On spectral density-based goodness-of-fit tests for time series models

Linyuan Li* and New Name

Department of Mathematics and Statistics, University of New Hampshire, Durham, NH, 03824, USA

July 9, 2019

Abstract

A new goodness-of-fit test for time series models is proposed. The test statistic is based on the ratio between the periodogram and the parametric spectral density or its estimator under the null. The asymptotic distribution of the statistic proposed is derived and its power properties are discussed. Unlike most of current tests of goodness of fit, the asymptotic distribution of our test statistic allows the null hypothesis to be either a short- or long-range dependence model. Our test is in the frequency domain, is easy to compute, and does not require the calculation of residuals from the fitted model. The finite sample performance of the test is investigated through a simulation experiment and it shows that our test has power comparable to current ones.

Short title: Tests for time series models

2010 Mathematics Subject Classification: Primary: 62M10; Secondary: 62G10, 62M15

Keywords: Linear processes; periodogram; spectral density; Bartlett approximation; long memory model.

1 Introduction

This paper deals with more general goodness-of-test procedures, i.e. procedures which can be applied when no a priori information is available about what departures from the null should be anticipated; the proposed procedures are then consistent against all correlation structures different from that of the postulated model.

The spectral distribution function is only one of many different ways for describing the second order characteristics of a stationary sequence. Another way is by means of the autocorrelation sequence. In applications, overall tests of fit are frequently based on the examination of the correlation behavior of model residuals, such as the Portmanteau type tests which are based on the examination of the M, squared residual autocorrelations. A particular (and popular) example in this context is the test of fit of autoregressive moving average (ARMA) models proposed by Box Pierce (1970) and Ljung Box (1978). Note that for M such tests are not consistent

*Corresponding author: Linyuan Li, Phone: (603) 862-4592; Fax: (603) 862-4096; E-mail: Linyuan.Li@unh.edu

against all correlation alternatives, i.e. they have no power if the lack of is not captured by the behavior of that M residual autocorrelations considered.

In the present paper a different approach for testing the goodness-of of a time series model is proposed. The test is based on the property according to which, if the model is correct, then for each non-zero frequency the asymptotic expected value of the ratio between the sample spectral density (periodogram) and the spectral density of the ted model equals one. The idea is to estimate the expected value of this ratio non-parametrically and then to compare the estimate obtained with its expected value under the assumption that the parametric model is correct. To evaluate the discrepancies between these two quantities globally an integrated squared deviation measure is used. Moreover, the distribution of the test statistic considered is not affected even if the parameters specifying the spectral density under the null are not known a priori.

The paper is organized as follows. In the next section, we recall briefly the elements of Fourier expansion. The asymptotic distributions of the coefficients of cosine series expansion are derived for testing a simple and a composite hypothesis, which are asymptotically normal distributions. Based on those asymptotic results for those empirical coefficients, new test statistics are proposed. In Section 4, we provide some simulation studies and demonstrate that our proposed tests are very comparable to the current tests. We conclude in Section 5 with some remarks, while all technical proofs are provided in Appendix.

Throughout this paper, we use C to denote positive and finite constants whose value may change from line to line. Specific constants are denoted by $\tau_0, \tau_1, \tau_2, \gamma_0, \gamma_1, \gamma_2$ and so on.

2 The Test Statistics

2.1 Testing a simple hypothesis

In this paper, we assume that time series X_t, $t = 1, 2, \cdots, n$ is a realization of a linear process

$$X_t = \sum_{j=0}^{\infty} \psi_j \zeta_{t-j}, \quad \sum_{j=0}^{\infty} \psi_j^2 < \infty, \tag{2.1}$$

where $\{\zeta_t\}$ is a sequence of white noise. The class of above linear processes $\{X_t\}$ is very large. From Giraitis et al. (2012, p.38), every stationary process $\{X_t\}$ with zero mean, whose spectral density $f(\cdot)$ satisfies $\int_{[-\pi,\pi]} \log f(u)du > -\infty$, can be represented as a linear process as above. Denote a finite dimensional parametric class of spectral densities by $\mathcal{F} = \{f(\cdot, \boldsymbol{\theta}), \boldsymbol{\theta} \in \Theta\}$, where Θ is a parameter space for $\boldsymbol{\theta}$. To present the essential ideas of the test statistic, we consider first the basic case of testing a simple hypothesis

$$\begin{aligned} &H_0 : f(\mu) = f(\mu, \boldsymbol{\theta}), \quad \text{for all } \mu \in (0, \pi), \\ &H_1 : f(\mu) \neq f(\mu, \boldsymbol{\theta}), \quad \text{for some } \mu \in (0, \pi), \end{aligned} \tag{2.2}$$

where $f(\cdot)$ is the spectral density for observed process $\{X_t\}$ and $f(\cdot, \boldsymbol{\theta}) \in \mathcal{F}$ with that $\boldsymbol{\theta}$ is assumed as known. Let $h(\mu) = f(\mu)/f(\mu, \boldsymbol{\theta})$. Then, above hypothesis (2.2) becomes

$$H_0 : h(\mu) = 1, \quad \text{for all } \mu \in (0, \pi); \quad H_1 : h(\mu) \neq 1, \quad \text{for some } \mu \in (0, \pi). \tag{2.3}$$

For any periodic even function $h(\cdot) \in L^2([0,\, \pi])$, using basis $\{1/\sqrt{\pi}, \sqrt{2/\pi}\cos(k\mu), k = 1, 2, \cdots\}$, we have it's following Fourier expansion

$$h(\mu) = \beta_0 \frac{1}{\sqrt{\pi}} + \frac{\sqrt{2}}{\sqrt{\pi}} \sum_{k=1}^{\infty} \beta_k \cos(k\mu), \tag{2.4}$$

where $\beta_0 = \int_0^\pi h(\mu)/\sqrt{\pi} d\mu$ and $\beta_k = (\sqrt{2}/\sqrt{\pi}) \int_0^\pi h(\mu)\cos(k\mu) d\mu$.

Under null hypothesis H_0 in (2.3) (or in (2.2)), one has $\beta_0 = \sqrt{\pi}$ and $\beta_k = 0$ for all $k > 1$, i.e., all the true coefficients except β_0 vanish. Under alternative hypothesis H_1, β_0 may not be $\sqrt{\pi}$ or there exist some k, such that $\beta_k \neq 0$. Therefore, the original null hypothesis H_0 in (2.3) is equivalent to following hypothesis with above coefficients:

$$\begin{aligned} &H_0 : \beta_0 = \sqrt{\pi},\ \ \beta_k = 0 \quad \text{for all } k \geq 1, \\ &H_1 : \beta_0 \neq \sqrt{\pi},\ \text{ or } \beta_k \neq 0 \quad \text{for some } k \geq 1. \end{aligned} \tag{2.5}$$

Regarding the hypothesis testing problem (2.2), it is well known that the basic tool to make inference on the spectral density $f(\mu)$ of $\{X_t\}$ is its periodogram

$$I_X(\mu) = \frac{1}{2\pi n} \left| \sum_{t=1}^{n} X_t e^{-\mathbf{i}t\mu} \right|^2.$$

Also, denote $\mu_j = 2\pi j/n,$ the Fourier frequencies, $j = 1, 2, \cdots, \nu$, $\nu = \nu_n = [n/2] - 1$. We propose estimators for above coefficients as

$$\tilde{\beta}_0 = \frac{2\sqrt{\pi}}{n} \sum_{j=1}^{\nu} \frac{I_X(\mu_j)}{f(\mu_j, \boldsymbol{\theta})}, \tilde{\beta}_k = \frac{2\sqrt{2\pi}}{n} \sum_{j=1}^{\nu} \cos(k\mu_j) \frac{I_X(\mu_j)}{f(\mu_j, \boldsymbol{\theta})},\ k = 1, 2, \cdots, \nu/2. \tag{2.6}$$

In order to derive properties of above empirical coefficients and control the bound on the covariance of periodogram ordinates, we need to impose some conditions on those coefficients ψ_j and innovations ζ_t in (2.1). In specific, we require following assumptions.

A1: $\sum_j |j|^{1/2+\tau} |\psi_j| < \infty$ for an arbitrary small positive $0 < \tau < 1/2$.

Note above assumption **A1** is slightly stronger than $\sum_j |j|^{1/2} |\psi_j| < \infty$, which is usually assumed for short memory linear processes.

A2: Innovations $\{\zeta_t,\, t \in \mathbb{Z}\}$ are *i.i.d.* random variables with $E(\zeta_t) = 0$, $E(\zeta_t^2) = 1$ and $\mu_4 = E(\zeta_t^4) < \infty$.

A3: The time series X_t is fourth-order stationary, i.e., $E(X^4) < \infty$, and for any $h_1, h_2, h_3 \in \mathbb{Z}$ and $a_0, a_1, a_2, a_3 \in \{0, 1\}$, the product moment $E(X_t^{a_0} X_{t+h_1}^{a_1} X_{t+h_2}^{a_2} X_{t+h_3}^{a_3})$ does not depend on t.

In the above context, a typical weak dependence condition is to assume that

$$\sum_{h_1, h_2, h_3 \in \mathbb{Z}} |K(h_1, h_2, h_3)| < \infty \tag{2.7}$$

where $K(h_1, h_2, h_3) = \mathrm{Cum}\,(X_0, X_{h_1}, X_{h_2}, X_{h_3})$ for $h_1, h_2, h_3 \in \mathbb{Z}$ is the fourth order cumulant function of the process X. Under (2.7), we may define the fourth order cumulant spectral density function as

$$\kappa(\lambda_1, \lambda_2, \lambda_3) = (2\pi)^{-3} \sum_{h_1 \in \mathbb{Z}} \sum_{h_2 \in \mathbb{Z}} \sum_{h_3 \in \mathbb{Z}} K(h_1, h_2, h_3)\, e^{-\mathbf{i}(\lambda_1 h_1 + \lambda_2 h_2 + \lambda_3 h_3)}. \tag{2.8}$$

Now, we are ready to present following limit distributions for above empirical coefficients in (2.6).

Theorem 2.1 *Under null hypothesis H_0 in (2.2) and assumptions* **A1**, **A2**, **A3**, *we have,*

$$\begin{aligned}
&\sqrt{n}\,(\tilde{\beta}_0-\sqrt{\pi}) \longrightarrow_d N(0,\,2\pi q_0),\\
&\sqrt{n}\,\tilde{\beta}_k \longrightarrow_d N(0,\,2\pi), \quad \textit{for all}\ \ k=1,2,\cdots,\nu/2,\\
&Cov\left(\sqrt{n}\,\tilde{\beta}_{k_1},\ \sqrt{n}\,\tilde{\beta}_{k_2}\right) \longrightarrow 0, \quad \textit{for all}\ \ k_1\neq k_2, k_1,k_2=1,2,\cdots\nu/2,\\
&Cov\left(\sqrt{n}\,(\tilde{\beta}_0-\sqrt{\pi}),\ \sqrt{n}\,\tilde{\beta}_k\right) \longrightarrow 0, \quad \textit{for all}\ \ k=1,2,\cdots,\nu/2,
\end{aligned}$$

where $q_0=1+\eta_4/2$, $\eta_4=Cum_4(\zeta_0)$ and "$\longrightarrow_d$" stands for convergence in distribution.

Remark 2.1 If innovations ζ_t follow normal distribution, one has $\eta_4=\mathrm{Cum}_4(\zeta_0)=E(\zeta_0^4)-3\sigma_{\zeta_0}^4=\mu_4-3=0$. Thus one has $q_0=1$. For non-Gaussian linear processes, there is a literature on how to estimate $\mathrm{Cum}_4(\zeta_0)$, see, e.g., Fragkeskou and Paparoditis (2016).

Since above empirical coefficients $\tilde{\beta}_0$ and $\tilde{\beta}_k$ are consistent estimators of the β_0 and β_k's, it is intuitive to propose the following test statistic $X_m^2(\boldsymbol{\theta})$ for the hypothesis testing problem (2.3):

$$X_m^2(\boldsymbol{\theta}) = \frac{n}{2\pi q_0}\left(\tilde{\beta}_0-\sqrt{\pi}\right)^2+\frac{n}{2\pi}\sum_{k=1}^{m}\tilde{\beta}_k^2. \tag{2.9}$$

From Theorem 2.1 and applying continuous mapping theorem, we obtain following result.

Theorem 2.2 *Under null hypothesis H_0 in (2.2) and assumptions* **A1**, **A2**, **A3**, *we have, for any fixed m such that $0\le m<\nu/2$,*

$$X_m^2(\boldsymbol{\theta}) \longrightarrow_d \chi^2(m+1), \quad \textit{as}\ n\to\infty.$$

For long memory process case, we need additional assumption on innovation that ζ_j's are normal random variables. Then we have similar results on those coefficients.

Theorem 2.3 *Under null hypothesis H_0 in (2.2) and assumptions* **A1**, **A2**, **A3**, *we have,*

$$\begin{aligned}
&\sqrt{n}\,(\tilde{\beta}_0-\sqrt{\pi}) \longrightarrow_d N(0,\,2\pi q_0),\\
&\sqrt{n}\,\tilde{\beta}_k \longrightarrow_d N(0,\,2\pi), \quad \textit{for all}\ \ k=1,2,\cdots,\nu/2,\\
&Cov\left(\sqrt{n}\,\tilde{\beta}_{k_1},\ \sqrt{n}\,\tilde{\beta}_{k_2}\right) \longrightarrow 0, \quad \textit{for all}\ \ k_1\neq k_2, k_1,k_2=1,2,\cdots\nu/2,\\
&Cov\left(\sqrt{n}\,(\tilde{\beta}_0-\sqrt{\pi}),\ \sqrt{n}\,\tilde{\beta}_k\right) \longrightarrow 0, \quad \textit{for all}\ \ k=1,2,\cdots,\nu/2,
\end{aligned}$$

where $q_0=1+\eta_4/2$, $\eta_4=Cum_4(\zeta_0)=0$ for normal innovation.

2.2 Testing a composite hypothesis

In applications, it is perhaps more important to test a composite hypothesis, i.e., $f(\cdot,\boldsymbol{\theta}) \in \mathcal{F}$ with $\boldsymbol{\theta}$ is unknown. In this case, the null hypothesis is given by $H_0 : f(\mu) = f(\mu,\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta$. Let $\hat{\boldsymbol{\theta}}$ be the estimator of $\boldsymbol{\theta}$, we propose estimators for above coefficients as

$$\hat{\beta}_0 = \frac{2\sqrt{\pi}}{n}\sum_{j=1}^{\nu}\frac{I_X(\mu_j)}{f(\mu_j,\hat{\boldsymbol{\theta}})}, \quad \hat{\beta}_k = \frac{2\sqrt{2\pi}}{n}\sum_{j=1}^{\nu}\cos(k\mu_j)\frac{I_X(\mu_j)}{f(\mu_j,\hat{\boldsymbol{\theta}})},\ k = 1,2,\cdots,\nu/2. \tag{2.10}$$

Then we have

Theorem 2.4 *Under the null hypothesis H_0, we have,*

$$\begin{aligned}
&\sqrt{n}\,(\hat{\beta}_0 - \sqrt{\pi}) \longrightarrow_d \mathcal{N}(0,\,2\pi q_0),\\
&\sqrt{n}\,\hat{\beta}_k \longrightarrow_d \mathcal{N}(0,\,2\pi), \quad \textit{for all} \ \ k = 1,2,\cdots,\nu/2,\\
&Cov\Big(\sqrt{n}\,\hat{\beta}_{k_1},\ \sqrt{n}\,\hat{\beta}_{k_2}\Big) \longrightarrow 0, \quad \textit{for all} \ \ k_1 \neq k_2, k_1, k_2 = 1,2,\cdots\nu/2,\\
&Cov\Big(\sqrt{n}\,(\hat{\beta}_0 - \sqrt{\pi}),\ \sqrt{n}\,\hat{\beta}_k\Big) \longrightarrow 0, \quad \textit{for all} \ \ k = 1,2,\cdots,\nu/2.
\end{aligned}$$

Let test statistic $X_m^2(\hat{\boldsymbol{\theta}})$ be defined as

$$X_m^2(\hat{\boldsymbol{\theta}}) = \frac{n}{2\pi\hat{q}_0}\big(\hat{\beta}_0 - \sqrt{\pi}\big)^2 + \frac{n}{2\pi}\sum_{k=1}^{m}\hat{\beta}_k^2, \tag{2.11}$$

where $\hat{q}_0$ is a consistent estimate of q_0. Then, we have

Theorem 2.5 *Under the null hypothesis H_0, we have, for any fixed m such that $0 \le m < \nu/2$,*

$$X_m^2(\hat{\boldsymbol{\theta}}) \longrightarrow_d \chi^2(m+1), \quad \textit{as } n \to \infty.$$

For long memory process case, we have similar result.

Theorem 2.6 *Under the null hypothesis H_0, we have,*

$$\begin{aligned}
&\sqrt{n}\,(\hat{\beta}_0 - \sqrt{\pi}) \longrightarrow_d \mathcal{N}(0,\,2\pi q_0),\\
&\sqrt{n}\,\hat{\beta}_k \longrightarrow_d \mathcal{N}(0,\,2\pi), \quad \textit{for all} \ \ k = 1,2,\cdots,\nu/2,\\
&Cov\Big(\sqrt{n}\,\hat{\beta}_{k_1},\ \sqrt{n}\,\hat{\beta}_{k_2}\Big) \longrightarrow 0, \quad \textit{for all} \ \ k_1 \neq k_2, k_1, k_2 = 1,2,\cdots\nu/2,\\
&Cov\Big(\sqrt{n}\,(\hat{\beta}_0 - \sqrt{\pi}),\ \sqrt{n}\,\hat{\beta}_k\Big) \longrightarrow 0, \quad \textit{for all} \ \ k = 1,2,\cdots,\nu/2.
\end{aligned}$$

Propose adaptive test statistics

$$\hat{m} = \underset{m}{\operatorname{argmax}}\Big\{\frac{n}{2\pi\hat{q}_0}\big(\hat{\beta}_0 - \sqrt{\pi}\big)^2 + \frac{n}{2\pi}\sum_{k=1}^{m}\hat{\beta}_k^2 - 2(m+1)\Big\}, \tag{2.12}$$

where m is determined from BIC or AIC. The adaptive test statistics is defined as

$$X_{\hat{m}}^2(\hat{\boldsymbol{\theta}}) = \frac{n}{2\pi\hat{q}_0}\big(\hat{\beta}_0 - \sqrt{\pi}\big)^2 + \frac{n}{2\pi}\sum_{k=1}^{\hat{m}}\hat{\beta}_k^2, \tag{2.13}$$

3 Implementation and Simulation Studies

4 Conclusion

5 Appendix

We need some preparatory results.

Lemma 5.1 *Let $\mu_j = 2\pi j/n, j = 1, 2, \cdots, \nu$ and $\nu = [n/2] - 1$. Then, we have, as $n \to \infty$,*

$$\frac{1}{\nu}\sum_{j=1}^{\nu}\cos(2k\mu_j) \to 0, \text{ uniformly for all integers } 1 \le k \le \nu/2,$$

$$\frac{1}{\sqrt{\nu}}\sum_{j=1}^{\nu}\cos(k\mu_j) \to 0, \text{ uniformly for all integers } 1 \le k \le \nu.$$

Proof: Since proofs to the above two results are similar, we only provide a proof to the second result here. Use a result that $\sum_{j=1}^{\nu}\cos(\lambda j) = \frac{\sin[(\nu+\frac{1}{2})\lambda]}{2\sin(\lambda/2)} - \frac{1}{2}$ with $\lambda = 2k\pi/n$, we have

$$\frac{1}{\sqrt{\nu}}\sum_{j=1}^{\nu}\cos(k\mu_j) = \frac{1}{2\sqrt{\nu}}\frac{\sin(2k\pi\nu/n)\cos(k\pi/n)}{\sin(k\pi/n)} + \frac{1}{2\sqrt{\nu}}\cos(2k\pi\nu/n) - \frac{1}{2\sqrt{\nu}}. \qquad (5.1)$$

Since $\nu = [n/2] - 1 \to \infty$ as $n \to \infty$, above 2nd and 3rd terms on the RHS of (5.1) go to zero. Thus, in order to prove the Lemma, we only need to show that the first term on the RHS of (5.1) goes to zero uniformly for all integers $1 \le k \le \nu$. Furthermore, it suffices to show $\nu^{-1/2}\sin(2\pi k\nu/n)/\sin(\pi k/n) \to 0$ to prove the Lemma. From $2/\pi \le \sin(x)/x \le 1$ for all $x \in (0, \pi/2)$, it is equivalent to show that $n^{1/2}|\sin(2\pi k\nu/n)|/k \to 0$, ignoring a multiple of a constant. Applying the mean value theorem, we could write $n^{1/2}|\sin(2k\pi\nu/n)|/k = n^{1/2}|\sin(2k\pi\nu/n) - \sin(k\pi)|/k \le Cn^{1/2}|2\nu/n - 1||\cos\{k\pi[1 + \eta(2\nu/n - 1)]\}|$, where $0 < \eta < 1$. From $\nu = [n/2] - 1$ we have $n^{1/2}|2\nu/n - 1| \to 0$. Thus we have $n^{1/2}|\sin(2\pi k\nu/n)|/k \to 0$, which proves the Lemma.

Next Lemma quantifies the covariance for periodogram ordinates for short memory processes.

Lemma 5.2 *Let $\mu_j = 2\pi j/n, \mu_k = 2\pi k/n, j, k = 1, 2, \cdots, \nu$ and $\nu = [n/2] - 1$. Under assumptions* **A1**, **A2**, **A3**, *we have,*

$$Cov\big(I_X(\mu_j),\, I_X(\mu_k)\big) = n^{-1}\eta_4 f(\mu_j, \boldsymbol{\theta})f(\mu_k, \boldsymbol{\theta}) + o(n^{-1}),$$

for all μ_j and μ_k such that $\mu_j \neq \mu_k$. Above $o(\cdot)$ term is uniform in μ_j and μ_k.

Proof: The proof is adapted from Krogstad (1982). For completeness of exposition, we provide an outline below. From (4.2) in Krogstad (1982, p.198), we have $\mathrm{Cov}\big(I_X(\mu_j),\, I_X(\mu_k)\big) = Q_n(\mu_j, \mu_k) + C_n(\mu_j, \mu_k)$, where the 1st term $Q_n(\mu_j, \mu_k) = n^{-2}\sum\sum\sum_{j,k,l,m=0}^{n-1} K(k-j, l-j, m-j)\, e^{-\mathbf{i}(j-k)\mu_j - \mathbf{i}(l-m)\mu_k}$ and $C_n(\mu_j, \mu_k) = \big\{r(j-l)r(k-m) + r(j-m)r(k-l)\big\}e^{-\mathbf{i}(j-k)\mu_j - \mathbf{i}(l-m)\mu_k}$.

Above three variables function $K(\cdot,\cdot,\cdot)$ is fourth order cumulant function for the process. From (4.7) in Krogstad (1982, p.199), we further have $Q_n(\mu_j,\mu_k) = n^{-1}\kappa(\mu_j,\mu_k,-\mu_k)(1+o(1))$, where $\kappa(\cdot,\cdot,\cdot)$ is fourth order cumulant spectral density function defined as

$$\kappa(\lambda_1,\lambda_2,\lambda_3) = (2\pi)^{-3}\sum_{h_1}\sum_{h_2}\sum_{h_3} K(h_1,h_2,h_3)\, e^{-\mathbf{i}(\lambda_1 h_1+\lambda_2 h_2+\lambda_3 h_3)},$$

and $o(\cdot)$ term is uniform in μ_j and μ_k. From Rosenblatt (1985, p.130), $Q_n(\mu_j,\mu_k)$ cab be further written as $Q_n(\mu_j,\mu_k) = n^{-1}\eta_4 f(\mu_j,\boldsymbol{\theta})f(\mu_k,\boldsymbol{\theta})(1+o(1))$, where $\eta_4 = E(\zeta_0^4)-3$ is the fourth-order cumulant of ζ_0. Also see Fragkeskou and Paparoditis (2016, p.241) about this statement.

As to 2nd term, we see that $C_n(\mu_j,\mu_k)$ is exactly same as the sum of term in (10.3.15) and term in (10.3.16) in Brockwell and Davis (1991, p.349). Note that $0 < \mu_j+\mu_k < 2\pi$ for our μ_j and μ_k. They showed that $|C_n(\mu_j,\mu_k)| = O(n^{-1})$ under a weaker assumption that $\sum_j |j|^{1/2}|\psi_j| < \infty$. Nevertheless, trace their proof, it is not hard to derive that $|C_n(\mu_j,\mu_k)| = o(n^{-1})$ under our slightly stronger assumption that $\sum_j |j|^{1/2+\tau}|\psi_j| < \infty$. In specific, the right side of (10.3.18) in Brockwell and Davis (1991, p.349) can be bounded by

$$\frac{1}{n}\sum_{|s|<n}|s||r(s)| \le \frac{1}{n^{1/2+\tau}}\sum_{|s|<n}\sum_{j=-\infty}^{\infty}|s|^{1/2+\tau}|\psi_j\psi_{j+s}| \le \frac{2}{n^{1/2+\tau}}\Big(\sum_{s=-\infty}^{\infty}|s|^{1/2+\tau}|\psi_s|\Big)\Big(\sum_{j=-\infty}^{\infty}|\psi_j|\Big).$$

Now multiply two factors of (10.3.15) in Brockwell and Davis (1991, p.348) together, we obtain it is bounded by $O(n^{-1-2\tau}) = o(n^{-1})$. Thus we prove that the term in (10.3.15) is $o(n^{-1})$, which proves $|C_n(\mu_j,\mu_k)| = o(n^{-1})$.

The following result shows that the covariance of X^2 and Y^2 can be bounded by the covariance of X and Y themselves, when $(X,\,Y)$ follows a bivariate normal distribution.

Lemma 5.3 *Let* $(X,\,Y)\sim N(0,0,\sigma_1^2,\sigma_2^2,\sigma_{12})$ *with* $Var(X)=\sigma_1^2$, $Var(Y)=\sigma_2^2$, $Cov(X,Y)=\sigma_{12}$, *then*

$$Cov(X^2,\,Y^2) = 2\sigma_{12}^2.$$

Proof: Let $U = X/\sigma_1$ and $V = Y/\sigma_2$, then $(U,\,V)\sim N(0,0,1,1,\rho)$, where $\rho = \sigma_{12}/(\sigma_1\sigma_2)$. From this relation, in order to prove the Lemma, it is equivalent to show $\mathrm{Cov}\,(U^2,\,V^2) = 2\rho^2$. Let $W_1 = U-\rho V$, $W_2 = \sqrt{1-\rho^2}\,V$, then it is not hard to show that W_1, W_2 are independent, and follow the same normal distribution $N(0,\,1-\rho^2)$. Write U and V in terms of W_1 and W_2, we have $U = W_1+\rho(1-\rho^2)^{-1/2}W_2$ and $V = (1-\rho^2)^{-1/2}W_2$. Now

$$\begin{aligned}\mathrm{Cov}\,(U^2,\,V^2) &= \mathrm{Cov}\,(W_1^2+2\rho(1-\rho^2)^{-1/2}W_1W_2+\rho^2(1-\rho^2)^{-1}W_2^2,\,(1-\rho^2)^{-1}W_2^2)\\ &= 0+0+\rho^2(1-\rho^2)^{-2}\mathrm{Cov}\,(W_2^2,\,W_2^2) = 2\rho^2,\end{aligned}$$

above first equality follows from the independence of W_1 and W_2, and last equality follows from $W_2\sim N(0,\,1-\rho^2)$.

The following Lemma provides a bound on covariance for periodograms for long memory Gaussian processes.

Lemma 5.4 *Let $\mu_j = 2\pi j/n, \mu_k = 2\pi k/n, j, k = 1, 2, \cdots, \nu$ and $\nu = [n/2] - 1$. Then, for long memory case with Gaussian innovation, we have,*

$$Cov\Big(\frac{I_X(\mu_j)}{f(\mu_j, \boldsymbol{\theta})}, \frac{I_X(\mu_k)}{f(\mu_k, \boldsymbol{\theta})}\Big) = O\Big(\Big(\frac{k}{j}\Big)^{2d}\frac{\log^2 k}{k^2}\Big),$$

uniformly for $1 \leq j \leq k \leq \nu$.

Proof: Let

$$A_j = \frac{1}{(2\pi n)^{1/2}}\sum_{t=1}^{n} X_t \cos(\mu_j t), \quad B_j = \frac{1}{(2\pi n)^{1/2}}\sum_{t=1}^{n} X_t \sin(\mu_j t).$$

Then

$$\frac{I_X(\mu_j)}{f(\mu_j, \boldsymbol{\theta})} = \Big(\frac{A_j}{f^{1/2}(\mu_j, \boldsymbol{\theta})}\Big)^2 + \Big(\frac{B_j}{f^{1/2}(\mu_j, \boldsymbol{\theta})}\Big)^2.$$

Similarly, we define A_k, B_k and $I_X(\mu_k)/f(\mu_k, \boldsymbol{\theta})]$ as A_j, B_j, etc. Also define

$$\mathbf{V} = \Big(\frac{A_j}{f^{1/2}(\mu_j, \boldsymbol{\theta})}, \frac{B_j}{f^{1/2}(\mu_j, \boldsymbol{\theta})}, \frac{A_k}{f^{1/2}(\mu_k, \boldsymbol{\theta})}, \frac{B_k}{f^{1/2}(\mu_k, \boldsymbol{\theta})}\Big)^T = \big(V_1,\, V_2,\, V_3,\, V_4\big)^T,$$

then we have

$$\begin{aligned}\mathrm{Cov}\Big(\frac{I_X(\mu_j)}{f(\mu_j, \boldsymbol{\theta})}, \frac{I_X(\mu_k)}{f(\mu_k, \boldsymbol{\theta})}\Big) &= \mathrm{Cov}\,(V_1^2 + V_2^2,\, V_3^2 + V_4^2) \\ &= \mathrm{Cov}\,(V_1^2,\, V_3^2) + \mathrm{Cov}\,(V_1^2,\, V_4^2) + \mathrm{Cov}\,(V_2^2,\, V_3^2) + \mathrm{Cov}\,(V_2^2,\, V_4^2).\end{aligned} \tag{5.2}$$

Above $(V_{l_1},\, V_{l_2})$, $l_1 = 1, 2$; $l_2 = 3, 4$, follow bivariate normal distributions, denoted with $N(0, 0, \sigma_{l_1}^2, \sigma_{l_2}^2, \sigma_{l_1 l_2})$, where $\mathrm{Cov}(V_{l_1},\, V_{l_2}) = \sigma_{l_1 l_2}$. From Lemma 9 of Moulines and Soulier (1999, p.1436) or (A.37) in Chen and Deo (2004, p.412), we have $|\mathrm{Cov}(V_{l_1},\, V_{l_2})| = O(j^{-d}k^{d-1}\log k)$ uniformly for $1 \leq j \leq k \leq \nu$. Now apply Lemma 5.3, we have $|\mathrm{Cov}\,(V_{l_1}^2,\, V_{l_2}^2)| = O\big((k/j)^{2d}\log^2 k/k^2\big)$. Thus, we prove the lemma.

Proof to Theorem 2.1: In view of (2.6), we write

$$\sqrt{n}\,\tilde{\beta}_0 = \frac{\sqrt{2\nu}}{\sqrt{n}}\sum_{j=1}^{\nu}\frac{\sqrt{2\pi}}{\sqrt{\nu}}\frac{I_X(\mu_j)}{f(\mu_j, \boldsymbol{\theta})}.$$

From $\nu = [n/2] - 1$, we have $\sqrt{2\nu}/\sqrt{n} \to 1$. From Slutsky theorem, we only need to verify the limit distribution for the summation term. In order to do that, we apply a result from Giraitis and Koul (2013) on the limit distribution for $S_{n,X} = \sum_{j=1}^{\nu} b_{n,j}\frac{I_X(\mu_j)}{f(\mu_j, \boldsymbol{\theta})}$. Thus we need to verify the corresponding assumptions. In our case, we have

$$b_{n,j} = \frac{\sqrt{2\pi}}{\sqrt{\nu}},\ b_n := \max_j\{|b_{n,j}|\} = \frac{\sqrt{2\pi}}{\sqrt{\nu}},\ B_n = \Big(\sum_{j=1}^{\nu} b_{n,j}^2\Big)^{1/2} = (2\pi)^{1/2},$$

$$q_n^2 = B_n^2 + \mathrm{Cum}_4(\zeta_0)\frac{1}{n}(\sum_{j=1}^{\nu} b_{n,j})^2 = 2\pi\Big(1 + \mathrm{Cum}_4(\zeta_0)\frac{\nu}{n}\Big).$$

From $b_n/B_n = \nu^{-1/2} \to 0$, the assumptions for the Theorem 2.1 in Giraitis and Koul (2013) are satisfied. Combining $q_n^2 \to 2\pi(1+\mathrm{Cum}_4(\zeta_0)/2) = 2\pi q_0$ from $\nu/n \to 1/2$ and $\sum_{j=1}^{\nu} b_{n,j} = \sqrt{2\nu\pi}$, applying their Theorem 2.1, we have

$$q_n^{-1}\big(\sqrt{n}\tilde{\beta}_0 - \sqrt{2\nu\pi}\big) \longrightarrow_d N(0,\, 1).$$

Now from $\sqrt{n\pi} - \sqrt{2\nu\pi} \to 0$, $q_n^2 \to 2\pi q_0$ and Slutsky theorem, we prove $\sqrt{n}(\tilde{\beta}_0 - \sqrt{\pi}) \longrightarrow_d N(0,\, 2\pi q_0)$, which is the first result in Theorem 2.1.

As to the second result, we write

$$\sqrt{n}\,\tilde{\beta}_k = \frac{\sqrt{2\nu}}{\sqrt{n}} \sum_{j=1}^{\nu} \frac{2\sqrt{\pi}}{\sqrt{\nu}} \cos(k\mu_j) \frac{I_X(\mu_j)}{f(\mu_j, \boldsymbol{\theta})}.$$

Again, we only need to verify the limit distribution for the summation term. In this case, we have

$$b_{n,j} = \frac{2\sqrt{\pi}}{\sqrt{\nu}} \cos(k\mu_j), \quad b_n := \max_j\{|b_{n,j}|\} \le \frac{2\sqrt{\pi}}{\sqrt{\nu}},$$

and $B_n^2 = \sum_{j=1}^{\nu} b_{n,j}^2 = 2\pi + 2\pi\nu^{-1}\sum_{j=1}^{\nu}\cos(2k\mu_j)$. From above Lemma 5.1, we have $B_n^2 \to 2\pi$. Thus, we have $b_n/B_n \to 0$. Hence, the assumptions for the Theorem 2.1 in Giraitis and Koul (2013) are satisfied. Therefore, we have

$$q_n^{-1}\big(\sqrt{n}\tilde{\beta}_k - \sum_{j=1}^{\nu} b_{n,j}\big) \longrightarrow_d N(0,\, 1),$$

where $q_n^2 = B_n^2 + \mathrm{Cum}_4(\zeta_0)\frac{1}{n}(\sum_{j=1}^{\nu} b_{n,j})^2$. Now from above Lemma 5.1 that $\sum_{j=1}^{\nu} b_{n,j} \to 0$, we have $q_n^2 \to 2\pi$ and $\sqrt{n}\tilde{\beta}_k \longrightarrow_d N(0,\, 2\pi)$, which is the second result in Theorem 2.1.

Next, we show that those empirical coefficients are uncorrelated asymptotically.

$$\begin{aligned}
\mathrm{Cov}\left(n^{1/2}\tilde{\beta}_{k_1},\; n^{1/2}\tilde{\beta}_{k_2}\right) &= \frac{8\pi}{n}\sum_{j_1=1}^{\nu}\sum_{j_2=1}^{\nu}\cos(k_1\mu_{j_1})\cos(k_2\mu_{j_2})\mathrm{Cov}\left(\frac{I_X(\mu_{j_1})}{f(\mu_{j_1},\boldsymbol{\theta})},\; \frac{I_X(\mu_{j_2})}{f(\mu_{j_2},\boldsymbol{\theta})}\right) \\
&= \frac{8\pi}{n}\sum_{j=1}^{\nu}\cos(k_1\mu_j)\cos(k_2\mu_j)\mathrm{Var}\left(\frac{I_X(\mu_j)}{f(\mu_j,\boldsymbol{\theta})}\right) \\
&\quad + \frac{8\pi}{n}\sum\sum_{j_1\ne j_2}\cos(k_1\mu_{j_1})\cos(k_2\mu_{j_2})\mathrm{Cov}\left(\frac{I_X(\mu_{j_1})}{f(\mu_{j_1},\boldsymbol{\theta})},\; \frac{I_X(\mu_{j_2})}{f(\mu_{j_2},\boldsymbol{\theta})}\right) \\
&=: I_{1n} + I_{2n},
\end{aligned}$$

Using $\mathrm{Var}(I_X(\mu_j)) = (2\pi)^2 f^2(\mu_j, \boldsymbol{\theta}) + O(n^{-1/2})$, one has

$$I_{1n} =: \frac{32\pi^3}{n}\sum_{j=1}^{\nu}\cos(k_1\mu_j)\cos(k_2\mu_j) + R_{1n}, \tag{5.3}$$

where $|R_{1n}| \le Cn^{-1}\nu n^{-1/2} \le Cn^{-1/2} \to 0$. Write $\cos(k_1\mu_j)\cos(k_2\mu_j) = \{\cos[(k_1+k_2)\mu_j] - \cos[(k_1-k_2)\mu_j]\}/2$, together with $k_1 \ne k_2$ and $1 \le k_1 + k_2 \le \nu$, one could derive the first term of

RHS of (5.3) goes to 0 just as the proof to Lemma 5.1. Thus, we prove that $I_{1n} \to 0$ as $n \to \infty$. As to I_{2n}, applying Lemma 5.2, we have

$$I_{2n} =: \frac{8\pi\eta_4}{n^2} \sum\sum_{j_1 \neq j_2} \cos(k_1\mu_j)\cos(k_2\mu_j) + R_{2n}, \tag{5.4}$$

where $|R_{2n}| = o(n^{-1}\nu^2 n^{-1})$, which implies that $|R_{2n}| \to 0$. As to the first term of RHS of (5.4), write it as

$$\frac{8\pi\eta_4}{n^2} \sum_{j_1=1}^{\nu} \sum_{j_2=1}^{\nu} \cos(k_1\mu_{j_1})\cos(k_2\mu_{j_2}) - \frac{8\pi\eta_4}{n^2} \sum_{j=1}^{\nu} \cos(k_1\mu_j)\cos(k_2\mu_j).$$

Now, apply Lemmas 5.1, one could show the above first term goes to 0. The second term is bounded by $Cn^{-2}\nu$, which goes to 0 too. Thus we complete the proof to the third result in the theorem. The last result could be proved in the same fashion, we omit it here.

Proof to Theorem 2.3: The proof to the first two results in Theorem 2.3 is the same as that the proof to Theorem 2.1, since the result we used from Giraitis and Koul (2013) holds for both short and long memory cases. Therefore, we only need to prove the covariance result in Theorem 2.3. Similar to short memory case, we write

$$\begin{aligned}
\operatorname{Cov}\left(n^{1/2}\tilde{\beta}_{k_1},\ n^{1/2}\tilde{\beta}_{k_2}\right) &= \frac{8\pi}{n} \sum_{j=1}^{\nu} \cos(k_1\mu_j)\cos(k_2\mu_j) \operatorname{Var}\left(\frac{I_X(\mu_j)}{f(\mu_j,\boldsymbol{\theta})}\right) \\
&\quad + \frac{8\pi}{n} \sum\sum_{j_1 \neq j_2} \cos(k_1\mu_{j_1})\cos(k_2\mu_{j_2}) \operatorname{Cov}\left(\frac{I_X(\mu_{j_1})}{f(\mu_{j_1},\boldsymbol{\theta})},\ \frac{I_X(\mu_{j_2})}{f(\mu_{j_2},\boldsymbol{\theta})}\right) \\
&=: J_{1n} + J_{2n},
\end{aligned}$$

As to term J_{1n}, consider the Bartlett approximation for a standardized periodogram $I_X(\mu_j)/f(\mu_j,\boldsymbol{\theta})$ of a linear process, we can write it as in Theorem 5.3.2 in Koul, et al (2012, p.128)

$$\frac{I_X(\mu_j)}{f(\mu_j,\boldsymbol{\theta})} = \frac{I_\zeta(\mu_j)}{f_\zeta(\mu_j,\boldsymbol{\theta})} + r_{n,j}.$$

Thus, we have

$$\operatorname{Var}\left(\frac{I_X(\mu_j)}{f(\mu_j,\boldsymbol{\theta})}\right) = \operatorname{Var}\left(\frac{I_\zeta(\mu_j)}{f_\zeta(\mu_j,\boldsymbol{\theta})}\right) + 2\operatorname{Cov}\left(\frac{I_\zeta(\mu_j)}{f_\zeta(\mu_j,\boldsymbol{\theta})},\ r_{n,j}\right) + \operatorname{Cov}\left(r_{n,j},\ r_{n,j}\right). \tag{5.5}$$

Hence, above J_{1n} could be written as $J_{1n} = J_{11n} + J_{12n} + J_{13n}$ in terms of (5.5). From the proof to Theorem 5.3.2 in Koul, et al (2012, p.131), we have $\operatorname{Var}(I_\zeta(\mu_j)/f_\zeta(\mu_j,\boldsymbol{\theta})) = 1 + O(n^{-1})$. Thus, apply the same argument as that to I_{1n}, we prove $J_{11n} \to 0$ as $n \to \infty$. From $E|r_{n,j}|^2 \le C\log j/j$ in Theorem 5.3.2 in Koul, et al (2012, p.128), it is easy to see that $|J_{13n}| \le Cn^{-1}\log^2 n \to 0$. As to J_{12n}, from Cauchy-Schwarz inequality, we have $|\operatorname{Cov}(I_\zeta(\mu_j)/f_\zeta(\mu_j,\boldsymbol{\theta}),\ r_{n,j})| \le Cj^{-1/2}\log^{1/2} j$. Thus, we have $|J_{12n}| \le Cn^{-1}(\log n)^{1/2} n^{1/2} \to 0$. Hence, we prove that $J_{1n} \to 0$.

As to J_{2n}, we have

$$
\begin{aligned}
|J_{2n}| &\leq \frac{C}{n} \sum_{j_1=1}^{n} \sum_{j_2>j_1}^{n} \left| \mathrm{Cov}\left(\frac{I_X(\mu_{j_1})}{f(\mu_{j_1}, \boldsymbol{\theta})}, \ \frac{I_X(\mu_{j_2})}{f(\mu_{j_2}, \boldsymbol{\theta})} \right) \right| \\
&\leq \frac{C}{n} \sum_{j_1=1}^{n} j_1^{-2d} \sum_{j_2>j_1}^{n} j_2^{2d-2} (\log j_2)^2 \\
&\leq \frac{C(\log n)^2}{n} \sum_{j_1=1}^{n} j_1^{-2d} j_1^{2d-1} \\
&\leq \frac{C(\log n)^3}{n}.
\end{aligned}
$$

The above 2nd inequality follows from Lemma 5.4. Thus, we prove that $J_{2n} \to 0$. Hence, we complete the proof.

REFERENCES

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods.* Second edition. Springer Series in Statistics. Springer-Verlag, New York, 1991.

Chen, W. and Deo, R. S. (2004). A generalized Portmanteau goodness-of-fit test for time series models. *Econometric Theory* **20**, 382–416.

Fragkeskou, M. and Paparoditis, E. (2016). Inference for the fourth-order innovation cumulant in linear time series. *J. Time Series Anal.* **37**, 240–266.

Giraitis, L. and Koul, H. L. (2013). On asymptotic distributions of weighted sums of periodograms. *Bernoulli* **19**, 2389–2413.

Giraitis, L.; Koul, H. L. and Surgailis, D. *Large sample inference for long memory processes.* Imperial College Press, London, 2012.

Hurvich, C. M., Deo, R. and Brodsky, J. (1998). The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long-memory time series. *J. Time Ser. Anal.* **19**, 19–46.

Krogstad, H. E. (1982). On the covariance of the periodogram. *J. Time Ser. Anal.* **3**, 195–207.

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

Robinson, P. M. (1995). Log-periodogram regression of time series with long range dependence. *Ann. Statist.* **23**, 1048–1072.

Rosenblatt M. *Stationary Sequences and Random Fields.* Boston: Birkhäuser, 1985.