_1)dQ_n^{(2)}(\boldsymbol{\eta}_1;\mathbf{t})$, $\pi Q(\boldsymbol{\eta}_1)dQ_n^{(3)}(\boldsymbol{\eta}_1;\mathbf{t})$ and $(1/2)dQ_n^{(4)}(\boldsymbol{\eta}_1;\mathbf{t})$. As with the FML estimator, we then find that the convergence rates and asymptotic approximations given in lemmas 4, 6, 7, 10, 11 and 12 of Chen and Deo (2006) remain valid, thus establishing Theorem 3 for the Whittle, TML and CSS estimators, and hence confirming that the four estimators $\widehat{\boldsymbol{\eta}}_1^{(1)}$, $\widehat{\boldsymbol{\eta}}_1^{(2)}$, $\widehat{\boldsymbol{\eta}}_1^{(3)}$ and $\widehat{\boldsymbol{\eta}}_1^{(4)}$ are asymptotically equivalent.

Appendix B: Evaluation of Bias Correction Term

For the FML estimator we have

$$\begin{aligned} E_0\left(\frac{\partial Q_n^{(1)}(\boldsymbol{\eta})}{\partial\boldsymbol{\eta}}\right) &= \frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor} E_0(I(\lambda_j))\frac{\partial f_1(\boldsymbol{\eta},\lambda_j)^{-1}}{\partial\boldsymbol{\eta}} \\ &= \frac{2\pi}{n}\sum_{j=1}^{\lfloor n/2\rfloor}\left(\sum_{|k|<n}\left(1-\frac{|k|}{n}\right)\gamma_0(k)\exp(ik\lambda_j)\right)\frac{\partial f_1(\boldsymbol{\eta},\lambda_j)^{-1}}{\partial\boldsymbol{\eta}}\,, \end{aligned}$$

where $\gamma_0(k)$ denotes the autocovariance at lag k of the TDGP (see, for example, Brockwell and Davis, 1991, Proposition 10.3.1). Similarly, for the Whittle estimator we have

$$\begin{aligned} E_0\left(\frac{\partial Q_n^{(2)}(\sigma_\varepsilon^2,\boldsymbol{\eta})}{\partial\boldsymbol{\eta}}\right) = \frac{4}{n}\sum_{j=1}^{\lfloor n/2\rfloor}\frac{\partial\log f_1(\boldsymbol{\eta}_1,\lambda_j)}{\partial\boldsymbol{\eta}} \\ + \frac{8\pi}{\sigma^2 n}\sum_{j=1}^{\lfloor n/2\rfloor}\left(\sum_{|k|<n}\left(1-\frac{|k|}{n}\right)\gamma_0(k)\exp(ik\lambda_j)\right)\frac{\partial f_1(\boldsymbol{\eta},\lambda_j)^{-1}}{\partial\boldsymbol{\eta}}. \end{aligned}$$

Differentiating the TML criterion function with respect to $\boldsymbol{\eta}$ gives

$$\frac{\partial Q_n^{(3)}(\sigma^2,\boldsymbol{\eta})}{\partial\boldsymbol{\eta}} = \frac{1}{n}\mathrm{tr}\boldsymbol{\Sigma}_\eta^{-1}\frac{\partial\boldsymbol{\Sigma}_\eta}{\partial\boldsymbol{\eta}} + \frac{1}{n\sigma^2}\mathbf{Y}^T\frac{\partial\boldsymbol{\Sigma}_\eta^{-1}}{\partial\boldsymbol{\eta}}\mathbf{Y}\,,$$

which has expectation

$$E_0\left(\frac{\partial Q_n^{(3)}(\sigma^2,\boldsymbol{\eta})}{\partial\boldsymbol{\eta}}\right) = \frac{1}{n}\mathrm{tr}\boldsymbol{\Sigma}_\eta^{-1}\frac{\partial\boldsymbol{\Sigma}_\eta}{\partial\boldsymbol{\eta}} - \frac{1}{n\sigma^2}\mathrm{tr}\boldsymbol{\Sigma}_\eta^{-1}\frac{\partial\boldsymbol{\Sigma}_\eta}{\partial\boldsymbol{\eta}}\boldsymbol{\Sigma}_\eta^{-1}\boldsymbol{\Sigma}_0\,,$$

where $\mathbf{\Sigma}_0 = [\gamma_0 (|i-j|)]$ and $\sigma^2 \mathbf{\Sigma}_\eta = [\gamma_1 (|i-j|)]$, $i, j = 1, 2, ..., n$. The criterion function for the CSS estimator can be re-written as

$$Q_n^{(4)}(\boldsymbol{\eta}) = \frac{1}{n} \sum_{t=1}^{n} \left(\sum_{i=0}^{t-1} \tau_i y_{t-i} \right)^2 = \frac{1}{n} \sum_{t=1}^{n} \sum_{i=0}^{t-1} \sum_{j=0}^{t-1} \tau_i \tau_j y_{t-i} y_{t-j} \, ,$$

where τ_i is as defined in (10). The gradient of $Q_n^{(4)}(\boldsymbol{\eta})$, recalling that $\tau_i = \tau_i(\boldsymbol{\eta})$, is thus given by

$$\frac{\partial Q_n^{(4)}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} = \frac{1}{n} \sum_{t=1}^{n} \sum_{i=0}^{t-1} \sum_{j=0}^{t-1} \left(\tau_i \frac{\partial \tau_j}{\partial \boldsymbol{\eta}} + \tau_j \frac{\partial \tau_i}{\partial \boldsymbol{\eta}} \right) y_{t-i} y_{t-j} \, ,$$

and the expected value of the gradient is

$$E_0 \left(\frac{\partial Q_n^{(4)}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} \right) = \frac{1}{n} \sum_{t=1}^{n} \sum_{i=0}^{t-1} \sum_{j=0}^{t-1} \left(\tau_i \frac{\partial \tau_j}{\partial \boldsymbol{\eta}} + \tau_j \frac{\partial \tau_i}{\partial \boldsymbol{\eta}} \right) \gamma_0(i-j) \, .$$

References

Beran, J. (1994). *Statistics for long-memory processes.* Chapman and Hall, New York, 1st edition.

Brockwell, P. J. and Davis, R. A. (1991). *Time series: Theory and Methods.* Springer, New York, 2nd edition.

Chen, W. W. and Deo, R. S. (2006). Estimation of mis-specified long memory models. *Journal of Econometrics*, **134**(1), 257–281.

Fox, R. and Taqqu, S. M. (1986). Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *The Annals of Statistics*, **14**(2), 517–532.

Gradshtein, I. S. and Ryzhik, I. M. (2007). *Tables of Integrals, Series and Products.* Academic Press, Sydney.

Grenander, U. and Szego, G. (1958). *Toeplitz Forms and Their Application.* University of California Press, Berkeley.

Hannan, E. J. (1973). The asymptotic theory of linear time-series models. *Journal of Applied Probability*, **10**(1), 130–145.

Lahiri, S. N. (2003). A necessary and sufficient condition for asymptotic independence of discrete Fourier transforms under short- and long-range dependence. *Annals of Statistics*, **31**(2), 613–641.

Moulines, E. and Soulier, P. (1999). Broadband log-periodogram regression of time series with long-range dependence. *Annals of Statistics*, **27**(4), 1415–1439.