Deo, Rohit S.; Chen, Willa W.

Working Paper

Estimation of Mis-Specified Long Memory Models

Papers / Humboldt-Universität Berlin, Center for Applied Statistics and Economics (CASE), No. 2004,03

Provided in cooperation with:
Humboldt-Universität Berlin

Suggested citation: Deo, Rohit S.; Chen, Willa W. (2003) : Estimation of Mis-Specified Long Memory Models, Papers / Humboldt-Universität Berlin, Center for Applied Statistics and Economics (CASE), No. 2004,03, http://hdl.handle.net/10419/22177

Estimation of Mis-Specified Long Memory Models

Willa W. Chen * Rohit S. Deo †

November 7, 2003

Abstract: We study the asymptotic behaviour of frequency domain maximum likelihood estimators of mis-specified models of long memory Gaussian series. We show that even if the long memory structure of the time series is correctly specified, mis-specification of the short memory dynamics may result in parameter estimators which are slower than $\sqrt{n}$ consistent. The conditions under which this happens are provided and the asymptotic distribution of the estimators is shown to be non-Gaussian. Conditions under which estimators of the parameters of the mis-specified model have the standard $\sqrt{n}$ consistent and asymptotically normal behaviour are also provided.

1 Introduction

The asymptotic behaviour of maximum likelihood parameter estimators when the model being estimated is mis-specified is often of interest for various reasons. Some interesting questions which arise, including those raised by White (1982), are: Do the estimators still converge to some limit and does this limit have meaning? If the estimators are consistent for some value, are they still asymptotically normal? Is the standard $\sqrt{n}$ rate of convergence still retained? These questions are not just of theoretical interest but also of practical importance. For example, the Efficient Method of Moments (EMM) estimation procedure (Gallant and Tauchen, 1996) estimates the parameters of a correctly specified model whose likelihood can not be written analytically by deliberately estimating a mis-specified model whose likelihood has a simple analytically form. Naturally, properties of the estimators of the true model parameters, such as their rate of convergence, will depend on the properties of the estimators of the mis-specified model.

Though White (1982) considers the consequences of model mis-specification when the data are identically independently distributed, there has also been considerable work in the literature where the data are assumed to follow a time series. Most of the research in this area (see, for example, Taniguchi, 1979) has assumed that the true data generating process of the series is such that the covariances are summable, implying that the series has short memory. A notable exception to this framework is the work by Yajima (1993), in which he considers model mis-specification of long memory time series which have non-summable covariances. Yajima (1993) studies the consequence of fitting short memory Autoregressive Moving Average (ARMA) models to long memory time series. He shows that when the value of the

*Department of Statistics, Texas A&M University, College Station, Texas 77843, USA
†New York University, 44 W. 4'th Street, New York NY 10012, USA

long memory parameter, d, is greater than 0.25, the estimators of the parameters of the fitted ARMA models will converge to some pseudo-true value at a rate which is slower than $\sqrt{n}$ and depends on d. Furthermore, Yajima (1993) shows that in such cases, the limiting distribution will be non-Gaussian. In our paper, we study the asymptotic distribution of estimators of mis-specified long memory models for a long memory time series. More specifically, we assume that the long memory dynamics of the fitted model are specified correctly but that the short memory dynamics are not. If the short memory dynamics are sufficiently mis-specified, we show that the estimators of the fitted model converge to some pseudo-true value at a rate which is slower than $\sqrt{n}$ and the asymptotic distribution is non-Gaussian. This result shows that even correct specification of merely the long memory dynamic need not be enough to guarantee $\sqrt{n}$ rates of convergence of the estimators and an asymptotic Gaussian distribution. We also establish the condition under which the estimators of the mis-specified model will have the usual $\sqrt{n}$ consistent and asymptotically normal behaviour. In the next section, we state our assumptions and the theoretical results that we have obtained.

2 Asymptotic Results

We will assume that we have n observations $X_1, ..., X_n$ from a stationary Gaussian time series with a spectral density given by

$$f_0(\lambda) = \frac{\sigma_0^2}{2\pi} g_0(\lambda) |2\sin(\lambda/2)|^{-2d_0}, \tag{1}$$

where $\sigma_0^2 > 0$, $0 < d_0 < 0.5$ and $g_0(\lambda)$ is a spectral density continuous on $[-\pi, \pi]$, bounded above and bounded away from zero with continuous second derivatives. An example of a spectral density that is of the form (1) is that of an Autoregressive Fractionally Integrated Moving Average (ARFIMA) process. We are interested in the asymptotic properties of estimators of parameters of mis-specified models which are fit to the data from the process given by (1). We will assume that the mis-specified model that is estimated has a spectral density given by

$$f_1^*\left(\boldsymbol{\theta}, \sigma^2, \lambda\right) = \frac{\sigma^2}{2\pi} g_1(\boldsymbol{\beta}, \lambda) |2\sin(\lambda/2)|^{-2d},$$

where $\sigma^2 > 0$, $\boldsymbol{\theta} = \left(d, \boldsymbol{\beta}'\right) \in \boldsymbol{\Theta}$, $\boldsymbol{\Theta} = [\delta, 0.5 - \delta] \times \boldsymbol{\Phi}$ for some $0 < \delta < 0.25$ such that $d_0 \in \boldsymbol{\Theta}$, $\boldsymbol{\Phi}$ is a p dimensional compact convex set and $g_1(\beta, \lambda)$ is a spectral density is a spectral density such that $g_1(\beta, \lambda) \neq g_0(\lambda)$ for all β. Thus, the short memory component $g_0(\lambda)$ of the true spectral density is mis-specified as $g_1(\beta, \lambda)$ in the family of models that is to be estimated. In this paper, we will study the estimator $\hat{\boldsymbol{\theta}} = \left(\hat{d}, \hat{\boldsymbol{\beta}}'\right)'$ of the parameter vector $\boldsymbol{\theta} = \left(d, \boldsymbol{\beta}'\right)'$ obtained by minimising the objective function

$$Q_n(\boldsymbol{\theta}) = \frac{2\pi}{n} \sum_{j=1}^{n/2} \frac{I(\lambda_j)}{f_1(\boldsymbol{\theta}, \lambda_j)}, \tag{2}$$

where $I(\lambda) = (2\pi n)^{-1} \left|\sum_{t=1}^{n} X_t \exp(-i\lambda t)\right|^2$ is the periodogram, $\lambda_j = 2\pi j/n$ are the Fourier frequencies and $f_1(\boldsymbol{\theta}, \lambda) = g_1(\boldsymbol{\beta}, \lambda) |2\sin(\lambda/2)|^{-2d}$. The objective function $Q_n(\boldsymbol{\theta})$ is an approximation to the negative of the exact Gaussian log-likelihood (Whittle 1953, Brockwell and Davis 1996) for estimating the parameters $\boldsymbol{\theta}$. Furthermore, when the model being estimated is a correctly specified ARMA model, the estimator $\hat{\boldsymbol{\theta}}$ has the same asymptotic distribution as the exact Gaussian maximum likelihood estimator and is thus asymptotically efficient. See Chapter 10, Brockwell and Davis (1996). This equivalence of the asymptotic distribution of $\hat{\boldsymbol{\theta}}$ and the exact Gaussian maximum likelihood estimator continues to hold when estimating the parameters of a correctly specified long memory Gaussian time series (Fox and Taqqu, 1986, Dahlhaus, 1989). We need some technical assumptions on $g_1(\boldsymbol{\beta}, \lambda)$ which we state next.

A. 1 $g_1(\boldsymbol{\beta}, \lambda)$ *is thrice differentiable with continuous third derivatives.*

A. 2

$$\inf_{\boldsymbol{\beta}} \inf_{\lambda} g_1(\boldsymbol{\beta}, \lambda) > 0$$

and

$$\sup_{\boldsymbol{\beta}} \sup_{\lambda} g_1(\boldsymbol{\beta}, \lambda) < \infty$$

A. 3

$$\sup_{\lambda} \sup_{\boldsymbol{\beta}} \left| \frac{\partial g_1(\boldsymbol{\beta}, \lambda)}{\partial \beta_i} \right| < \infty, \qquad 1 \leq i \leq p.$$

A. 4

$$\sup_{\lambda} \sup_{\boldsymbol{\beta}} \left| \frac{\partial^2 g_1(\boldsymbol{\beta}, \lambda)}{\partial \beta_i \partial \beta_j} \right| < \infty, \qquad \sup_{\lambda} \sup_{\boldsymbol{\beta}} \left| \frac{\partial^2 g_1(\boldsymbol{\beta}, \lambda)}{\partial \beta_i \partial \lambda} \right| < \infty \qquad 1 \leq i, \ j \leq p.$$

A. 5

$$\sup_{\lambda} \sup_{\boldsymbol{\beta}} \left| \frac{\partial^3 g_1(\boldsymbol{\beta}, \lambda)}{\partial \beta_i \partial \beta_j \partial \beta_k} \right| < \infty, \qquad 1 \leq i, \ j, \ k \leq p.$$

A. 6 $\int_{-\pi}^{\pi} \log g_1(\boldsymbol{\beta}, \lambda)\, d\lambda = 0$ *for all* $\boldsymbol{\beta} \in \boldsymbol{\Theta}$.

A. 7 *There exists a unique vector* $\boldsymbol{\theta}_1 = \left(d_1, \boldsymbol{\beta}_1'\right) \in \boldsymbol{\Theta}$ *which satisfies*

$$\boldsymbol{\theta}_1 = \arg\min_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}),$$

where

$$Q(\boldsymbol{\theta}) = \int_0^{\pi} \frac{f_0(\lambda)}{f_1(\boldsymbol{\theta}, \lambda)} d\lambda$$

and $f_1(\boldsymbol{\theta}, \lambda) = g_1(\boldsymbol{\beta}, \lambda)\left(2\sin(\lambda/2)\right)^{-2d}$.

It is easy to check that assumptions A.1 - A.6 are satisfied by the class of spectral densities of stationary invertible Autoregressive Moving Average (ARMA) processes with roots bounded away from the unit circle. Assumption A.7 assumes that there exists a pseudo-true parameter value θ_1 such that among all the spectral densities $f_1(\theta, \lambda)$ of the mis-specified family, the member $f_1(\theta_1, \lambda)$ is closest to the true spectral density $f_0(\lambda)$ with respect to the distance $Q(\theta)$. Such an assumption is standard in the literature on mis-specified models fit to time series (See, for example, Taniguchi, 1979, white, 1982 and Yajima, 1993). In the literature, the estimator $\hat{\theta}$ is generally shown to converge to this pseudo-true parameter value θ_1 at a $\sqrt{n}$ rate and is proved to be asymptotically normal. However, in the framework studied in this paper, we will show that though $\hat{\theta}$ still converges to θ_1, both its rate of convergence as well as its asymptotic distribution depend on the value $d_0 - d_1$, i.e. on the difference between the true value and the pseudo-true value of the long memory parameter. More specifically, depending on whether $d_0 - d_1$ is greater than 0.25, less than 0.25 or equal to 0.25, we get three different rates of convergence and limiting distributions of $\hat{\theta}$. The difference $d_0 - d_1$ between the true value and the pseudo-true value of the long memory parameter depends on the extent to which the mis-specified short memory component $g_1(\boldsymbol{\beta},\lambda)$ differs from the true short memory component $g_0(\lambda)$. This point is illustrated in the following example.

Assume that the true spectral density is an ARFIMA$(0, d_0, 1)$ given by

$$f_0(\lambda) = \frac{1}{2\pi} |1 + \alpha_0 \exp(i\lambda)|^2 \left(2\sin\frac{\lambda}{2}\right)^{-2d_0}$$

where the Moving Average (MA) parameter is α_0, the long memory parameter is d_0 and the innovation variance σ_0^2 is 1. Suppose that the mis-specified model is an ARFIMA$(0, d, 0)$ model given by

$$f_1^*(\lambda, d) = \frac{\sigma^2}{2\pi}\left(2\sin\frac{\lambda}{2}\right)^{-2d},$$

where $d \in (0, 0.5)$. In this example, the short memory component in the true model is the MA part given by $g_0(\lambda) = |1 + \alpha_0 \exp(i\lambda)|^2$, whereas in the mis-specified model the short memory component is $g_1(\beta, \lambda) \equiv 1$. Thus, the mis-specified model fails to capture the short memory MA component of the true spectral density. Now

$$Q(d) = \int_0^\pi \frac{f_0(\lambda)}{f_1(d,\lambda)} d\lambda = \int_0^\pi |1 + \alpha_0 \exp(i\lambda)|^2 \left(2\sin\frac{\lambda}{2}\right)^{-2(d_0-d)} d\lambda$$

and hence

$$\frac{\partial^2 Q(d)}{\partial d^2} = \int_0^\pi |1 + \alpha_0 \exp(i\lambda)|^2 \left(2\sin\frac{\lambda}{2}\right)^{-2(d_0-d)} \left(\log\left(2\sin\frac{\lambda}{2}\right)\right)^2 d\lambda.$$

Since the second derivative $\frac{\partial^2 Q(d)}{\partial d^2}$, being the integral of a positive function, is trivially positive for all d it follows that $Q(d)$ is a convex function and hence the value of d that minimises $Q(d)$ is found by setting $\frac{\partial \log Q(d)}{\partial d} = 0$. Using expressions for the covariance function of an ARFIMA$(0, d, 0)$ process given on page 522 of Brockwell and Davis (1996), we get

$$\begin{aligned} Q(d) &= \int_0^\pi |1 + \alpha_0 \exp(i\lambda)|^2 \left(2\sin\frac{\lambda}{2}\right)^{-2(d_0-d)} d\lambda \\ &= \frac{\Gamma(1 - 2(d_0 - d))}{\Gamma^2(1 - (d_0 - d))} \left\{1 + \alpha_0^2 + +2\alpha_0 \left(\frac{d_0 - d}{1 - (d_0 - d)}\right)\right\} \end{aligned}$$

and hence, taking the logarithm of $Q(d)$ and letting Ψ denote the di-gamma function, simple calculus shows that

$$\frac{\partial \log Q(d)}{\partial d} = 2\Psi[1 - 2(d_0 - d)] - 2\Psi[1 - (d_0 - d)] - \frac{2\alpha_0}{[1 - (d_0 - d)]^2}\left(1 + \alpha_0^2 + +2\alpha_0\left(\frac{d_0 - d}{1 - (d_0 - d)}\right)\right)^{-1}.$$

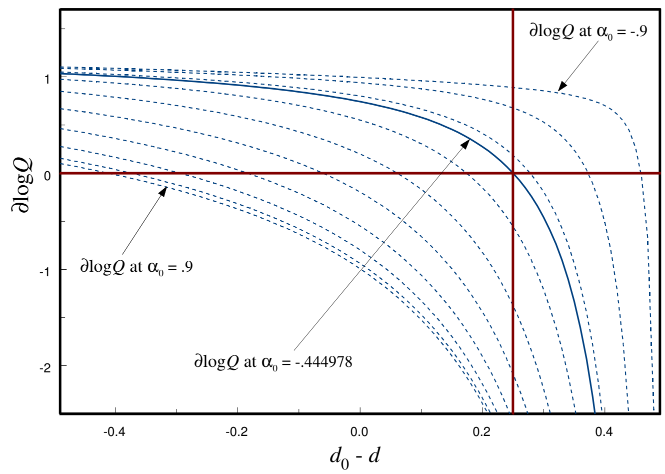

Figure 1: The $\frac{\partial \log Q(d)}{\partial d}$ curves of fitting ARFIMA(0,d,0) to an ARFIMA(0,d,1). The dash lines are $\frac{\partial \log Q(d)}{\partial d}$ at $\alpha_0 = -0.9, \ -0.7, \ \ldots, -0.1, \ 0.1, \ldots, 0.7, \ 0.9$ and the solid line is $\frac{\partial \log Q(d)}{\partial d}$ at $\alpha_0 = -0.444978$.

Noting that $\frac{\partial Q(d)}{\partial d}$ is a function of $\widetilde{d} \equiv d_0 - d$, we plot in Figure 1 the function $\frac{\partial \log Q(d)}{\partial d}$ as a function of $\widetilde{d} \in (-0.5, 0.5)$ for α_0 taking values $\{-0.9, \ -0.7, \ \ldots, -0.1, \ 0.1, \ldots, 0.7, \ 0.9\} \cup \{-0.444978\}$. The vertical line in the plot is drawn at $\widetilde{d} = 0.25$ whereas the horizontal line marks the origin. The curve which intersects the horizontal zero at exactly $\widetilde{d} = 0.25$ corresponds to $\alpha_0 = -0.444978$. We now make some key observations which will help us to understand the nature of the zeroes of $\frac{\partial Q(d)}{\partial d}$ for all values of α_0 from this plot. First, observe that for every fixed α_0, $\frac{\partial \log Q(d)}{\partial d}$ is a decreasing function in $\widetilde{d}$ due to its convexity in d. Also, from elementary calculus, it is seen that for any fixed $\widetilde{d}$, $\frac{\partial \log Q(d)}{\partial d}$ is a decreasing function of α_0 for all $|\alpha_0| < 1$. Thus, the curve which is furthest to the left corresponds to $\alpha_0 = -0.9$ while the curve furthest to the right corresponds to $\alpha_0 = 0.9$. From these remarks it follows that for any $\alpha_0 < -0.444978$, the zeroes of $\frac{\partial \log Q(d)}{\partial d}$ will occur at $\widetilde{d} > 0.25$, whereas for any $-0.444978 < \alpha_0 < 1$ the zeroes of $\frac{\partial \log Q(d)}{\partial d}$ will occur at $\widetilde{d} < 0.25$. Thus, this example illustrates that if the true spectral density is an ARFIMA$(0, d_0, 1)$ and if the misspecified model is chosen to be an ARFIMA$(0, d, 0)$, then the resultant pseudo-true long memory parameter d_1 which satisfies A.7 will be such that $d_0 - d_1 > 0.25$ if the true MA parameter has value less then -0.444978, $d_0 - d_1 = 0.25$ if the true MA parameter has value equal to -0.444978 and $d_0 - d_1 < 0.25$ if the true MA parameter has value greater than -0.444978.

A slightly more complicated example can also be given, where the true spectral density is an ARFIMA$(0, d_0, 1)$ and the mis-specified model is an ARFIMA$(1, d, 0)$. In this example, the short memory component of the true model is an MA of order 1, whereas the short memory component of the mis-specified model is an AutoRegressive (AR) model of order 1. It can be shown that $d_0 - d_1 > 0.25$ if the true MA parameter has value less then -0.637014, $d_0 - d_1 = 0.25$ if the true MA parameter has value equal to -0.637014 and $d_0 - d_1 < 0.25$ if the true MA parameter has value greater than -0.637014. Thus, for this value of the MA parameter, the AR mis-specification is not too serious and we get $d_0 - d_1 < 0.25$. As we shall see, the limiting distribution of $\hat{\boldsymbol{\theta}}$ depends on the value of d_1 which in turn depends on the degree to which the fitted model is mis-specified.

We now present our main results on the asymptotic distribution of $\hat{\boldsymbol{\theta}}$.

Theorem 1 *Assume that $f_0(\lambda)$ and the family $f_1(\theta,\lambda)$ satisfy the assumptions A.1 -A.7 stated in this section and $d_0 - d_1 > 0.25$, where d_1 is defined in A.7. Define the matrix $\mathbf{B} = (b_{i,j})$ by*

$$b_{i,j} = -2\int_{-\pi}^{\pi} \frac{f_0(\lambda)}{f_1^3(\boldsymbol{\theta}_1,\lambda)} \frac{\partial f_1(\boldsymbol{\theta}_1,\lambda)}{\partial \theta_i} \frac{\partial f_1(\boldsymbol{\theta}_1,\lambda)}{\partial \theta_j} + \int_{-\pi}^{\pi} \frac{f_0(\lambda)}{f_1^2(\boldsymbol{\theta}_1,\lambda)} \frac{\partial^2 f_1(\boldsymbol{\theta}_1,\lambda)}{\partial \theta_i \partial \theta_j}.$$

$d^* = d_0 - d_1$. *Then*

$$\frac{n^{1-2d^*}}{\log n}\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_1 - \boldsymbol{\mu}_n\right) \xrightarrow{D} \mathbf{B}^{-1}\left(\sum_{j=1}^{\infty} W_j, 0, \ldots, 0\right)',$$

where $\boldsymbol{\mu}_n = \mathbf{B}^{-1}E\left[Q_n(\boldsymbol{\theta})\right]$, $\boldsymbol{\mu}_n \xrightarrow{P} 0$ and

$$W_j = \frac{(2\pi)^{1-2d_*} g_0(\boldsymbol{\alpha}_0, 0)}{j^{2d_*} g_1(\boldsymbol{\beta}_1, 0)}\left[\left(A_j^2 + B_j^2\right) - E\left(A_j^2 + B_j^2\right)\right],$$

where $\{A_j, B_k\}_{j,k=1}^{\infty}$ are a sequence of normal random variables with mean zero and

$$\text{cov}(A_j, A_k) = \mathcal{I}(\phi_{A,j}, \phi_{A,k}), \quad \text{cov}(B_j, B_k) = \mathcal{I}(\phi_{B,j}, \phi_{B,k}), \ \text{cov}(A_j, B_k) = \mathcal{I}(\phi_{A,j}, \phi_{B,k}),$$

where

$$\phi_{A,j}(u) = \sin(2\pi j u), \qquad \phi_{B,k}(u) = \sin(2\pi k u)$$

and

$$\mathcal{I}(\phi_1, \phi_2) = \iint_{[0,1]^2} \{\phi_1(x)\phi_2(y) + \phi_1(y)\phi_2(x)\} |x-y|^{2d_0-1} dxdy.$$

There are several elements in the result that we obtain in Theorem 1 that are quite non-standard compared to results that one generally obtains for the asymptotic distribution of parameter estimators. Firstly, the rate of convergence of the estimators is slower than $\sqrt{n}$ and can actually be arbitrarily close to zero depending on the value of d^*. Secondly, the asymptotic distribution of the estimators is degenerate in the sense that all the different parameters' estimators converge to multiples of the same limit random variable. This happens due to the fact that the vector of derivatives of the objective function $Q_n(\boldsymbol{\theta})$ is dominated by one random variable. Thirdly, the asymptotic distribution of the estimators is not Gaussian. These results are similar in spirit to those obtained by Yajima (1993), who showed that if a short memory ARMA process were fit to a long memory series with memory parameter $d > 0.25$, the resultant estimators would be n^{1-2d} consistent with non-Gaussian limiting distributions. Our result shows that this continues to be the case even when the mis-specified model has a long memory component, as long as the short memory component is sufficiently ill specified. Fourthly, Theorem 1 implies that the asymptotic bias μ_n of $\hat{\boldsymbol{\theta}}$, though asymptotically negligible, converges to zero at the same rate as the standard deviation of $\hat{\theta}$. This happens due to the fact that we are using an objective function that is a discretised sum over Fourier frequencies and the rate at which the discrete sum approaches the limit integral is slow. We conjecture that if we were to use a slightly modified version of the objective function $Q_n(\boldsymbol{\theta})$, which used integrals instead of discrete sums, we might be able to eliminate the bias term μ_n. However, a distinct advantage of the discrete sum is that it is mean invariant whereas the integral version is not.

Our next Theorem states the asymptotic distribution of $\hat{\boldsymbol{\theta}}$ when the short memory component is not sufficiently mis-specified, resulting in a value of d^* that is less than 0.25.

Theorem 2 *Assume that $f_0(\lambda)$ and the family $f_1(\theta,\lambda)$ satisfy the assumptions A.1 -A.7 and that $d^* = d_0 - d_1 < 0.25$, where d_1 is defined in A.7. Then*

$$n^{1/2}\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_1\right) \stackrel{D}{\rightarrow} N\left(\mathbf{0}, \boldsymbol{\Sigma}\right),$$

where $\boldsymbol{\Sigma} = \mathbf{B}^{-1}\Lambda\mathbf{B}^{-1}$, and $\mathbf{B}$ is as in Theorem 1.

$$\Lambda{=}2\pi\int_0^{\pi}\left(\frac{f_0(\lambda)}{f_1(\boldsymbol{\theta}_1,\lambda)}\right)^2\left(\frac{\partial\log f_1(\boldsymbol{\theta}_1,\lambda)}{\partial\boldsymbol{\theta}}\right)\left(\frac{\partial\log f_1(\boldsymbol{\theta}_1,\lambda)}{\partial\boldsymbol{\theta}}\right)'d\lambda.$$

Theorem 2 shows that when $d_0 - d_1 < 0.25$, the estimators of the parameters of the mis-specified model are $\sqrt{n}$ consistent and asymptotically normal. Our final theoretical result, stated in the following Theorem, states the asymptotic distribution of $\hat{\theta}$ for the "borderline" case when d^* =0.25.

Theorem 3 *Assume that $f_0(\lambda)$ and the family $f_1(\theta,\lambda)$ satisfy the assumptions A.1 -A.7 and that $d^* = d_0 - d_1 = 0.25$, where d_1 is defined in A.7. Then*

$$n\left(\sum_{j=1}^{n/2}\left(\frac{f_0(\lambda_j)}{f_1(\boldsymbol{\theta}_1,\lambda_j)}\frac{\partial\log f_1(\boldsymbol{\theta}_1,\lambda)}{\partial d}\right)^2\right)^{-1/2}\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_1\right) \stackrel{D}{\rightarrow} \mathbf{B}^{-1}(Z, 0, \ldots 0)',$$

where Z is a standard normal random variable and $\mathbf{B}$ is as in Theorem 1. Furthermore,

$$n\left(\sum_{j=1}^{n/2}\left(\frac{f_0(\lambda_j)}{f_1(\boldsymbol{\theta}_1,\lambda_j)}\frac{\partial\log f_1(\boldsymbol{\theta}_1,\lambda)}{\partial d}\right)^2\right)^{-1/2} \propto \left(\frac{n}{\log^3 n}\right)^{1/2}.$$

The above result shows that when $d^* = 0.25$, the estimator $\hat{\boldsymbol{\theta}}$ falls short of $\sqrt{n}$ consistency by a logarithmic rate, though asymptotic normality is still retained.

3 Appendix

Throughout the Appendix, when we are deriving the asymptotic distribution of $\hat{\boldsymbol{\theta}}$, we will use the fact that $\hat{\boldsymbol{\theta}}$ is the minimiser of (2) and appeal to the Taylor series expansion about $\boldsymbol{\theta}_1$,

$$0 = \frac{\partial Q_n(\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}} + \frac{\partial^2 Q_n(\grave{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_1\right), \tag{3}$$

where $\grave{\theta}$ lies between $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}_1$.

Proof of Theorem 1: Note that, by (3)

$$\begin{aligned}\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_1\right) &= -\left(\frac{\partial^2 Q_n(\grave{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right)^{-1}\frac{\partial Q_n(\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}} \\ &= -\left(\frac{\partial^2 Q_n(\grave{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right)^{-1}\left(\frac{\partial Q_n(\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}} - E\frac{\partial Q_n(\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}}\right) - \left(\frac{\partial^2 Q_n(\grave{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right)^{-1}E\frac{\partial Q_n(\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}}.\end{aligned}$$

Hence

$$\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_1\right)+\left(\frac{\partial^2 Q_n(\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right)^{-1}E\frac{\partial Q_n\left(\boldsymbol{\theta}_1\right)}{\partial\boldsymbol{\theta}}=-\left(\frac{\partial^2 Q_n(\dot{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right)^{-1}\left(\frac{\partial Q_n\left(\boldsymbol{\theta}_1\right)}{\partial\boldsymbol{\theta}}-E\frac{\partial Q_n\left(\boldsymbol{\theta}_1\right)}{\partial\boldsymbol{\theta}}\right)$$
$$-\left\{\left(\frac{\partial^2 Q_n(\dot{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right)^{-1}-\left(\frac{\partial^2 Q_n(\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right)^{-1}\right\}E\frac{\partial Q_n\left(\boldsymbol{\theta}_1\right)}{\partial\boldsymbol{\theta}}.$$

The second term on the right hand side is $o_p\left(\log n/n^{1-2d^*}\right)$ by Corollary 2. The Theorem follows from Lemmas 3, 6 and 7. □

Proof of Theorem 2: Note that

$$n^{1/2}E\left(Q_n\left(\boldsymbol{\theta}_1\right)\right)=o\left(1\right)$$

by Lemma 4. The Theorem follows from Lemmas 3, 9 and equation (3). □

Proof of Theorem 3: Note that

$$n\left(\sum_{j=1}^{n/2}\left(\frac{f_0\left(\lambda_j\right)}{f_1\left(\boldsymbol{\theta}_1,\lambda_j\right)}\frac{\partial\log f_1\left(\boldsymbol{\theta}_1,\lambda\right)}{\partial d}\right)^2\right)^{-1/2}E\left(Q_n\left(\boldsymbol{\theta}_1\right)\right)=o\left(1\right)$$

by Lemma 4 and 10. The Theorem follows from Lemmas 3, 10, 11, 12 and equation (3). □

It is convenient to write the normalised periodogram as,

$$\frac{I\left(\lambda_j\right)}{f_0\left(\lambda_j\right)}=A_{n,j}^2+B_{n,j}^2, \tag{4}$$

where

$$A_{n,j}=\frac{1}{\sqrt{2\pi n}}\frac{\sum_{t=1}^n X_t\cos(\lambda_j)}{f_0^{1/2}\left(\lambda_j\right)}\qquad\text{and}\qquad B_{n,j}=\frac{1}{\sqrt{2\pi n}}\frac{\sum_{t=1}^n X_t\sin(\lambda_j)}{f_0^{1/2}\left(\lambda_j\right)}.$$

Let $\xi_{n,j}=A_{n,j}$ or $B_{n,j}$. It was shown in Lemma 4 of Moulines and Soulier (1999) that

$$\operatorname{cov}\left(\xi_{n,j},\xi_{n,k}\right)=O\left(j^{-d_0}k^{d_0-1}\log k\right) \tag{5}$$

for $1\leq j<k\leq n/2$. This bound yields the following Lemma.

Lemma 1 $\operatorname{cov}\left(\frac{I(\lambda_j)}{f_0(\lambda_j)},\frac{I(\lambda_k)}{f_0(\lambda_k)}\right)=O\left(j^{-2d_0}k^{2d_0-2}\log^2 k\right)$ *for* $1\leq j<k\leq n/2$.

Proof. Since $\left(\xi_{n,j},\xi_{n,j},\xi_{n,k},\xi_{n,k}\right)'$ is a normal random vector, we have

$$\operatorname{cov}\left(\xi_{n,j}^2,\xi_{n,k}^2\right)=2\operatorname{cov}\left(\xi_{n,j},\xi_{n,k}\right)\operatorname{cov}\left(\xi_{n,j},\xi_{n,k}\right)=O\left(j^{-2d_0}k^{2d_0-2}\log^2 k\right),$$

by Isserlis' formula (1918) and the lemma follows.□

Lemma 2 *Under assumptions A.1 - 3,*

$$Q_n\left(\boldsymbol{\theta}\right)\xrightarrow{P}Q\left(\boldsymbol{\theta}\right),$$

uniformly in $\boldsymbol{\theta}\in\boldsymbol{\Theta}$.

Proof. We will prove this lemma by verifying the two conditions of Lemma 5.5.5 in Fuller (1996), (i) $Q_n(\boldsymbol{\theta}) - Q(\boldsymbol{\theta})$, for each $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ and (ii) There exists a sequence of positive random variables $\{L_n\}$ and L such that for $\boldsymbol{\theta}_a, \boldsymbol{\theta}_b \in \boldsymbol{\Theta}$, $|Q_n(\boldsymbol{\theta}_a) - Q_n(\boldsymbol{\theta}_b)| \leq \|\boldsymbol{\theta}_a - \boldsymbol{\theta}_b\| L_n$ and $|Q(\boldsymbol{\theta}_a) - Q(\boldsymbol{\theta}_b)| \leq \|\boldsymbol{\theta}_a - \boldsymbol{\theta}_b\| L$, where L_n and L are $O_p(1)$. We now verify the first condition. Note that

$$\begin{aligned} E(Q_n(\boldsymbol{\theta})) &= \frac{2\pi}{n}\sum_{j=1}^{[n/2]} E\left(\frac{I(\lambda_j)}{f_1(\boldsymbol{\theta},\lambda_j)}\right) = \frac{2\pi}{n}\sum_{j=1}^{[n/2]} E\left(\frac{I(\lambda_j)}{f_0(\lambda_j)}\right)\frac{f_0(\lambda_j)}{f_1(\boldsymbol{\theta},\lambda_j)} \\ &= \frac{2\pi}{n}\sum_{j=1}^{[n/2]}\frac{f_0(\lambda_j)}{f_1(\boldsymbol{\theta},\lambda_j)} + \frac{2\pi}{n}\sum_{j=1}^{[n/2]}\left\{E\left(\frac{I(\lambda_j)}{f_0(\lambda_j)}\right) - 1\right\}\frac{f_0(\lambda_j)}{f_1(\boldsymbol{\theta},\lambda_j)} \longrightarrow Q(\boldsymbol{\theta}), \end{aligned}$$

since by Lemma 6 of Moulines and Soulier (1999),

$$\begin{aligned} &\frac{2\pi}{n}\sum_{j=1}^{[n/2]}\left\{E\left(\frac{I(\lambda_j)}{f_0(\lambda_j)}\right) - 1\right\}\frac{f_0(\lambda_j)}{f_1(\boldsymbol{\theta},\lambda_j)} \\ &= O\left(\frac{2\pi}{n}\sum_{j=1}^{[n/2]}\left\{\frac{\log j}{j}\right\}\frac{f_0(\lambda_j)}{f_1(\boldsymbol{\theta},\lambda_j)}\right) = O\left(\frac{2\pi}{n}\sum_{j=1}^{[n/2]}\frac{\log j}{j}\lambda_j^{-2(d_0-d)}\right) = o(1). \end{aligned} \tag{6}$$

Condition (i) will follow if

$$\operatorname{var}(Q_n(\boldsymbol{\theta})) \longrightarrow 0.$$

We write

$$Q_n(\boldsymbol{\theta}) = \frac{2\pi}{n}\sum_{j=1}^{n}\frac{I(\lambda_j)}{f_0(\lambda_j)}\frac{f_0(\lambda_j)}{f_1(\boldsymbol{\theta},\lambda_j)}.$$

Using Lemma 1, we have

$$\begin{aligned} \operatorname{var}(Q_n(\boldsymbol{\theta})) &= \left(\frac{2\pi}{n}\right)^2\sum_{j,k=1}^{[n/2]}\left(\frac{f_0(\lambda_j)f_0(\lambda_k)}{f_1(\boldsymbol{\theta},\lambda_j)f_1(\boldsymbol{\theta},\lambda_k)}\right)\operatorname{cov}\left(\frac{I(\lambda_j)}{f_0(\lambda_j)},\frac{I(\lambda_k)}{f_0(\lambda_k)}\right) \\ &= O\left(\frac{1}{n^2}\sum_{j\leq k=1}^{n/2}\lambda_j^{-2(d_0-d)}\lambda_k^{-2(d_0-d)}j^{-2d_0}k^{2d_0-2}\log^2 k\right) \\ &= O\left(n^{4(d_0-d)-2}\sum_{j=1}^{n/2}j^{-2(d_0-d)-2d_0}\sum_{k>j}k^{-2(d_0-d)+2d_0-2}\log^2 k\right) \\ &= O\left(n^{4(d_0-d)-2-2(d_0-d)-2d_0+1}\mathbf{1}_{\left\{d<-\frac{1}{2}+2d_0\right\}} + n^{4(d_0-d)-2}\mathbf{1}_{\left\{d>-\frac{1}{2}+2d_0\right\}}\right) \\ &= O\left(n^{-2d-1}\mathbf{1}_{\left\{d<-\frac{1}{2}+2d_0\right\}} + n^{4(d_0-d)-2}\log n\mathbf{1}_{\left\{d\geq-\frac{1}{2}+2d_0\right\}}\right) \to 0. \end{aligned} \tag{7}$$

Next we verify condition (ii). Now

$$\partial Q_n(\boldsymbol{\theta}) = -\frac{2\pi}{n}\sum_{j=1}^{[n/2]}\frac{I(\lambda_j)}{f_1^2(\boldsymbol{\theta},\lambda_j)}\partial f_1(\boldsymbol{\theta},\lambda_j) = -\frac{2\pi}{n}\sum_{j=1}^{[n/2]}\frac{I(\lambda_j)}{f_0(\lambda_j)}\frac{f_0(\lambda_j)}{f_1(\boldsymbol{\theta},\lambda_j)}\frac{\partial f_1(\boldsymbol{\theta},\lambda_j)}{f_1(\boldsymbol{\theta},\lambda_j)},$$

where $\sup_{\boldsymbol{\theta}} \left|\partial f_1(\boldsymbol{\theta}, \lambda_j) / f_1(\boldsymbol{\theta}, \lambda_j)\right|$ is of $O(\log \lambda_j)$ by assumption A. 3. Thus there exist a $\widetilde{d}$, $0 < \widetilde{d} < 0.5$ such that

$$E \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \left|\partial Q_n(\boldsymbol{\theta})\right| = O\left(\frac{2\pi}{n} \sum_{j=1}^{[n/2]} \lambda_j^{-2d_0} \lambda_j^{2\widetilde{d}} \log \lambda_j\right) = O(1). \tag{8}$$

For any $\boldsymbol{\theta}_a, \boldsymbol{\theta}_b \in \boldsymbol{\Theta}$, we have

$$Q_n(\boldsymbol{\theta}_a) - Q_n(\boldsymbol{\theta}_b) = (\boldsymbol{\theta}_a - \boldsymbol{\theta}_b)' \partial Q_n\left(\boldsymbol{\theta}^{\dagger}\right),$$

where $\boldsymbol{\theta}^{\dagger}$ lies between $\boldsymbol{\theta}_a$ and $\boldsymbol{\theta}_b$. Hence

$$\left|Q_n(\boldsymbol{\theta}_a) - Q_n(\boldsymbol{\theta}_b)\right| \leq \|\boldsymbol{\theta}_a - \boldsymbol{\theta}_b\| L_n,$$

where $L_n \leq \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \left|\partial Q_n(\boldsymbol{\theta})\right| = O_p(1)$ by (8). Similarly,

$$\left|Q(\boldsymbol{\theta}_a) - Q(\boldsymbol{\theta}_b)\right| \leq \|\boldsymbol{\theta}_a - \boldsymbol{\theta}_b\| L,$$

where $L \leq \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \left|\partial Q(\boldsymbol{\theta})\right| = O(1)$ by assumption A.3. The proof is completed by Lemma 5.5.5 of Fuller (1996). □

Corollary 1 *Under the assumptions of Lemma 2,* $\widehat{\boldsymbol{\theta}} \xrightarrow{p} \boldsymbol{\theta}_1$.

Proof. See (ii) of Lemma 5.5.1 in Fuller (1996).□

Lemma 3 *Under assumptions A.1 -A. 5,*

$$\partial^2 Q_n(\boldsymbol{\theta}) \xrightarrow{P} \partial^2 Q(\boldsymbol{\theta}),$$

uniformly in $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, *where* $\boldsymbol{\Theta}$ *is a convex compact subset.*

Proof. The proof of this lemma follows along the same lines as the proof of Lemma 2. We show that

$$\partial^2 Q_n(\boldsymbol{\theta}) \xrightarrow{P} \partial^2 Q(\boldsymbol{\theta}), \text{ for each } \boldsymbol{\theta} \in \boldsymbol{\Theta}. \tag{9}$$

and

$$\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \frac{\partial}{\partial \boldsymbol{\theta}} \left[\partial^2 Q_{n,ij}(\boldsymbol{\theta})\right] = O_p(1) \tag{10}$$

where $\partial^2 Q_{n,ij}(\boldsymbol{\theta})$ denote the (i,j)th entry of $\partial^2 Q_n(\boldsymbol{\theta})$. It can be shown, by assumption A.4, that

$$\partial^2 Q_n(\boldsymbol{\theta}) = \frac{2\pi}{n} \sum_{j=1}^{[n/2]} \frac{I(\lambda_j)}{f_0(\lambda_j)} \frac{f_0(\lambda_j)}{f_1(\boldsymbol{\theta}, \lambda_j)} \mathbf{H}(\boldsymbol{\theta}, \lambda_j), \tag{11}$$

where $\mathbf{H}(\boldsymbol{\theta}, \lambda_j)$ is a matrix function such that $\|\mathbf{H}(\boldsymbol{\theta}, \lambda_j)\| = O\left(\log^2 \lambda_j\right)$. Together with the fact that $E(I(\lambda_j)) / f_0(\lambda_j) - 1 = O(\log j / j)$ by lemma 6 of Moulines and Soulier (1999), we have

$$E\left[\partial^2 Q_n(\boldsymbol{\theta})\right] \longrightarrow \partial^2 Q(\boldsymbol{\theta}).$$

Furthermore it can be shown, by Lemma 1, that

$$E\left\|\partial^2 Q_n(\boldsymbol{\theta}) - E\partial^2 Q_n(\boldsymbol{\theta})\right\|^2 = \text{trace}\left[\text{cov}\left(\text{vec}\,\partial^2 Q_n(\boldsymbol{\theta})\right)\right] \to 0.$$

We skip the proof of the above equation since it is similar to the proof of (7). Thus (9) is established. Now

$$\frac{\partial}{\partial\boldsymbol{\theta}}\left[\partial^2 Q_{n,ij}(\boldsymbol{\theta})\right] = \frac{2\pi}{n}\sum_{j=1}^{[n/2]}\frac{I(\lambda_j)}{f_0(\lambda_j)}\frac{f_0(\lambda_j)}{f_1(\boldsymbol{\theta},\lambda_j)}\mathbf{h}(\boldsymbol{\theta},\lambda_j),$$

where $\mathbf{h}(\boldsymbol{\theta},\lambda_j)$ is a vector function such that $\|\mathbf{h}(\boldsymbol{\theta},\lambda_j)\| = O\left(\log^3\lambda_j\right)$ by assumption A.5. Thus,

$$E\sup\frac{\partial}{\partial\boldsymbol{\theta}}\left[\partial^2 Q_{n,ij}(\boldsymbol{\theta})\right] = O\left(\frac{2\pi}{n}\sum_{h=1}^{[n/2]}\lambda_h^{-2d}\lambda_h^{2\widetilde{d}}\log^3\lambda_h\right) = O_p(1),$$

where $0 < \widetilde{d} < 0.5$. By assumption A.1 and A.5, for any $\boldsymbol{\theta}_a, \boldsymbol{\theta}_b \in \boldsymbol{\Theta}$,

$$\partial^2 Q_{n,ij}(\boldsymbol{\theta}_a) - \partial^2 Q_{n,ij}(\boldsymbol{\theta}_b) = (\boldsymbol{\theta}_a - \boldsymbol{\theta}_b)'\frac{\partial}{\partial\boldsymbol{\theta}}\left[\partial^2 Q_{n,ij}(\boldsymbol{\theta})\right],$$

where $\boldsymbol{\theta}^\dagger$ lies between $\boldsymbol{\theta}_a$ and $\boldsymbol{\theta}_b$. We have

$$\left|\partial^2 Q_n(\boldsymbol{\theta}_a) - \partial^2 Q_n(\boldsymbol{\theta}_b)\right| \le \|\boldsymbol{\theta}_a - \boldsymbol{\theta}_b\| K_n,$$

where

$$K_n \le \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\sum_{i,j}\left\|\frac{\partial}{\partial\boldsymbol{\theta}}\left[\partial^2 Q_{n,ij}(\boldsymbol{\theta})\right]\right\| = O_p(1).$$

Similarly,

$$\left|\partial^2 Q(\boldsymbol{\theta}_a) - \partial^2 Q(\boldsymbol{\theta}_b)\right| \le \|\boldsymbol{\theta}_a - \boldsymbol{\theta}_b\| K,$$

where

$$K \le \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\sum_{i,j}\left\|\frac{\partial}{\partial\boldsymbol{\theta}}\left[\partial^2 Q_{ij}(\boldsymbol{\theta})\right]\right\|.$$

Thus, the two conditions in Lemma 5.5.5 of Fuller are shown and the result is proved. □

Lemma 4 *Under assumptions of Lemma 2,*

$$E\partial Q_n(\boldsymbol{\theta}_1) = \begin{cases} O\left(n^{2d^*-1}\log n\right), & 0 < d^* < 0.5 \\ O\left(n^{-1}\log^4 n\right), & -0.5 < d^* \le 0 \end{cases}.$$

Proof. Since

$$\partial Q_n(\boldsymbol{\theta}) = -\frac{2\pi}{n}\sum_{j=1}^{[n/2]}\frac{I(\lambda_j)}{f_1^2(\boldsymbol{\theta},\lambda_j)}\partial f_1(\boldsymbol{\theta},\lambda_j) = -\frac{2\pi}{n}\sum_{j=1}^{[n/2]}\frac{I(\lambda_j)}{f_0(\lambda_j)}f_0(\lambda_j)\,\partial f_1^{-1}(\boldsymbol{\theta}_1,\lambda_j),$$

we have

$$E\partial Q_n(\boldsymbol{\theta}_1) = -\frac{2\pi}{n}\sum_{j=1}^{[n/2]}f_0(\lambda_j)\,\partial f_1^{-1}(\boldsymbol{\theta}_1,\lambda_j) - \frac{2\pi}{n}\sum_{j=1}^{[n/2]}E\left(\frac{I(\lambda_j)}{f_0(\lambda_j)} - 1\right)f_0(\lambda_j)\,\partial f_1^{-1}(\boldsymbol{\theta}_1,\lambda_j). \qquad (12)$$

By Lemma 6 of Moulines and Soulier (1999), the second term of (12) is

$$O\left(\frac{2\pi}{n}\sum_{j=1}^{[n/2]}\frac{\log j}{j}f_0\left(\lambda_j\right)\partial f_1^{-1}\left(\boldsymbol{\theta}_1,\lambda_j\right)\right)=O\left(\frac{2\pi}{n}\sum_{j=1}^{[n/2]}\frac{\log j}{j}\lambda_j^{-2d^*}\log\lambda_j\right)$$
$$=\begin{cases} O\left(n^{2d^*-1}\log n\right), & 0<d^*<0.5\\ O\left(n^{-1}\log^4 n\right), & -0.5<d^*\leq 0\end{cases}.$$

Since $\partial Q\left(\boldsymbol{\theta}_1\right)=0$, the first term of (12) is

$$\begin{aligned}
E\partial Q_n\left(\boldsymbol{\theta}_1\right) &= -\frac{2\pi}{n}\sum_{j=1}^{[n/2]} f_0\left(\lambda_j\right)\partial f_1^{-1}\left(\boldsymbol{\theta}_1,\lambda_j\right)+\int_0^{\pi} f_0\left(\lambda\right)\partial f_1^{-1}\left(\boldsymbol{\theta}_1,\lambda\right)d\lambda\\
&= O\left(\sum_{j=1}^{[n/2]}\int_{\lambda_{j-1}}^{\lambda_j}\left[f_0\left(\lambda_j\right)\partial f_1^{-1}\left(\boldsymbol{\theta}_1,\lambda_j\right)-f_0\left(\lambda\right)\partial f_1^{-1}\left(\boldsymbol{\theta}_1,\lambda\right)\right]d\lambda\right)\\
&= O\left(\sum_{j=1}^{[n/2]}\left|f_0\left(\lambda_{\tilde{j}}\right)\frac{\partial^2}{\partial\boldsymbol{\theta}\partial\lambda}f_1^{-1}\left(\boldsymbol{\theta}_1,\lambda_{\tilde{j}}\right)+\frac{\partial}{\partial\lambda}f_0\left(\lambda_{\tilde{j}}\right)\frac{\partial}{\partial\boldsymbol{\theta}}f_1^{-1}\left(\boldsymbol{\theta}_1,\lambda_{\tilde{j}}\right)\right|\int_{\lambda_{j-1}}^{\lambda_j}\left(\lambda_j-\lambda\right)d\lambda\right)\\
&= O\left(\sum_{j=1}^{[n/2]}\lambda_{\tilde{j}}^{-2d^*-1}\lambda_1^2\log\lambda_j\right)=\begin{cases} O\left(n^{2d^*-1}\log n\right), & 0<d^*<0.5\\ O\left(n^{-1}\log^4 n\right), & -0.5<d^*\leq 0\end{cases},
\end{aligned}$$

where $\lambda_{j-1}<\lambda_{\tilde{j}}<\lambda_j$.□

Corollary 2 *Under assumptions of Lemma 3 and $d^*>0.25$,*

$$\left\{\left(\partial^2 Q_n\left(\boldsymbol{\theta}^*\right)\right)^{-1}-\left(\partial^2 Q\left(\boldsymbol{\theta}_1\right)\right)^{-1}\right\}E\partial Q_n\left(\boldsymbol{\theta}_1\right)=o_p\left(\frac{\log n}{n^{1-2d^*}}\right),$$

where $\boldsymbol{\theta}^*$ lies between $\widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}_1$.

Proof. The corollary follows by Lemma 3 and 4.□

Lemma 5 *The normalised periodogram,*

$$\left(\frac{I\left(\lambda_1\right)}{f_0\left(\lambda_1\right)},\cdots,\frac{I\left(\lambda_s\right)}{f_0\left(\lambda_s\right)}\right)\xrightarrow{D}\left(Z_1,\ldots,Z_s\right)$$

for any fixed integer s, where $Z_j=A_j^2+B_j^2$, A_j and B_j are normal random variables with mean zero and

$$\operatorname{cov}\left(A_j,B_k\right)=\iint_{[0,1]^2}\left\{\phi_{A,j}\left(x\right)\phi_{B,k}\left(y\right)+\phi_{B,k}\left(x\right)\phi_{A,j}\left(y\right)\right\}\left|x-y\right|^{2d_0-1}dxdy$$

where

$$\phi_{A,j}\left(u\right)=\cos\left(2\pi uj\right)\quad\text{and}\quad\phi_{B,k}\left(u\right)=\sin\left(2\pi uk\right).$$

Furthermore

$$\text{cov}\left(Z_j, Z_k\right) = O\left(j^{-2d_0} k^{2d_0-2} \log^2 k\right) \text{ and } \text{var}\left(Z_j\right) = O\left(1 + j^{-1} \log j\right),$$

for $1 \le j < k \le s$.

Proof. See Deo (1997) for the expression $\text{cov}\left(A_j, B_k\right)$ and the first part of the lemma. Since $A_{n,j}$ and $B_{n,j}$ are normal random variables whose variances are bounded above for all n by (5), $E\left(A_{n,j}^p\right)$ and $E\left(B_{n,j}^p\right)$ are also bounded above for any $p > 0$. Hence, $A_{n,j}^p$ and $B_{n,j}^p$ are uniformly integrable for any $p > 0$. This fact in conjunction with the result that

$$A_{n,j} \xrightarrow{D} A_j \quad \text{and} \quad B_{n,j} \xrightarrow{D} B_j$$

implies (See the Corollary of Theorem 25.12 in Billingsley (1995) that

$$E\left(A_{n,j}^p\right) \to E\left(A_j^p\right) \text{ and } E\left(B_{n,j}^p\right) \to E\left(B_j^p\right).$$

Thus, $\text{cov}\left(Z_j, Z_k\right) = \lim_{n\to\infty} \text{cov}\left(\frac{I(\lambda_j)}{f_0(\lambda_j)}, \frac{I(\lambda_k)}{f_0(\lambda_k)}\right)$ and the covariance bound follows by applying Lemma 1.□

The rest of this appendix is dedicated to the limiting distribution of $\frac{\partial}{\partial \boldsymbol{\theta}} Q_n\left(\boldsymbol{\theta}_1\right) - E \frac{\partial}{\partial \boldsymbol{\theta}} Q_n\left(\boldsymbol{\theta}_1\right)$ for three cases, $d^* > 0.25$, $d^* < 0.25$ and $d^* = 0.25$. We will use the following notations,

$$w\left(\boldsymbol{\theta}_1, \lambda\right) = \frac{f_0\left(\lambda\right) \partial f_1\left(\boldsymbol{\theta}_1, \lambda\right)}{f_1^2\left(\boldsymbol{\theta}_1, \lambda\right)} = \left(w\left(d_1, \lambda\right), w'\left(\boldsymbol{\beta}_1, \lambda\right)\right)', \tag{13}$$

where

$$w\left(d_1, \lambda\right) = \frac{f_0\left(\lambda\right)}{f_1^2\left(\boldsymbol{\theta}_1, \lambda\right)} \frac{\partial f_1\left(\boldsymbol{\theta}_1, \lambda\right)}{\partial d} = (-2) \cdot \frac{f_0\left(\lambda\right)}{f_1\left(\boldsymbol{\theta}_1, \lambda\right)} \log\left(2 \sin\left(\lambda/2\right)\right) \tag{14}$$

and

$$w\left(\boldsymbol{\beta}_1, \lambda\right) = \frac{f_0\left(\lambda\right)}{f_1^2\left(\boldsymbol{\theta}_1, \lambda\right)} \frac{\partial f_1\left(\boldsymbol{\theta}_1, \lambda\right)}{\partial \boldsymbol{\beta}} = \frac{f_0\left(\lambda\right)}{f_1\left(\boldsymbol{\theta}_1, \lambda\right)} \frac{\partial g_1\left(\boldsymbol{\beta}_1, \lambda_j\right)}{g_1\left(\boldsymbol{\beta}_1, \lambda_j\right)}. \tag{15}$$

Lemma 6 *Under assumptions of Theorem 1* ($d^* > 0.25$),,

$$\frac{n^{1-2d^*}}{\log n} \left\{ \frac{\partial Q_n\left(\boldsymbol{\theta}_1\right)}{\partial d} - E \frac{\partial Q_n\left(\boldsymbol{\theta}_1\right)}{\partial d} \right\} \xrightarrow{D} \sum_{j=1}^{\infty} W_j,$$

where

$$W_j = \frac{2\pi g_0\left(0\right)}{\left(2\pi j\right)^{2d^*} g_1\left(\boldsymbol{\beta}_1, 0\right)} \left(Z_j - E Z_j\right)$$

and Z_j *is the same as defined in Lemma 5.*

Proof. Using the notation in (14), we have

$$\begin{aligned} \frac{n^{1-2d^*}}{\log n} \left\{ \frac{\partial Q_n\left(\boldsymbol{\theta}_1\right)}{\partial d} - E \frac{\partial Q_n\left(\boldsymbol{\theta}_1\right)}{\partial d} \right\} &= \frac{n^{1-2d^*}}{\log n} \frac{2\pi}{n} \sum_{j=1}^{[n/2]} \left(\frac{I\left(\lambda_j\right)}{f_0\left(\lambda_j\right)} - E \frac{I\left(\lambda_j\right)}{f_0\left(\lambda_j\right)} \right) w(d_1, \lambda_j) \\ &= \alpha_{s,n} + \beta_{s,n} \end{aligned}$$

where

$$\alpha_{s,n} = \frac{2\pi n^{-2d^*}}{\log n} \sum_{j=1}^{s} \left(\frac{I(\lambda_j)}{f_0(\lambda_j)} - E\frac{I(\lambda_j)}{f_0(\lambda_j)} \right) w(d_1, \lambda_j)$$

and

$$\beta_{s,n} = \frac{2\pi n^{-2d^*}}{\log n} \sum_{j=s}^{[n/2]} \left(\frac{I(\lambda_j)}{f_0(\lambda_j)} - E\frac{I(\lambda_j)}{f_0(\lambda_j)} \right) w(d_1, \lambda_j).$$

Since

$$\frac{n^{-2d^*}}{\log n} w(d_1, \lambda_j) \longrightarrow (2\pi j)^{-2d^*} \frac{g_0(0)}{g_1(\boldsymbol{\beta}_1, 0)},$$

we have for each s,

$$\alpha_{s,n} \xrightarrow{D} \sum_{j=1}^{s} W_j \text{ as } n \to \infty, \tag{16}$$

by Lemma 5.

The proof can be completed by verifying that

$$\sum_{j=1}^{s} W_j \xrightarrow{D} \sum_{j=1}^{\infty} W_j \quad \text{as} \quad s \to \infty \tag{17}$$

and

$$\lim_{s \longrightarrow \infty} \lim \sup_{n \to \infty} P\{|\beta_{s,n}| \geq \epsilon\} = 0. \tag{18}$$

See Theorem 4.2 of Billingsley (1968). We first show the Cauchy Convergence Criterion for (17),

$$\lim_{s\to\infty, m\to\infty} E\left(\sum_{j=s}^{m} W_j\right)^2 = o(1).$$

By Lemma 5,

$$\text{cov}(W_j, W_k) = O\left(j^{-2d^*-2d_0} k^{-2d^*+2d_0-2} \log^2 k\right), \text{ for } 1 \leq j < k \leq n/2$$

and

$$\text{var}(W_j) = O\left(j^{-4d^*}\right), \text{ for } 1 \leq j \leq n/2.$$

Hence

$$\begin{aligned} E\left(\sum_{j=s}^{m} W_j\right)^2 &= O\left(\sum_{j=s}^{m} j^{-4d^*} + \sum_{j<k=s}^{m} j^{-2d^*-2d_0} k^{-2d^*+2d_0-2} \log^2 k\right) \\ &= O\left(s^{-4d^*+1}\right) \to 0 \text{ as } s \to \infty, \end{aligned}$$

since $d^* > 0.25$. Using (5) and Chebyshev inequality, (18) follows by a similar computation as above. □

Lemma 7 *Under assumptions of Theorem 1 ($d^* > 0.25$),*

$$\frac{n^{1-2d^*}}{\log n}\left\{\frac{\partial Q_n(\boldsymbol{\theta}_1)}{\partial\boldsymbol{\beta}} - E\frac{\partial Q_n(\boldsymbol{\theta}_1)}{\partial\boldsymbol{\beta}}\right\} = o_p(1).$$

Proof. The LHS of the above equation is

$$\frac{n^{1-2d^*}}{\log n}\frac{2\pi}{n}\sum_{j=1}^{[n/2]}\left(\frac{I(\lambda_j)}{f_0(\lambda_j)} - E\frac{I(\lambda_j)}{f_0(\lambda_j)}\right)w(\boldsymbol{\beta}_1,\lambda_j),$$

where $w(\boldsymbol{\beta}_1,\lambda_j)$ is defined as (15). Now

$$w(\boldsymbol{\beta}_1,\lambda_j) = O\left(|\sin\lambda_j/2|^{-2d^*}\right) = O\left(\lambda_j^{-2d^*}\right),$$

since

$$\frac{\partial g_1(\boldsymbol{\beta}_1,\lambda_j)}{g_1(\boldsymbol{\beta}_1,\lambda_j)} = O(1),$$

by Assumption A.2 and A.3. Hence

$$E\left\|\frac{n^{1-2d^*}}{\log n}\frac{2\pi}{n}\sum_{j=1}^{[n/2]}\left(\frac{I(\lambda_j)}{f_0(\lambda_j)} - E\frac{I(\lambda_j)}{f_0(\lambda_j)}\right)w(\boldsymbol{\beta}_1,\lambda_j)\right\|^2$$

$$= \frac{n^{-4d^*}}{\log^2 n}\sum_{j,k=1}^{[n/2]}\left(\operatorname{cov}\frac{I(\lambda_j)}{f_0(\lambda_j)}\frac{I(\lambda_k)}{f_0(\lambda_k)}\right)w(\boldsymbol{\beta}_1,\lambda_j)\,w'(\boldsymbol{\beta}_1,\lambda_k)$$

$$= O\left(\frac{1}{\log^2 n}\sum_{j,k=1}^{[n/2]} j^{-2d^*-2d_0}k^{-2d^*+2d_0-2}\log^2 k\right) = O\left(\frac{1}{\log^2 n}\right).$$

□

Since X_t is Gaussian, it has MA(∞) expression, $X_t = \sum_{u=-\infty}^{\infty}\varphi_u\varepsilon_{t-u}$ where ε_t are *i.i.d.* $N(0,\sigma^2)$. We need the next lemma for the proof of Lemmas 9 and 11.

Lemma 8 *Let $I_\varepsilon(\lambda)$ be the periodogram of $\{\varepsilon_t\}_{t=1}^n$ and*

$$R(\lambda) = \frac{I(\lambda)}{f_0(\lambda)} - \frac{2\pi}{\sigma^2}I_\varepsilon(\lambda).$$

Then

$$E[R(\lambda_j)R(\lambda_k)] = O\left(j^{-1}k^{-1}\log k\log j + j^{-2d_0}k^{2d_0-2}\log^2 k\right)$$

and

$$E\left[R^2(\lambda_j)\right] = O\left(j^{-1}\log j\right)$$

for $\delta_n \le j < k \le n/2$, where

$$\delta_n \to \infty\,,\ \frac{\log\delta_n}{\delta_n} \to 0 \text{ and } \frac{\delta_n}{n} \to 0 \text{ as } n \to \infty.$$

Proof. This lemma is identical to Lemma 5 of Chen and Deo (2003) except δ_n was chosen to be $\log^2 n$ in that paper for convenience. □

Lemma 9 *Under assumptions of Theorem 2 ($d^* < 0.25$),*

$$n^{1/2}\left\{\frac{\partial Q_n(\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}} - E\frac{\partial Q_n(\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}}\right\} \xrightarrow{D} Y$$

where Y is a normal random vector with zero mean and the variance,

$$\mathrm{var}(Y) = 2\pi\int_0^\pi \left(\frac{f_0(\lambda)}{f_1(\boldsymbol{\theta}_1,\lambda)}\right)^2 \partial\log f_1(\boldsymbol{\theta}_1,\lambda)\,\partial\log f_1(\boldsymbol{\theta}_1,\lambda)'\,d\lambda$$

Proof. Using the notation in (13), we denote

$$Y_n = \left\{\frac{2\pi}{n}\sum_{j=1}^{[n/2]}\frac{2\pi}{\sigma^2}I_\varepsilon(\lambda_j)\,w(\boldsymbol{\theta}_1,\lambda_j)\right\}$$

and

$$\widetilde{Y}_n = \left\{\frac{2\pi}{n}\sum_{j=1}^{[n/2]}\left(\frac{2\pi}{\sigma^2}I_\varepsilon(\lambda_j) - E\frac{2\pi}{\sigma^2}I_\varepsilon(\lambda_j)\right)w(\boldsymbol{\theta}_1,\lambda_j)\right\}.$$

We will show that

$$n^{1/2}\left(\frac{\partial Q_n(\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}} - Y_n\right) \xrightarrow{p} 0, \tag{19}$$

and

$$n^{1/2}\widetilde{Y}_n \xrightarrow{D} Y. \tag{20}$$

Let $R(\lambda)$ be as in Lemma 8, we have

$$\begin{aligned}\left\|n^{1/2}\left(\frac{\partial Q_n(\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}} - Y_n\right)\right\|^2 &= n\left(\frac{2\pi}{n}\right)^2\sum_{j,k=1}^{[n/2]}R(\lambda_j)\,w'(\boldsymbol{\theta}_1,\lambda_j)R(\lambda_k)\,w(\boldsymbol{\theta}_1,\lambda_k)\\ &= n(H_1+H_2+H_3).\end{aligned}$$

where

$$H_1 = \left(\frac{2\pi}{n}\right)^2\sum_{j,k=1}^{\log\log n-1}R(\lambda_j)\,w'(\boldsymbol{\theta}_1,\lambda_j)R(\lambda_k)\,w(\boldsymbol{\theta}_1,\lambda_k),$$

$$H_2 = \left(\frac{2\pi}{n}\right)^2\sum_{j,k=\log\log n}^{[n/2]}R(\lambda_j)\,w'(\boldsymbol{\theta}_1,\lambda_j)R(\lambda_k)\,w(\boldsymbol{\theta}_1,\lambda_k)$$

and

$$H_3 = 2\cdot\left(\frac{2\pi}{n}\right)^2\sum_{j=1}^{\log\log n-1}R(\lambda_j)\,w'(\boldsymbol{\theta}_1,\lambda_j)\sum_{k=\log\log n}^{[n/2]}R(\lambda_k)\,w(\boldsymbol{\theta}_1,\lambda_k).$$

Note that

$$w(\boldsymbol{\theta}_1, \lambda_j) = O\left(\lambda_j^{-2d^*} \log \lambda_j\right).$$

By Lemma 1 and the fact that $\max_j E\left[I_\varepsilon^2(\lambda_j)\right] < \infty$, we have

$$nH_1 = O_p\left(n^{-1} \sum_{j,k=1}^{\log\log n - 1} w'(\boldsymbol{\theta}_1, \lambda_j) w(\boldsymbol{\theta}_1, \lambda_k)\right) = O_p\left(n^{-1+4d^*} \log^2 n \sum_{j,k=1}^{\log\log n - 1} j^{-2d^*} k^{-2d^*} \log j \log k\right)$$

$$= O_p\left(n^{-1+4d^*}\left(\log^2 n\right)(\log\log n)^{2-4d^*}\right) = o_p(1),$$

since $d^* < 0.25$. By Lemma 8,

$$\begin{aligned} nH_2 = \; & O_p\left(n^{4d^*-1} \log^2 n \left\{ \sum_{j=\log\log n}^{[n/2]} j^{-1-4d^*} \log j \right.\right. \\ & \left.\left. + \sum_{j=\log\log n}^{[n/2]} \sum_{k=j+1}^{[n/2]} \left(j^{-1-2d^*} k^{-1-2d^*} \log k \log j + j^{-2d_0-2d^*} k^{2d_0-2-2d^*} \log^2 k\right)\right\}\right) \\ = \; & O_p\left(n^{-1+4d^*}\left(\log^2 n\right)(\log\log n)^{-4d^*}\right) = o_p(1). \end{aligned}$$

Hence $nH_3 = o_p(1)$ by Cauchy-Schwartz inequality. We have proved (19).

We next show (20), or equivalently by the Cramer-Wold device, that

$$n^{1/2} c' \widetilde{Y}_n \xrightarrow{D} c'Y,$$

for all $c \in R^p$. Observe that

$$c'\widetilde{Y}_n = \frac{2\pi}{n} \sum_{j=1}^{[n/2]} \left(\frac{2\pi}{\sigma^2} I_\varepsilon(\lambda_j) - E\frac{2\pi}{\sigma^2} I_\varepsilon(\lambda_j)\right) c'w(\boldsymbol{\theta}_1, \lambda_j).$$

By Assumption A.2, $C_1 \lambda_j^{-2d^*} \log \lambda_j \le \|w(\boldsymbol{\theta}_1, \lambda_j)\| \le C_2 \lambda_j^{-2d^*} \log \lambda_j$ for some constants C_1, $C_2 > 0$. Hence,

$$\lim_{n\to\infty} \sup_{1 \le j \le n/2} \left(\sum_{j=1}^{n/2} \|c'w(\boldsymbol{\theta}_1, \lambda_j)\|^2\right)^{-1} \left[c'w(\boldsymbol{\theta}_1, \lambda_j)\right]^2 = 0. \tag{21}$$

Since ε_t is a Gaussian process, $I_\varepsilon(\lambda_j)$ are i.i.d exponential random variables, each with mean $\frac{\sigma^2}{2\pi}$ and variance $\left(\frac{\sigma^2}{2\pi}\right)^2$, see Brockwell and Davis (1996). We have

$$\frac{n}{2\pi}\left(\sum_{j=1}^{n/2} \left[c'w(\boldsymbol{\theta}_1, \lambda_j)\right]^2\right)^{-1/2} c'\widetilde{Y}_n \to N(0, 1)$$

by Corollary 5.3.4 of Fuller (1996). We have shown that

$$\frac{n}{2\pi}\left(\sum_{j=1}^{n/2} w(\boldsymbol{\theta}_1, \lambda_j) w'(\boldsymbol{\theta}_1, \lambda_j)\right)^{-1/2} \widetilde{Y}_n \to N(0, \mathbf{I}).$$

Equation (20) follows from the fact that

$$\frac{2\pi}{n}\sum_{j=1}^{[n/2]} w(\boldsymbol{\theta}_1,\lambda_j)w'(\boldsymbol{\theta}_1,\lambda_j) \to \int_0^{\pi} \frac{f_0^2(\lambda)}{f_1^2(\boldsymbol{\theta}_1,\lambda)}\partial \log f_1(\theta,\lambda)\,\partial \log f_1(\theta,\lambda)'\,d\lambda.$$

□

We need the following lemma for the proof of Lemmas 11 and 12.

Lemma 10 *Under the assumptions of Theorem 3 ($d^* = 0.25$), , there exist two constants $M_*, M^* > 0$ such that*

$$M_* < \frac{1}{n\log^3 n}\sum_{j=1}^{n/2} w^2(d,\lambda_j) < M^*,$$

where $w(d,\lambda)$ is defined as in (14).

Proof. Since $d^* = 0.25$,we'll show that

$$M_* < \frac{1}{n\log^3 n}\sum_{j=1}^{n/2}\left(\left|2\sin\frac{\lambda_j}{2}\right|^{-1/2}\log\left|2\sin\frac{\lambda_j}{2}\right|\frac{g_0(\lambda_j)}{g_1(\boldsymbol{\beta}_1,\lambda_j)}\right)^2 < M^*.$$

By assumption A. 2 and (1),

$$0 < m_* < \frac{g_0(\lambda_j)}{g_1(\boldsymbol{\beta}_1,\lambda_j)} < m^*,$$

for some positive constants m_* and m^*. Hence, it is sufficient to show that

$$\sum_{j=1}^{n/2}\left(\left|2\sin\frac{\lambda_j}{2}\right|^{-1/2}\log\left|2\sin\frac{\lambda_j}{2}\right|\right)^2 \sim Cn\log^3 n. \tag{22}$$

We will use the following formulae (see, for example, Gradshteyn and Ryzhik),

$$\log\sin x = \log x - \frac{x^2}{6} - \frac{x^4}{180} - \frac{x^4}{2835} - \cdots = \log x - \sum_{k=1}^{\infty}\frac{\varsigma(2k)\,x^2}{k\pi^2} = \log x - Cx^2,$$

where the zeta function $\varsigma(z) = \sum_{\ell=1}^{\infty}\frac{1}{\ell^z}$ and

$$\sin^{-1}\pi x = \frac{1}{\pi x} + \frac{2x}{\pi}\sum_{k=1}^{\infty}\frac{(-1)^k}{x^2 - k^2}.$$

Note that $\sin\frac{\lambda_j}{2} > 0$ for $j = 1,\ldots,n/2$. Applying the above formula by letting $x = j/n$, we have

$$\sin^{-1}\frac{\lambda_j}{2} = \frac{2}{\lambda_j} + \frac{2}{\pi}\frac{j}{n}\sum_{k=1}^{\infty}\frac{(-1)^k}{(j/n)^2 - k^2} = \frac{2}{\lambda_j} + O\left(\frac{j}{n}\right),$$

since

$$0<\left|\sum_{k=1}^{\infty}\frac{(-1)^k}{0-k^2}\right|\leq\left|\sum_{k=1}^{\infty}\frac{(-1)^k}{(j/n)^2-k^2}\right|\leq\left|\sum_{k=1}^{\infty}\frac{(-1)^k}{1/4-k^2}\right|<\sum_{k=1}^{\infty}\frac{1}{k^2}=\frac{\pi^2}{6}.$$

Hence the RHS of (22) multiplied by 1/4 is

$$\begin{aligned}
&\sum_{j=1}^{n/2}\sin^{-1}\frac{\lambda_j}{2}\log^2\left(2\sin\frac{\lambda_j}{2}\right)\\
=&\sum_{j=1}^{n/2}\frac{1}{\lambda_j}\left(\log^2 n-2\log n\log 2\pi j+\log^2 2\pi j\right)+O\left(\frac{j\log^2 n}{n}\right)\\
=&n\log^2 n\sum_{j=1}^{n/2}\frac{1}{2\pi j}-2n\log n\sum_{j=1}^{n/2}\frac{\log 2\pi j}{2\pi j}+n\sum_{j=1}^{n/2}\frac{\log^2 2\pi j}{2\pi j}+O\left(n\log^2 n\right)\\
=&n\log^2 n\left(\frac{\log(n/2)}{2\pi}\right)-2n\log n\left(\frac{1}{2}\frac{\log^2(2\pi n/2)}{2\pi}\right)+n\left(\frac{1}{3}\frac{\log^3(2\pi n/2)}{2\pi}\right)+O\left(n\log^2 n\right) \qquad (23)\\
=&\frac{1}{3}\frac{n\log^3 n}{2\pi}+O\left(n\log^2 n\right)
\end{aligned}$$

Equality (23) follows from

$$\sum_{j=1}^{n}\frac{1}{j}=\log n+O(1),\quad \sum_{j=1}^{n}\frac{\log j}{j}=\frac{1}{2}\log^2 n+O(1)\quad\text{and}\quad\sum_{j=1}^{n}\frac{\log^2 j}{j}=\frac{1}{3}\log^3 n+O(1).$$

We have

$$\frac{1}{n\log^3 n}\sum_{j=1}^{n/2}\sin^{-2}\frac{\lambda_j}{2}\log^2\left(2\sin\frac{\lambda_j}{2}\right)\to C.$$

□

Lemma 11 *Under assumptions of Theorem 3* ($d^*=0.25$),

$$\frac{n}{2\pi}\left(\sum_{j=1}^{n}w^2\left(d_1,\lambda_j\right)\right)^{-1/2}\left\{\frac{\partial Q_n\left(\boldsymbol{\theta}_1\right)}{\partial d}-E\frac{\partial Q_n\left(\boldsymbol{\theta}_1\right)}{\partial d}\right\}\xrightarrow{D}N(0,1)$$

where

$$w\left(d_1,\lambda\right)=\frac{f_0\left(\lambda_j\right)}{f_1^2\left(\boldsymbol{\theta}_1,\lambda_j\right)}\frac{\partial f_1\left(\boldsymbol{\theta}_1,\lambda_j\right)}{\partial d}.$$

Proof. The proof is similar to that of Lemma 9. Except now that $\int_0^\pi w^2\left(d_1,\lambda\right)d\lambda$ is not integrable since $d^*=.25$ and $w\left(d_1,\lambda\right)=O\left(\lambda^{-1/2}\log\lambda\right)$.

We will use similar notations in the proof of Lemma 9,

$$Y_n\left(d_1\right)=\left\{\frac{2\pi}{n}\sum_{j=1}^{[n/2]}\frac{2\pi}{\sigma^2}I_\varepsilon\left(\lambda_j\right)w(d_1,\lambda_j)\right\}$$

and

$$\widetilde{Y}_n(d_1) = \left\{ \frac{2\pi}{n} \sum_{j=1}^{[n/2]} \left(\frac{2\pi}{\sigma^2} I_\varepsilon(\lambda_j) - E\frac{2\pi}{\sigma^2} I_\varepsilon(\lambda_j) \right) w(d_1, \lambda_j) \right\}.$$

Following from lemma (10), we have

$$\left(\frac{n}{\log^3 n}\right)^{1/2} M_* < n \left(\sum_{j=1}^{n} w^2(d_1, \lambda_j) \right)^{-1/2} < \left(\frac{n}{\log^3 n}\right)^{1/2} M^*.$$

we will show that

$$\left(\frac{n}{\log^3 n}\right)^{1/2} \left(\frac{\partial Q_n(\boldsymbol{\theta}_1)}{\partial d} - Y_n(d_1) \right) \xrightarrow{p} 0, \tag{24}$$

and

$$\frac{n}{2\pi} \left(\sum_{j=1}^{n} w^2(d, \lambda_j) \right)^{-1/2} \widetilde{Y}_n(d_1) \xrightarrow{D} N(0,1). \tag{25}$$

Let

$$\begin{aligned}
\left| \left(\frac{n}{\log^3 n}\right)^{1/2} \left(\frac{\partial Q_n(\boldsymbol{\theta}_1)}{\partial d} - Y_n(d) \right) \right|^2 &= \left(\frac{n}{\log^3 n}\right) \left(\frac{2\pi}{n}\right)^2 \sum_{j,k=1}^{[n/2]} R(\lambda_j) w(d_1, \lambda_j) R(\lambda_k) w(d_1, \lambda_k) \\
&= \frac{n}{\log^3 n} \left(H_1(d_1) + H_2(d_1) + H_3(d_1) \right).
\end{aligned}$$

where H_1, H_2 and H_3 are the same as those in the proof of Lemma 9 with $w(\lambda)$ replacing by $w(d, \lambda)$. Note that

$$w(d_1, \lambda_j) = O\left(\lambda_j^{-1/2} \log \lambda_j \right)$$

By Lemma 1 and the fact that $\max_j E\left[I_\varepsilon^2(\lambda_j)\right] < \infty$, we have,

$$\begin{aligned}
\frac{n}{\log^3 n} H_1(d_1) &= O_p\left(\frac{n}{\log^3 n} \cdot \frac{1}{n^2} \sum_{j,k=1}^{\log\log n - 1} w(d_1, \lambda_j) w(d_1, \lambda_k) \right) \\
&= O_p\left(\frac{n}{\log^3 n} \cdot \frac{1}{n^2} \sum_{j,k=1}^{\log\log n - 1} \lambda_j^{-1/2} \lambda_k^{-1/2} \log \lambda_j \log \lambda_k \right) \\
&= O_p\left(\frac{n}{\log^3 n} \cdot \frac{n \log^2 n}{n^2} \cdot \log\log n \right) = o_p(1)
\end{aligned}$$

and

$$\begin{aligned}
&\frac{n}{\log^3 n} H_2(d_1) \\
=&O_p\left(\frac{n}{\log^3 n} \cdot \frac{n \log^2 n}{n^2} \left\{ \sum_{j=\log\log n}^{[n/2]} j^{-2} \log j + \sum_{j=\log\log n}^{[n/2]} \sum_{k=j+1}^{[n/2]} \left(j^{-3/2} k^{-3/2} \log k \log j + j^{-2d_0 - 1/2} k^{2d_0 - 5/2} \log^2 k \right) \right\} \right) \\
=&O_p\left(\frac{n}{\log^3 n} \cdot \frac{n \log^2 n}{n^2} \cdot \frac{1}{\log\log n} \right) = o_p(1)
\end{aligned}$$

by Lemma 8. Hence $\frac{n}{\log^3 n} H_3(d_1) = o_p(1)$ by Cauchy-Schwartz inequality. We have shown (24).

By Assumption A.2 and Lemma 20, hence,

$$\lim_{n\to\infty} \sup_{1\le j\le n/2} \left(\sum_{j=1} \|w(d_1,\lambda_j)\|^2 \right)^{-1} w^2\left(d,\lambda_j\right) = \lim_{n\to\infty} O\left(\log^{-1} n\right) = 0.$$

We have (25) by Corollary 5.3.4 of Fuller (1996).□

Lemma 12 *Under assumptions of Theorem (3),*

$$\left(\frac{n}{\log^3 n}\right)^{1/2} \left\{ \frac{\partial Q_n(\boldsymbol{\theta}_1)}{\partial \boldsymbol{\beta}} - E\frac{\partial Q_n(\boldsymbol{\theta}_1)}{\partial \boldsymbol{\beta}} \right\} = o_p(1).$$

Proof. The proof is similar to that of Lemma 7. Since $d^* = 0.25$,

$$w(\boldsymbol{\beta}_1, \lambda_j) = O\left(|\sin\lambda_j/2|^{-2d^*}\right) = O\left(\lambda_j^{-1/2}\right)$$

We have

$$\begin{aligned} &E\left\| \left(\frac{n}{\log^3 n}\right)^{1/2} \left\{ \frac{\partial Q_n(\boldsymbol{\theta}_1)}{\partial \boldsymbol{\beta}} - E\frac{\partial Q_n(\boldsymbol{\theta}_1)}{\partial \boldsymbol{\beta}} \right\} \right\|^2 \\ &= \frac{n}{\log^3 n}\left(\frac{2\pi}{n}\right)^2 \sum_{j,k=1}^{[n/2]} \left(\operatorname{cov}\frac{I(\lambda_j)}{f_0(\lambda_j)}\frac{I(\lambda_k)}{f_0(\lambda_k)}\right) w(\boldsymbol{\beta}_1,\lambda_j)\, w'(\boldsymbol{\beta}_1,\lambda_k) \\ &= O\left(\frac{1}{\log^3 n} \sum_{j,k=1}^{[n/2]} j^{-1/2-2d_0} k^{-1/2+2d_0-2}\log^2 k \right) = O\left(\frac{1}{\log^2 n}\right), \end{aligned}$$

since $0.25 \le d_0 < 0.5$. □

References

[1] Billingsley, P. (1995), Probability and Measure, 3'rd Ed., Wiley, New York.

[2] Brockwell, P. and Davis, R. (1996) Time Series: Theory and Methods, 2'st Ed., Springer, New York.

[3] Chen, W. and Deo, R.(2003) , "A Generalized Portmanteau Goodness-of-fit Test for Time Series Models," *Econometric Theory*, To appear.

[4] Dahlhaus, R. (1989), "Efficient parameter estimation for self-similar processes," *Annals of Statistics* **17**, 1749-1766.

[5] Deo, R. (1997), "Asymptotic theory for certain regression models with long memory errors," *J. Time Ser. Anal.* **18**, 385-393.

[6] Fox, R. and Taqqu, M. S. (1986), " Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series, " *Annals of Statistics*, **14**, 517-532.

[7] Fuller, W. (1996) , Introduction to Statistical Time Series, 2'nd Ed., Wiley, New York.

[8] Gallant, A. R. and Tauchen, G. (1996) "Which moments to match?" *Econometric Theory* **12** , 657–681.

[9] Isserlis, L. (1918),"On a formula for the product moment coefficient of any order of a normal frequency distribution in any number of variables," *Biometrika* **12**, 134-139.

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

[11] Taniguchi, M. (1979), " On estimation of parameters of Gaussian stationary processes, " *Journal of Applied Probability*, **50**, 575-591.

[12] Yajima, Y. (1993), "Asymptotic properties of estimates in incorrect ARMA models for long-memory time series," New directions in time series analysis, Part II, 375–382, Springer, New York.

[13] White, H. (1982), "Maximum likelihood estimation of misspecified model, " *Econometrica*, **16**, 1-26.

[14] Whittle, Y. (1953), "Estimation and information in stationary time series," *Arkiv för Matematik* , **2**, 423-434.