COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 1267

POOLED LOG PERIODOGRAM REGRESSION

Katsumi Shimotsu and Peter C. B. Phillips

July 2000

Pooled Log Periodogram Regression*

Katsumi Shimotsu
Department of Economics
Yale University

and

Peter C. B. Phillips
Cowles Foundation for Research in Economics
Yale University

12 June 2000

Abstract

Estimation of the memory parameter in time series with long range dependence is considered. A pooled log periodogram regression estimator is proposed that utilizes a set of mL periodogram ordinates with $L \to \infty$ rather than m ordinates used in the conventional log periodogram estimator. Consistency and asymptotic normality of the pooled regression estimator are established. The pooled estimator is shown to have smaller variance but larger bias than the conventional log periodogram estimator. Finite sample performance is assessed in simulations, and the methods are illustrated in an empirical application with inflation and stock returns.

*Thanks go to a referee and associate editor for helpful comments on an earlier version of the paper. Shimotsu thanks Donald Andrews and Oliver Linton for valuable comments and the Cowles Foundation for support under a Cowles prize. Phillips thanks the NSF for research support under Grant No. SBR97-30295. Simulations were performed in MATLAB. The paper was keyboarded by the authors in Scientific Word 2.5.

1 Introduction

The model we work with is a stationary Gaussian long-memory process X_t whose spectral density has the form

$$f_{XX}(\lambda) = \left|1 - e^{i\lambda}\right|^{-2d} f_{uu}(\lambda), \tag{1}$$

where $-1/2 < d < 1/2$ and $f_{uu}(\lambda)$ is a symmetric, periodic (with period 2π), positive, and continuous function bounded above and away from zero with a finite third derivative. Our objective is to estimate the parameter d in (1), which governs the long-memory property of X_t. The time domain version of (1) has the form $(1-L)^d X_t = u_t$, where u_t is a covariance stationary time series with spectral density $f_{uu}(\lambda)$. It is often preferable to leave the precise generating mechanism of u_t unspecified, so that the treatment of u_t is nonparametric. The estimation of d then falls within the framework of semiparametric methods. The most common estimator for d in this framework is provided by log periodogram regression, which was proposed by Geweke and Porter-Hudak (1983) and is sometimes called the GPH estimator. Rigorous analysis by Künsch (1986), Robinson (1995), and, most recently, Hurvich, Deo, and Brodsky (1998) followed the earlier work and established asymptotic properties of the estimator, including consistency and asymptotic normality and an optimal formula for the choice of the number of periodogram ordinates used in the regression. There is now a large and growing literature on the subject, the estimator is commonly used in empirical work, especially in economics, and it offers the computational convenience of least squares regression.

In view of (1), we have the following relation between the spectral density of X_t and u_t in logarithmic form

$$\ln(f_{XX}(\lambda)) = -2d \ln\left|1 - e^{i\lambda}\right| + \ln(f_{uu}(\lambda)).$$

Using periodogram ordinates in place of the actual spectra and evaluating these at the fundamental frequencies $\lambda_s = \frac{2\pi s}{n}$, $s = 1, ..., n-1$ leads to the 'regression' relationship

$$\ln(I_{XX}(\lambda_s)) = -2d \ln\left|1 - e^{i\lambda_s}\right| + \ln(f_{uu}(\lambda_s)) + U(\lambda_s), \tag{2}$$

where $I_{XX}(\lambda_s) = w_X(\lambda_s) w_X(\lambda_s)^*$ and $w_X(\lambda_s)$ is the discrete Fourier transform $w_X(\lambda_s) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} X_t e^{it\lambda_s}$. The error in (2) is

$$U(\lambda_s) = \ln\left[\frac{I_{XX}(\lambda_s)}{f_{XX}(\lambda_s)}\right].$$

By virtue of the continuity of f_{uu}, $f_{uu}(\lambda_s)$ is effectively constant for frequencies in a shrinking band around the origin. This motivates the log-periodogram regression estimator of d, which is based on a linear least squares regression over frequencies $s = \ell + 1, ..., m$ (with ℓ a trimming number and m a truncation number) leading to

$$\ln(I_{XX}(\lambda_s)) = -2\widehat{d}(\ell) \ln\left|1 - e^{i\lambda_s}\right| + \widehat{\mu} + \text{error} \tag{3}$$

and a class of estimates $\widehat{d}(\ell)$ and $\widehat{\mu}$ that depend on a subset of $m-\ell$ frequencies. Under the rate condition $\frac{\ell}{m}+\frac{m}{n} \to 0$, the regression effectively uses $O(m)$ periodogram ordinates as $n \to \infty$.

The heuristic motivation for this regression is based on the idea that the errors $U(\lambda_s)$ in (2) would be asymptotically independent across frequencies if the spectrum $f_{XX}(\lambda)$ were bounded. But, in the present case that is not so, and the errors $U(\lambda_s)$ are asymptotically correlated as shown by Künsch (1986), a feature that suggests the trimming of some (ℓ) observations away from the origin. Robinson (1995) proved that $\widehat{d}(\ell)$ is consistent and asymptotically normally distributed under some additional conditions on ℓ, m, and n. Hurvich, Deo, and Brodsky (1998) derived the asymptotic bias, variance, and the mean squared error of $\widehat{d}(0)$, the original GPH estimator, and showed under some stronger conditions on m, n, and $f_{uu}(\lambda)$ that

$$\sqrt{m}\left(\widehat{d}(0)-d\right) \to_d N\left(0, \frac{\pi^2}{24}\right).$$

The GPH estimator achieves consistency and asymptotic normality by using only m periodogram ordinates at frequencies $\lambda_s = 2\pi/n, ..., 2\pi m/n$ with $m/n \to 0$. The truncation at λ_m implies that as n increases, the estimator uses a smaller and smaller proportion of the full frequency band $[0,\pi]$, so that the effective band shrinks to the origin. The shrinking process is deliberate in the design of the GPH estimator because, given the nonparametric specification of $f_{uu}(\lambda)$, it is natural to confine attention in the regression to an immediate neighbourhood of the origin $\lambda \sim 0$, because in this case (1) has the simpler asymptotic form $f_{XX}(\lambda) \sim \lambda^{-2d}G$ as $\lambda \to 0+$, with $G = f_{uu}(0)$ constant. However, as is apparent from (2), the periodogram at higher frequencies λ_s $(s = m+1, .., [n/2])$ continues to contain some information about d, although the intercept involves $f_{uu}(\lambda_s)$ and will now vary over frequency bands to the extent that $f_{uu}(\lambda)$ is not constant. This intuition indicates that conventional log periodogram regression may discard some information in the data and gains may be achieved by using more frequency bands while at the same time allowing for variation in $f_{uu}(\lambda)$.

Accordingly, we now propose a new procedure for estimating d that builds on this idea. The method is a pooled log periodogram regression that is taken over the wider band of frequencies $\lambda_s = \frac{2\pi s}{n}, s = 1, ..., mL$ with $L \to \infty$ and $mL/n \to 0$. This method corrects for variation in the regression intercept by taking subgroup means in the regression, so that it allows that the error spectrum $f_{uu}(\lambda)$ may be nonconstant across bands. The new estimator treats $\ln(f_{uu}(\lambda_s))$ in (2) as an infinite dimensional nuisance parameter appearing in the regression intercept. The approach taken is then analogous to the treatment of fixed effects in panel data regression. The estimator of d pools the information about d obtained within each (shrinking) band over which the error spectrum is effectively constant as $n \to \infty$. We therefore call the new estimator a pooled log periodogram regression estimator.

The pooled estimator is shown to be consistent and asymptotically normally distributed. The pooled estimator has a smaller asymptotic variance than the GPH estimator, reflecting the greater number of periodogram ordinates used in the regression, but it also has larger asymptotic bias because of the nonconstancy of $f_{uu}(\lambda)$.

Simulations show that in finite samples the pooled estimator performs substantially better than the GPH estimator when $f_{uu}(\lambda)$ has spectral peaks near the origin. On the other hand, the pooled estimator generally performs worse than the GPH estimator when $f_{uu}(\lambda)$ changes monotonically from $\lambda = 0$ to $\lambda = \pi$, although the difference is small.

The paper is organized as follows. The new estimator is constructed in Section 2. Section 3 gives assumptions and derives some preliminary asymptotic results. Section 4 proves consistency of the pooled estimator and derives its asymptotic mean squared error. Section 5 demonstrates asymptotic normality. Section 6 discusses the optimal choice of m. Section 7 discusses the simulation results and gives an empirical illustration. Proofs are collected in Section 8.

2 Pooling Log Periodogram Ordinates in Regression

The idea of pooling ordinates in log periodogram regression can be explained as follows. First, we use an alternate form of the log periodogram representation, viz.

$$\begin{aligned} \ln\left(I_{XX}(\lambda_s)\right) &= \ln\left(f_{uu}(\lambda_s)\right) + \ln\left|1 - e^{i\lambda_s}\right|^{-2d} + \ln\left(\frac{I_{XX}(\lambda_s)}{f_{XX}(\lambda_s)}\right) \qquad (4) \\ &= \ln\left(f_{uu}(\omega_j)\right) - 2d\ln\left|1 - e^{i\lambda_s}\right| + \ln\left(\frac{I_{XX}(\lambda_s)}{f_{XX}(\lambda_s)}\right) + \ln\left(\frac{f_{uu}(\lambda_s)}{f_{uu}(\omega_j)}\right), \end{aligned}$$

which allows for periodogram ordinates λ_s in the neighbourhood of a set of frequencies ω_j for $j = 0, 1, ...M - 1$, where M is a parameter that determines the total number of distinct bands.

The implied log periodogram relation is now

$$\ln\left(I_{XX}(\lambda_s)\right) = \ln\left(f_{uu}(\omega_j)\right) - 2d\ln\left|1 - e^{i\lambda_s}\right| + V(\lambda_s), \quad \lambda_s \in B_j \qquad (5)$$

where

$$V(\lambda_s) = \ln\left(\frac{I_{XX}(\lambda_s)}{f_{XX}(\lambda_s)}\right) + \ln\left(\frac{f_{uu}(\lambda_s)}{f_{uu}(\omega_j)}\right),$$

and

$$B_j = \begin{cases} \left\{\lambda_s | \omega_j - \frac{\pi}{2M} < \lambda_s \le \omega_j + \frac{\pi}{2M}\right\}, & \omega_j = \frac{(2j+1)\pi}{2M}, j = 1, ..., M-1 \\ \left\{\lambda_s | 0 < \lambda_s \le \frac{\pi}{M}\right\}, & \omega_0 = 0, j = 0 \end{cases}$$

are the frequency bands, which are of width $\frac{\pi}{M}$. We compute the regressor sequence in (5) using $\left|1 - e^{i\lambda_s}\right|^2 = 4\sin^2\left(\frac{\lambda_s}{2}\right)$ and do not use the conventional replacement $\left|1 - e^{i\lambda_s}\right|^2 \sim \lambda^2$, which is appropriate only for λ_s in the vicinity of the zero frequency.

We propose to estimate the parameter d in (5) by linear regression using $(L+1)$ bands $B_0, ..., B_L$ where L is a number such that $L \to \infty$ and $L/M \to 0$. Thereby we remove the intercept in the regression by pooling observations over bands to fit d. Write (5) as

$$Y_{sj} = \mu_j + dX_{sj} + \eta_{sj} + \varepsilon_{sj} \quad s = 1, ..., m; j = 0, 1, ..., L \qquad (6)$$

with

$$
\begin{aligned}
Y_{sj} &= \ln\left(I_{XX}\left(\lambda_s\right)\right), \quad \lambda_s \in B_j \\
X_{sj} &= -2\ln\left|1-e^{i\lambda_s}\right| = -\ln\left[4\sin^2\left(\frac{\lambda_s}{2}\right)\right] = -\ln\left(2-2\cos\lambda_s\right), \quad \lambda_s \in B_j \\
\eta_{sj} &= \ln\left(\frac{f_{uu}\left(\lambda_s\right)}{f_{uu}\left(\omega_j\right)}\right) = \ln f_{uu}\left(\lambda_s\right) - \ln f_{uu}\left(\omega_j\right), \quad \lambda_s \in B_j \\
\varepsilon_{sj} &= \ln\left(\frac{I_{XX}\left(\lambda_s\right)}{f_{XX}\left(\lambda_s\right)}\right) - \psi\left(1\right), \quad \lambda_s \in B_j \\
\mu_j &= \ln\left(f_{uu}\left(\omega_j\right)\right) + \psi\left(1\right),
\end{aligned}
$$

where $\psi\left(1\right) = \Gamma'\left(1\right) = -\gamma$ and $\gamma = 0.57721566...$ is Euler's constant.

The pooled estimator $\widehat{d}$ is given by the formula

$$
\widehat{d} = \frac{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\left(Y_{sj}-\overline{Y}_{.j}\right)\left(X_{sj}-\overline{X}_{.j}\right)}{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{.j}\right)^2} = \frac{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}Y_{sj}\left(X_{sj}-\overline{X}_{.j}\right)}{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{.j}\right)^2}, \tag{7}
$$

where

$$
\begin{aligned}
\overline{Y}_{.j} &= \frac{1}{m}\sum_{\{s:\lambda_s\in B_j\}} Y_{sj} = \frac{1}{m}\sum_{\{s:\lambda_s\in B_j\}} \ln\left(I_{XX}\left(\lambda_s\right)\right), \\
\overline{X}_{.j} &= \frac{1}{m}\sum_{\{s:\lambda_s\in B_j\}} X_{sj} = -\frac{1}{m}\sum_{\{s:\lambda_s\in B_j\}} \ln\left[4\sin^2\left(\frac{\lambda_s}{2}\right)\right].
\end{aligned}
$$

Note that the estimator $\widehat{d}$ uses data over an increasing number of frequency bands, not just those frequencies in B_0. The estimator still uses frequencies only in the vicinity of the origin because $mL/n \to 0$. In other words, the pooled estimator retains semiparametric nature of the log periodogram regression while using increasing number of bands. Subgroup means are subtracted in order to allow for the fact that the intercept μ_j may change over frequency bands $j = 0, ..., L$.

Combining equations (6) and (7) gives the estimation error

$$
\widehat{d} - d = \frac{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\eta_{sj}\left(X_{sj}-\overline{X}_{.j}\right) + \sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right)}{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{.j}\right)^2}. \tag{8}
$$

The idea of pooling ordinates over the bands B_j while allowing for variation in the spectrum across bands can be applied to other semiparametric estimators of the long memory parameter d. In particular, it is readily implemented in the case of the local Whittle estimator suggested by Künsch (1986) to give a pooled Whittle estimator. Our attention in the present paper, however, will be confined to the pooled log periodogram procedure.

3 Assumptions and Asymptotic Results

To establish a limit theory for the pooled estimator, many of the results in Robinson (1995) and Hurvich, Deo, and Brodsky (1998) are relevant, and our approach draws substantially on their earlier work. We start by introducing the following assumptions.

Assumption 1 $m \to \infty$, $n \to \infty$, $\frac{m}{n} = \frac{1}{2M} \to 0$.

Assumption 2 $\frac{M}{m} + \frac{m \ln m}{n} + \frac{\ln^2 n}{m} \to 0$.

Assumption 3 $f'_{uu}(0) = 0$, $f_{uu}(\omega) > B_0 > 0, |f'_{uu}(\omega)| < B_1 < \infty, |f''_{uu}(\omega)| < B_2 < \infty$, $|f'''_{uu}(\omega)| < B_3 < \infty \ \forall \omega \in [0, \pi]$.

3.1 Remark In what follows, it will be taken as a convention that $n = 2mM$ holds exactly with both m and M integers (so that n is even). The convention is convenient, but not essential in what follows, and M is simply defined by the ratio $M = \frac{n}{2m}$ when n is odd. Any particular choice of m and expansion rate for m affect the bandwidth $\frac{\pi}{M}$, its rate of contraction, and the number of bands in the regression. The rate condition in Assumption 2 controls the relative rates at which m, M, and $n \to \infty$. Assumption 3 implies that $f_{uu}(\omega)$ is bounded away from zero and smooth with finite third derivative, much as in Hurvich et al. (1998).

3.2 Lemma *For a number ℓ such that $\ell \to \infty$ and $\ell^5/M^4 \to 0$, the following results hold:*

(a) $\sum_{\{s:\lambda_s \in B_0\}} \left(X_{sj} - \overline{X}_{.j}\right)^2 = 4m + o(m)$.

(b) $\sum_{j=1}^{\ell} \sum_{\{s:\lambda_s \in B_j\}} \left(X_{sj} - \overline{X}_{.j}\right)^2 = 4m\Xi + o(m)$,

where $\Xi = \sum_{j=1}^{\infty} \left[-j(j+1)(\ln(j+1) - \ln j)^2 + 1 \right] \doteqdot 0.0803$.

3.3 Lemma *For a number ℓ such that $\ell/M \to 0$,*

$$\sum_{j=0}^{\ell} \sum_{\{s:\lambda_s \in B_j\}} \eta_{sj} \left(X_{sj} - \overline{X}_{.j}\right) = -\frac{2\pi^2}{3} \frac{f''_{uu}(0)}{f_{uu}(0)} \frac{m^3 \ell}{n^2} + o\left(\frac{m^3 \ell}{n^2}\right).$$

3.4 Remarks

(a) The GPH estimator involves regression only over the band B_0. In place of (8) it satisfies

$$\widehat{d}_{GPH} - d = \frac{\sum_{\{s:\lambda_s \in B_0\}} \eta_{sj} \left(X_{sj} - \overline{X}_{.j}\right) + \sum_{\{s:\lambda_s \in B_0\}} \varepsilon_{sj} \left(X_{sj} - \overline{X}_{.j}\right)}{\sum_{\{s:\lambda_s \in B_0\}} \left(X_{sj} - \overline{X}_{.j}\right)^2}. \tag{9}$$

The denominator of $\widehat{d} - d$

$$\sum_{j=0}^{L} \sum_{\{s:\lambda_s \in B_j\}} \left(X_{sj} - \overline{X}_{.j}\right)^2 = 4\left(1 + \Xi\right) m + o\left(m\right), \tag{10}$$

is larger than the denominator of the GPH estimator

$$\sum_{\{s:\lambda_s \in B_0\}} \left(X_{sj} - \overline{X}_{.j}\right)^2 = 4m + o\left(m\right),$$

as $m \to \infty$. Roughly speaking, the denominator measures the excitation level of the regressors and indicates the information content in the regressors about the coefficient (d) in the regression (6). From (10), it is apparent that this information content is larger when the frequency band $B_0, ..., B_L$ is employed than when the immediate band around the zero frequency B_0 is used. As we will see, this increase in information content reduces the asymptotic variance of the pooled estimator relative to that of the GPH estimator.

(b) The optimal expansion rate of m for the GPH estimator is known from Hurvich et al. (1998) to be $O(n^{\frac{4}{5}})$, whereas the optimal rate for the pooled estimator is, as we will see later, $O(n^{\frac{4}{5}} L^{-\frac{2}{5}})$. Thus, if optimal rates were chosen the variance gains of the pooled estimator would vanish as $n \to \infty$. Issues of a joint optimal choice of m and L have not been considered by the authors.

(c) Lemma 3.3 shows that the nonrandom bias of the pooled estimator that arises from the presence of the first term in the numerator of (8) is $O\left(\frac{m^2 L}{n^2}\right)$ when $f''_{uu}\left(\lambda\right) \neq 0$ and hence this bias tends to zero as $n \to \infty$.

4 MSE and Consistency

We start with the following theorem, which is a variant of theorem 2 in Robinson (1995).

4.1 Theorem *Let Assumption 3 hold. Then, for any sequences of positive integers $j = j\left(n\right)$ and $k = k\left(n\right)$ such that $0 < k < j < n/2$, as $n \to \infty$*

(*a*) $E\left[w\left(\lambda_j\right) \overline{w}\left(\lambda_j\right) / f_{XX}\left(\lambda_j\right)\right] = 1 + O\left(j^{-1} \ln n\right),$

(*b*) $E\left[w\left(\lambda_j\right) w\left(\lambda_j\right) / f_{XX}\left(\lambda_j\right)\right] = O\left(j^{-1} \ln n\right),$

(*c*) $E\left[w\left(\lambda_j\right) \overline{w}\left(\lambda_k\right) / \left(f_{XX}\left(\lambda_j\right) f_{XX}\left(\lambda_k\right)\right)^{1/2}\right] = O\left(k^{-1} \ln n\right),$

(*d*) $E\left[w\left(\lambda_j\right) w\left(\lambda_k\right) / \left(f_{XX}\left(\lambda_j\right) f_{XX}\left(\lambda_k\right)\right)^{1/2}\right] = O\left(k^{-1} \ln n\right),$

where $\lambda_j = 2\pi j/n$ *and* $w\left(\lambda\right) = \left(2\pi n\right)^{-1/2} \sum_1^n X_t e^{it\lambda}$.

Define the quantities

$$A_s = \frac{1}{\sqrt{2\pi n}} \sum_{t=0}^{n-1} X_t \cos \lambda_s t, \quad C_s = \frac{1}{\sqrt{2\pi n}} \sum_{t=0}^{n-1} X_t \sin \lambda_s t,$$

and

$$\begin{aligned}\alpha_{st} &= \max\left\{\left|\operatorname{cov}\left(A_s/f_{XX}^{1/2}(\lambda_s), A_t/f_{XX}^{1/2}(\lambda_t)\right)\right|, \left|\operatorname{cov}\left(A_s/f_{XX}^{1/2}(\lambda_s), C_t/f_{XX}^{1/2}(\lambda_t)\right)\right|,\right.\\ &\left.\left|\operatorname{cov}\left(C_s/f_{XX}^{1/2}(\lambda_s), A_t/f_{XX}^{1/2}(\lambda_t)\right)\right|, \left|\operatorname{cov}\left(C_s/f_{XX}^{1/2}(\lambda_s), C_t/f_{XX}^{1/2}(\lambda_t)\right)\right|\right\}.\end{aligned}$$

Theorem 4.1, combined with the Gaussianity of X_t, enables us to evaluate the means, variances, and covariances of

$$\varepsilon_{sj} = \ln\left(\frac{I_{XX}(\lambda_s)}{f_{XX}(\lambda_s)}\right) - \psi(1) = \ln\left(\frac{A_j^2}{f_{XX}(\lambda_j)} + \frac{C_j^2}{f_{XX}(\lambda_j)}\right) - \psi(1),$$

because the distribution of the normal vector

$$\left(A_j/\left(f_{XX}(\lambda_j)\right)^{1/2}, \quad C_j/\left(f_{XX}(\lambda_j)\right)^{1/2}, \quad A_k/\left(f_{XX}(\lambda_k)\right)^{1/2}, \quad C_k/\left(f_{XX}(\lambda_k)\right)^{1/2} \right)$$

is solely determined by its covariance matrix. In particular, it follows directly from theorem 4.1 that

$$\begin{aligned} E\left(\frac{A_j^2}{f_{XX}(\lambda_j)}\right) &= \frac{1}{2} + O\left(\frac{\ln n}{j}\right), \quad E\left(\frac{C_j^2}{f_{XX}(\lambda_j)}\right) = \frac{1}{2} + O\left(\frac{\ln n}{j}\right), \\ E\left(\frac{A_j C_j}{f_{XX}(\lambda_j)}\right) &= O\left(\frac{\ln n}{j}\right), \end{aligned}$$

uniformly for $1 \le t < s < n/2$.

The following lemma is also a consequence of theorem 4.1. Because the proofs of its four component parts are very similar to those of lemmas 2,3,6, and 7 in Hurvich, Deo, and Brodsky (1998), they are omitted here.

4.2 Lemma

(a) $\alpha_{st} = O(\ln n/t)$, *uniformly for* $1 \le t < s < n/2$.
(b) $Cov(\varepsilon_{sj}, \varepsilon_{tk}) = O\left(\alpha_{st}^2\right)$, *uniformly for* $\ln^2 n \le t < s < n/2$.
(c) $E(\varepsilon_{sj}) = O(\ln n/s)$, *uniformly for* $\ln^2 n \le s < n/2$.
(d) $Var(\varepsilon_{sj}) = \pi^2/6 + O(\ln n/s)$, *uniformly for* $\ln^2 n \le s < n/2$.

Lemma 4.2 shows that $\varepsilon_{sj}, \varepsilon_{tk}$ are asymptotically mean zero and independent and identically distributed for $\ln^2 n \le s,\ t \le n/2$. We now proceed to derive asymptotic representations of the bias and variances and covariances of $\sum_{\{s:\lambda_s \in B_j\}} \varepsilon_{sj}\left(X_{sj} - \overline{X}_{.j}\right)$ over different frequency bands.

4.3 Lemma (Bias) *For a number* $\ell < M$,

$$\sum_{j=1}^{\ell} \sum_{\{s:\lambda_s \in B_j\}} E(\varepsilon_{sj})\left(X_{sj} - \overline{X}_{.j}\right) = O(\ln m).$$

4.4 Lemma (Variance and covariances between bands B_j, B_k, $1 \le k < j < M$)

For $1 \le k < j < M$,

$$Var\left[\sum_{\{s:\lambda_s \in B_j\}} \varepsilon_{sj}\left(X_{sj} - \overline{X}_{.j}\right)\right] = \frac{\pi^2}{6}\sum_{\{s:\lambda_s \in B_j\}}\left(X_{sj} - \overline{X}_{.j}\right)^2 + O\left(\frac{\ln^2 m}{j^3}\right),$$

and

$$Cov\left[\sum_{\{s:\lambda_s \in B_j\}} \varepsilon_{sj}\left(X_{sj} - \overline{X}_{.j}\right), \sum_{\{t:\lambda_t \in B_k\}} \varepsilon_{tk}\left(X_{tk} - \overline{X}_{.k}\right)\right] = O\left(\frac{\ln^2 m}{jk^2}\right).$$

4.5 Lemma (Covariances between bands B_j, B_0, $1 \le j < M$)

$$Cov\left[\sum_{\{s:\lambda_s \in B_j\}} \varepsilon_{sj}\left(X_{sj} - \overline{X}_{.j}\right), \sum_{\{t:\lambda_t \in B_0\}} \varepsilon_{t0}\left(X_{t0} - \overline{X}_{.0}\right)\right] = O\left(\frac{m^{1/2}\ln^7 m}{j}\right) + O\left(\frac{m}{j\ln^3 m}\right).$$

4.6 Lemma (Asymptotic variance)

$$Var\left[\sum_{j=0}^{\ell}\sum_{\{s:\lambda_s \in B_j\}} \varepsilon_{sj}\left(X_{sj} - \overline{X}_{.j}\right)\right] = \frac{4\pi^2 m}{6}\left(1 + \Xi\right) + o\left(m\right).$$

4.7 Remark We can express $\widehat{d}$ as

$$\begin{aligned}
\widehat{d} &= \frac{\sum_{j=0}^{L}\sum_{\{s:\lambda_s \in B_j\}} \ln\left(I_{XX}\left(\lambda_s\right)\right)\left(X_{sj} - \overline{X}_{.j}\right)}{\sum_{j=0}^{L}\sum_{\{s:\lambda_s \in B_j\}}\left(X_{sj} - \overline{X}_{.j}\right)^2} \\
&= \sum_{j=0}^{L}\left(\frac{\sum_{\{s:\lambda_s \in B_j\}}\left(X_{sj} - \overline{X}_{.j}\right)^2}{\sum_{j=0}^{\ell}\sum_{\{s:\lambda_s \in B_j\}}\left(X_{sj} - \overline{X}_{.j}\right)^2}\frac{\sum_{\{s:\lambda_s \in B_j\}} \ln\left(I_{XX}\left(\lambda_s\right)\right)\left(X_{sj} - \overline{X}_{.j}\right)}{\sum_{\{s:\lambda_s \in B_j\}}\left(X_{sj} - \overline{X}_{.j}\right)^2}\right) \\
&= \sum_{j=0}^{L}\left(\frac{\sum_{\{s:\lambda_s \in B_j\}}\left(X_{sj} - \overline{X}_{.j}\right)^2}{\sum_{j=0}^{L}\sum_{\{s:\lambda_s \in B_j\}}\left(X_{sj} - \overline{X}_{.j}\right)^2}\right)\widehat{d}_j,
\end{aligned}$$

where $\widehat{d}_j$ is the estimator of d obtained from using the band B_j only. Lemmas 4.4, 4.5, and 4.6 imply that the $\widehat{d}_j$ are asymptotically independent. Therefore, $\widehat{d}$ is a weighted average of asymptotically independent component estimators $\widehat{d}_j$, and we may therefore anticipate that the variance of $\widehat{d}$ is smaller than that of $\widehat{d}_0 \equiv \widehat{d}_{GPH}$.

Specifically, lemmas 3.2, 3.3, 4.3, and 4.6 yield an asymptotic representation of the mean squared error of $\widehat{d}$, which is given in the next theorem.

4.8 Theorem *Let assumptions 1-3 hold. Then*

$$E\left(\widehat{d}-d\right) = -\frac{\pi^2}{6\left(1+\Xi\right)}\frac{f''_{uu}\left(0\right)}{f_{uu}\left(0\right)}\frac{m^2L}{n^2}+o\left(\frac{m^2L}{n^2}\right)+O\left(\frac{\ln^3 m}{m}\right), \qquad (11)$$

$$Var\left(\widehat{d}\right) = \frac{\pi^2}{24\left(1+\Xi\right)m}+o\left(\frac{1}{m}\right),$$

$$MSE\left(\widehat{d}\right) = \frac{\pi^4}{36\left(1+\Xi\right)^2}\left\{\frac{f''_{uu}\left(0\right)}{f_{uu}\left(0\right)}\right\}^2\frac{m^4L^2}{n^4}+\frac{\pi^2}{24\left(1+\Xi\right)m} \qquad (12)$$
$$+o\left(\frac{m^4L^2}{n^4}\right)+O\left(\frac{\ln^6 m}{m^2}\right)+O\left(\frac{mL\ln^3 m}{n^2}\right)+o\left(\frac{1}{m}\right).$$

4.9 Remarks

(a) The mean squared error tends to zero as $n \to \infty$ and $\widehat{d}$ is consistent.

(b) Hurvich, Deo, and Brodsky (1998) derive the following formulae for the asymptotic bias and variance of the GPH estimator:

$$E\left(\widehat{d}_{GPH}-d\right) = -\frac{2\pi^2}{9}\frac{f''_{uu}\left(0\right)}{f_{uu}\left(0\right)}\frac{m^2}{n^2}+O\left(\frac{\ln^3 m}{m}\right)+o\left(\frac{m^2}{n^2}\right),$$

$$Var\left(\widehat{d}_{GPH}\right) = \frac{\pi^2}{24m}+o\left(\frac{1}{m}\right).$$

Compared with the GPH estimator, the pooled estimator has a larger bias but smaller variance. Which effect dominates in finite samples will depend on the sample size and the shape of the error spectrum $f_{uu}\left(\omega\right)$. In the extreme case where the error spectrum is constant, $\eta_{sj}=0$, and both estimators are unbiased (the first term in the numerator of both (8) and (9) is zero).

5 Asymptotic Normality

To establish the asymptotic normality of $\widehat{d}$, we prove that the standardized quantity

$$m^{-1/2}\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right) \qquad (13)$$

with $L=O\left(\ln M\right)$ has a limiting normal distribution.

The following lemma gives the basis of the limiting distribution theory. Its proof draws heavily on the derivations in Theorem 3 of Robinson (1995) and applies the approach developed in that article to a case in which there are ℓm rather than m observations.

5.1 Lemma *Let $a_{kn}=a_k$ be a triangular array for which*

$$\max_k\left|a_k\right|=O\left(m\right), \qquad \sum_{k=1+m^{0.5+\delta}}^{\ell m}a_k^2\sim m, \qquad \sum_{k=1+m^{0.5+\delta}}^{\ell m}\left|a_k\right|^p=O\left(m\ln\ell\right), \qquad (14)$$

for all $p \geq 1$*, and let* ℓ *be a number that satisfies* $\ell \to \infty$, $\ell^2 m^2 m^{0.5+\Delta}/n^2 \to 0$ *for some* $0 < \Delta < \delta < 0.5$ *and* $\ln^K \ell/m^\Delta \to 0$ *for any* $K > 0$*. Then,*

$$\frac{1}{\sqrt{m}} \sum_{k=1+m^{0.5+\delta}}^{\ell m} a_k U_k \xrightarrow{d} N(0,\Omega) = N\left(0, \psi'(1)\right),$$

where

$$\begin{aligned} U_k &= \log\left[\left(v^R(\lambda_k)\right)^2 + \left(v^I(\lambda_k)\right)^2\right] - \psi(1) \\ &= \log\left(\frac{I_{XX}(\lambda_k)}{C_g \lambda_k^{-2d}}\right) - \psi(1) = \log\left(\frac{I_{XX}(\lambda_k)}{f_{uu}(0)\lambda_k^{-2d}}\right) - \psi(1). \end{aligned}$$

With this lemma in hand, we extract a limiting distributional result for the quantity (13). In particular, we have the following.

5.2 Lemma *Let assumptions 1-3 hold and additionally require that* $m = O\left(n^{\frac{4}{5}-\varepsilon}\right)$ *for some* $\varepsilon > 0$ *and* $L = O(\ln M)$*. Then, we have*

$$m^{-1/2} \sum_{j=0}^{L} \sum_{\{s:\lambda_s \in B_j\}} \varepsilon_{sj}\left(X_{sj} - \overline{X}_{.j}\right) \xrightarrow{d} N\left(0, 4\pi^2(1+\Xi)/6\right).$$

These preliminary results lead to the asymptotic normality of the estimator $\widehat{d}$.

5.3 Theorem *Let assumptions 1-3 hold. Moreover, if* $m = O\left(n^{\frac{4}{5}-\varepsilon}\right)$ *for some* $\varepsilon > 0$ *and* $L = O(\ln M)$*, we have*

$$m^{1/2}\left(\widehat{d} - d\right) \xrightarrow{d} N\left(0, \frac{\pi^2}{24(1+\Xi)}\right).$$

6 Simulations and Empirical Illustration

This section reports some simulations that were conducted to compare the finite sample performance of the two estimators $\widehat{d}_{GPH}$ and $\widehat{d}_{pooled}$. Because both estimators treat u_t nonparametrically, it seems desirable to examine their finite sample properties over models that allow for a variety of spectral shapes for $f_{uu}(\lambda)$. With this objective in mind, we used the following AR(2) generating mechanism for u_t

$$u_t - a_1 u_{t-1} - a_2 u_{t-2} = \varepsilon_t, \quad \varepsilon_t \sim iidN(0,1),$$

which permits a range of spectral shapes, including some with spectral peaks away from the origin. We generate the process $X_t = (1-L)^{-d} u_t$ by the algorithm of Davies and Harte (1987).

We set $m = n^{0.65}$ in the simulations. This amounts to using the frequency band $(0, 0.31\pi)$, $(0, 0.22\pi)$, and $(0, 0.09\pi)$ for sample size $n = 200$, 500, and 1000. We set

$L = 2$ in the construction of the pooled estimator. Although the pooled estimator requires a slightly stronger condition on m for asymptotic normality than the GPH estimator (specifically, $m = O(n^{\frac{4}{5}-\varepsilon})$ rather than $m = o(n^{\frac{4}{5}})$), we use the same m for comparison because m is rarely chosen to be as large as $n^{\frac{4}{5}}$ in practice.

6.1 Simulations over the (a_1, a_2) plane

First, we report some comprehensive simulations over the (a_1, a_2) parameter space, so that the effect of spectral shape on performance can be assessed. We take the region of the (a_1, a_2) plane for which u_t is stationary and use a grid with a step size of 0.1 in this plane. The bias, variance, and mean squared error (MSE) were computed using 1,000 replications. Sample size and long memory parameter were chosen to be $n = 500$ and $d = 0.3$, respectively. A second experiment, reported below, looks at performance for different values of n.

For the GPH estimator, we used the regressor $X_s = -2\ln(\lambda_s)$, instead of the exact regressor $-\ln\left(4\sin^2(\lambda_s/2)\right)$, as is common practice, and in our simulations the former regressor generally gave better results for the parameter values considered. The variances of the estimators were very similar and seemed to vary little across the different parameter values. Hence, most of the variation that appears in the MSE is due to differences in bias.

Figures 1 and 2 plot the MSE's. When a_1 and a_2 are close to the line $a_1 + a_2 = 1$, the MSE of both estimators becomes quite large. The MSE of the GPH estimator has a particularly large spike when a_1 is large and a_2 is small. The MSE of the pooled estimator also has a spike, although the magnitude of the spike is substantially smaller than that of the MSE of the GPH estimator. The MSE of the GPH estimator decreases monotonically as a_1 decreases, whereas the MSE of the pooled estimator has small bumps, especially when a_2 is small and negative.

To obtain a better idea of the differences between the two estimators, a contour plot of the MSE difference (MSE(GPH)−MSE(pooled)) is displayed in Figure 3. In general, the difference is small when it is negative, except in the area near the line $a_1 + a_2 = 1$. As expected from Figures 1 and 2, the difference is large when a_1 is large and a_2 is small. Figure 4 shows a contour plot of the logarithm of the relative efficiency $(= \log_2(MSE(GPH)/MSE(pooled)))$. The area of the (a_1, a_2) plane in which the MSE of the pooled estimator is smaller than that of the GPH estimator (i.e., the logarithm of the relative efficiency is greater than 0) is not very large. Because the pooled estimator has a larger MSE than the GPH estimator when the MSE of both estimators is small, the magnitudes of positive and negative relative efficiency are roughly equal to each other.

Figure 5 plots the spectral density of u_t for some of the more important parameter combinations in Figures 3 and 4. The values of a_1 and a_2 at these points are given in Table 2. At the points where the pooled estimator has a substantially smaller MSE than the GPH estimator, such as at B, C, and E, the spectral density has a peak near the origin. At the points where the spectral density has a peak away from the origin or changes monotonically, such as at A, D, and I, the GPH estimator has a smaller MSE than the pooled estimator. At the points F and L, where both estimators have

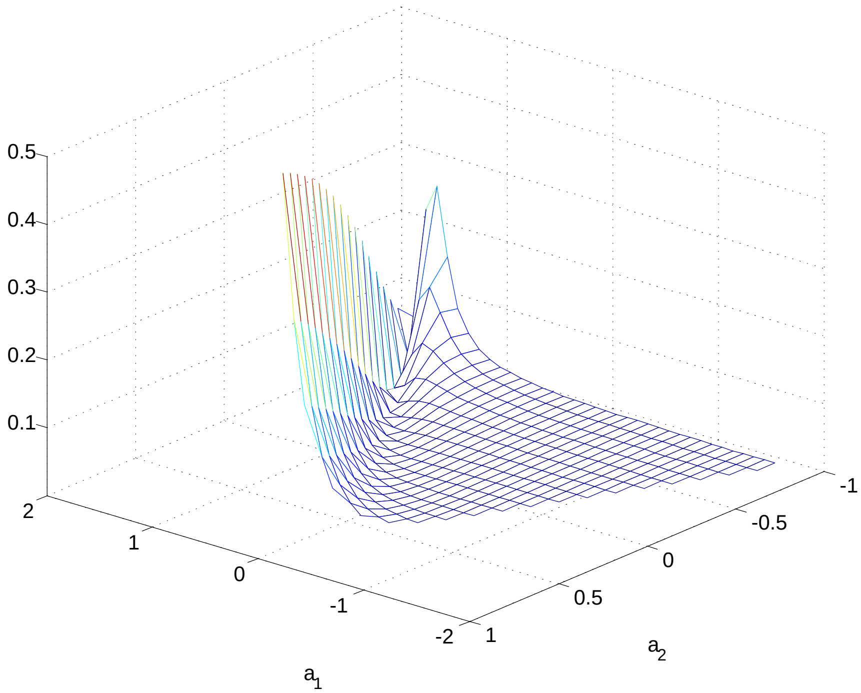

Figure 1: MSE of the GPH estimator ($d = 0.3, n = 500, m = 56$)

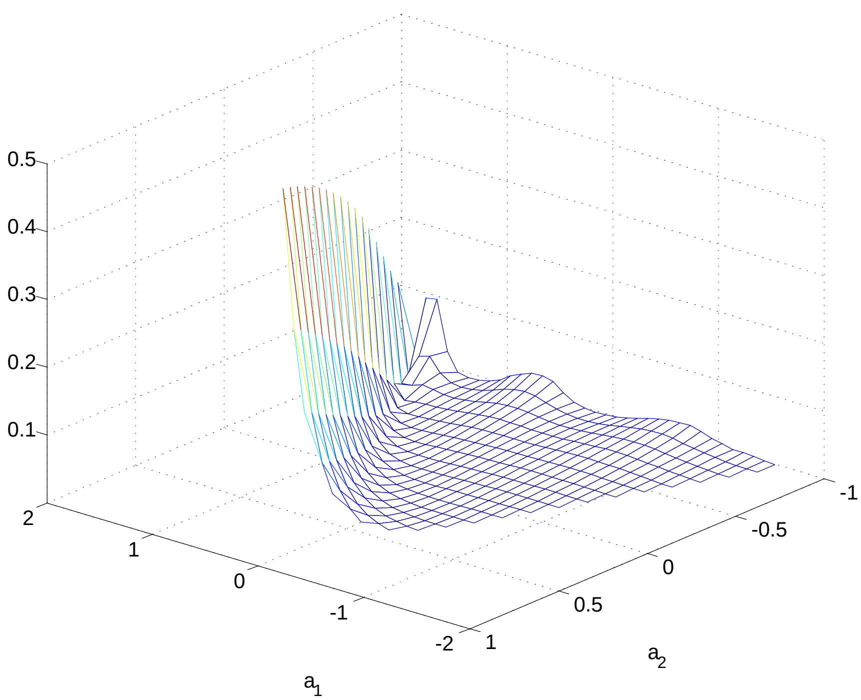

Figure 2: MSE of the pooled estimator ($d = 0.3, n = 500, m = 56$)

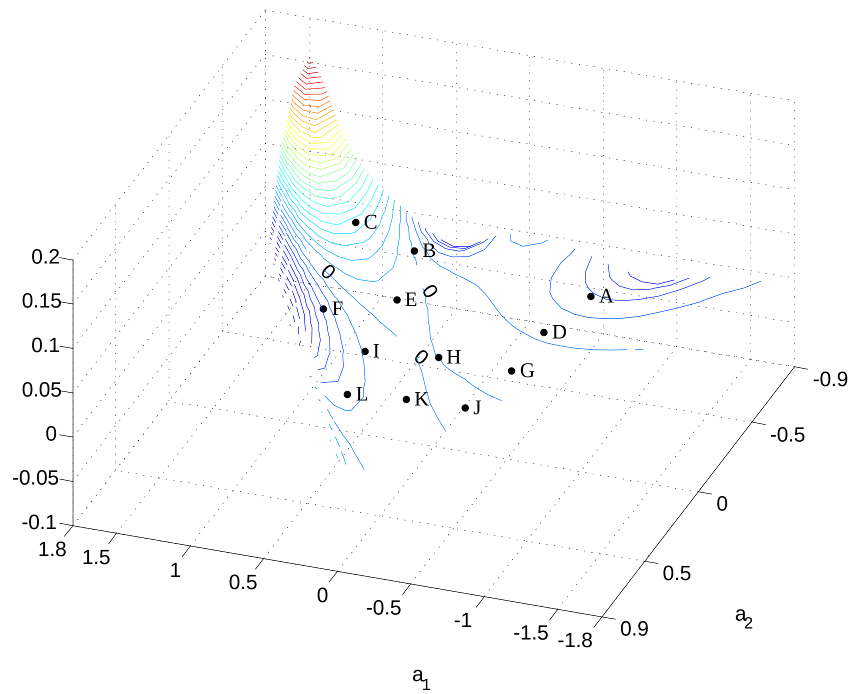

Figure 3: MSE(GPH)$-$MSE(pooled) $(d = 0.3, n = 500, m = 56)$

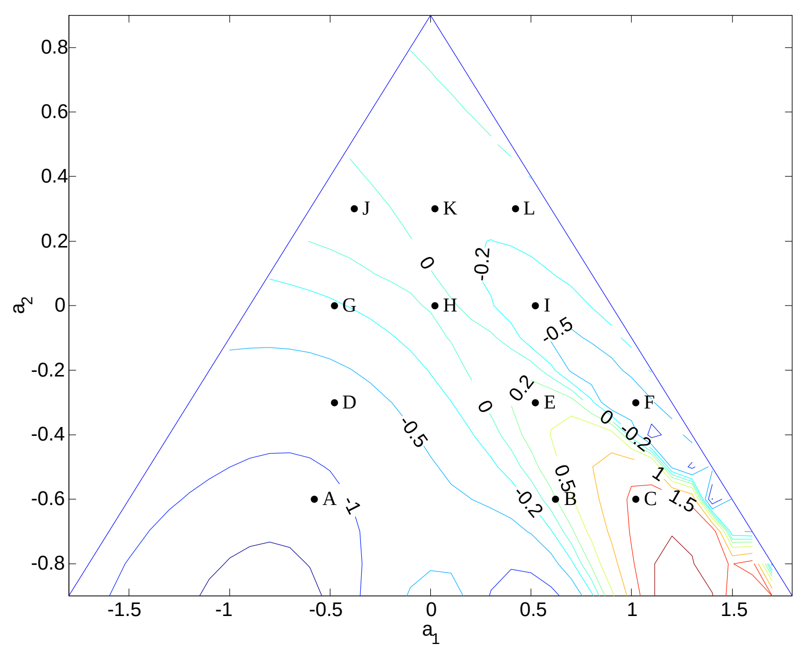

Figure 4: $\log_2 \left(MSE \left(GPH \right) / MSE \left(pooled \right) \right) (d = 0.3, n = 500, m = 56)$

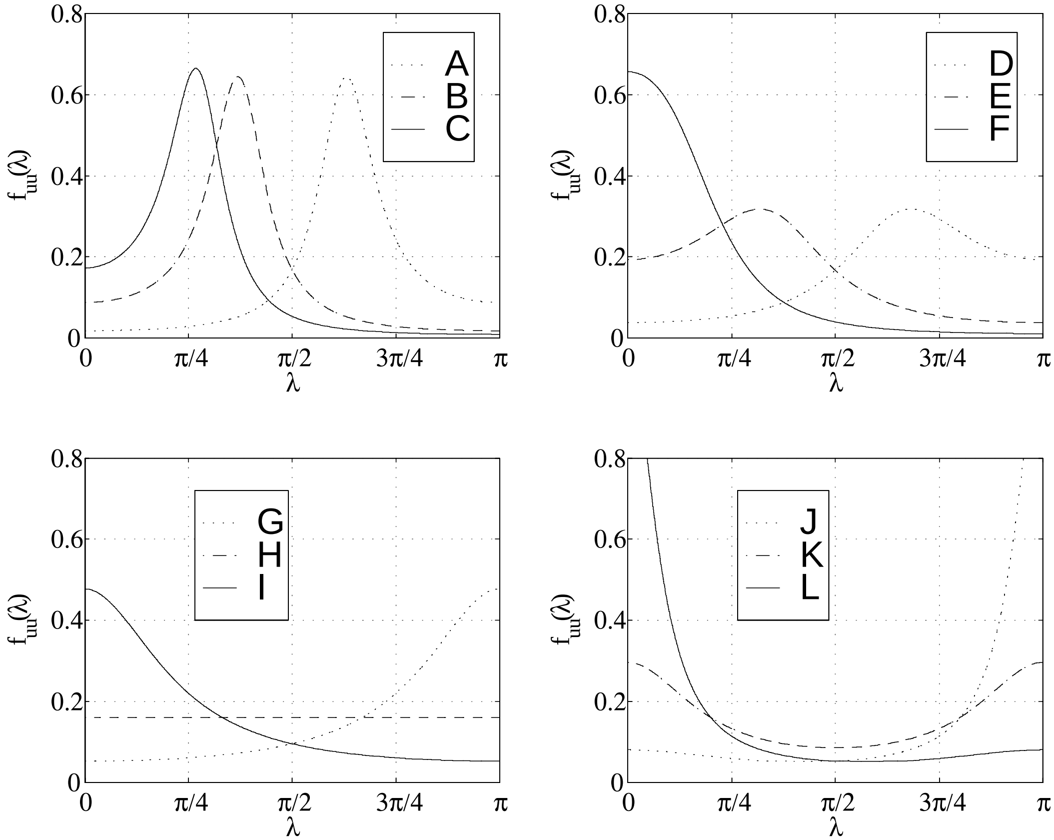

Figure 5: Spectral densities for several AR(2) processes

quite large MSE, the spectral density decreases sharply from the origin. Because the slope of the periodogram around the origin contains the strongest signal for both the GPH estimator and the pooled estimator, this result is hardly surprising, and neither procedure can be expected to work well. At point H, where u_t is a white noise, both estimators have similar MSE. The shape of the spectral density far away from the origin does not affect the MSE substantially, as indicated by the results on the points G and J.

Table 2. The values of a_1 and a_2 at the points A-L

	A	B	C	D	E	F	G	H	I	J	K	L
a_1	-0.6	0.6	1.0	-0.5	0.5	1.0	-0.5	0.0	0.5	-0.4	0.0	0.4
a_2	-0.6	-0.6	-0.6	-0.3	-0.3	-0.3	0.0	0.0	0.0	0.3	0.3	0.3

In sum, the pooled estimator has the advantage of being robust to the presence of a peak in the error spectrum $f_{uu}(\lambda)$ and produces substantial reductions in both bias and MSE when the peak is close to the regression frequency band. In such cases, of course, it is the peak in the short memory spectrum that exacerbates the bias in the GPH estimator.

6.2 Detailed simulation for several pairs of parameter values

For several points in Figures 3 and 4, we conducted a more detailed simulation covering different sample sizes[1]. Tables 3-5 show the simulation results for parameter combinations G, H, and I from Table 2, for which $a_2 = 0$ and u_t follows an AR(1) process. In general, for AR(1) processes with $a_1 \neq 0$, $\widehat{d}_{pooled}$ (the second row) is more biased than $\widehat{d}_{GPH}$ (the first row). The increase in the bias occurs because $f_{uu}(\lambda)$ is monotonically increasing (decreasing), and the bias effect of the nonzero slope of $f_{uu}(\lambda)$ on the estimator is accumulated across bands in the case of the pooled estimator $\widehat{d}_{pooled}$, as is apparent from the asymptotic formula (11). The variance of the two estimators are almost equal. MSE($\widehat{d}_{pooled}$) is similar to MSE($\widehat{d}_{GPH}$) when $a_1 = 0$, whereas the pooled estimator has a larger MSE than $\widehat{d}_{GPH}$ when $a_1 = \pm 0.5$. The difference in the MSE is smaller when a_1 is negative because the slope of $f_{uu}(\lambda)$ changes primarily where λ is far from the origin, and the effect of the shape of $f_{uu}(\lambda)$ far away from the origin is smaller. Thus, for AR(1) errors it appears that the GPH estimator is generally better than the pooled estimator in finite samples. The difference in the MSE is small when $n = 1000$, however, because the frequency band $[0, \lambda_m]$ becomes narrow relative to $[0, 2\pi]$.

Tables 6-8 show simulation results for the parameter combinations A, B, and C, for which the spectral density of u_t has a peak at the frequency $\lambda_p = \arccos\left\{-a_1 \frac{1-a_2}{4a_2}\right\}$. The values of λ_p are 1.98 ($\simeq 0.63\pi$), 1.16 ($\simeq 0.37\pi$), and 0.84 ($\simeq 0.27\pi/$) at A, B, and C, respectively. The simulation results for these parameter combinations are very different from the case of AR(1) errors. Now, the performance of the estimators depends very much on the location of the peak in the spectrum $f_{uu}(\lambda)$ relative to the frequency band being used in the regression. When the peak in the spectrum is near the frequency band $[\lambda_1, \lambda_m]$, as it is for B and C with $n = 200$ and 500, the estimator $\widehat{d}_{GPH}$ appears to be severely biased, and the pooled estimator $\widehat{d}_{pooled}$ has much smaller bias than $\widehat{d}_{GPH}$ (see Tables 7 and 8). On the other hand, at A, the peak in the spectrum of $f_{uu}(\lambda)$ is far from the origin, and $f_{uu}(\lambda)$ is close to constant around the origin. In this case (see Table 6), the bias of $\widehat{d}_{GPH}$ is relatively small, and $\widehat{d}_{pooled}$ has a larger bias. In all cases, although the variance of $\widehat{d}_{pooled}$ is smaller than that of $\widehat{d}_{GPH}$, the difference is not substantial. In terms of MSE, at parameter combinations B and C and for $n = 200$ and 500, MSE($\widehat{d}_{pooled}$) is decisively smaller than MSE($\widehat{d}_{GPH}$) due to the bias reduction. However, at A, $\widehat{d}_{pooled}$ has a larger MSE than $\widehat{d}_{GPH}$ because of its larger bias. When $n = 1000$, the difference between the two estimators becomes smaller.

In the simulations above, the pooled estimator uses a total $m(L+1) = 3m$ frequencies, while the GPH estimator uses m frequencies. We examine the effects of using the wider frequency band on the GPH estimator and see how this affects the comparison of the two estimators. The third row of Tables 3-8 reports results for $\widehat{d}_{GPH}(3m)$, the GPH estimator using $3m$ frequencies. Using $3m$ frequencies in GPH

[1] We report only the case $d = 0.3$, because the results for different memory parameter values are very similar.

leads to a dramatic increase in bias except for the parameter combinations H (where $f_{uu}(\lambda)$ is constant) and B with $m = 200$ and 500. The results in Tables 3, 5, 6, and 8 reveal that the GPH estimator based on the wider frequency band is very sensitive to the shape of the spectrum $f_{uu}(\lambda)$ around the origin and is generally much inferior to the pooled estimator. On the other hand, using only m frequencies, dividing the frequency band $[\lambda_1, \lambda_m]$ into bands and applying the pooled estimator does not provide a superior pooled estimator. The fourth row of Tables 3-8 shows results for $\widehat{d}_{pooled}(m)$, the pooled estimator that uses only m frequencies in total and two blocks, each block containing $[m/2]$ frequencies. Evidently, the increase in variance in this case more than offsets the reduction of bias.

In sum, the pooled estimator has advantages over the GPH estimator in finite samples, because the use of a wider frequency band ($m(L+1)$ rather than m) makes it less sensitive to the presence of peaks in the underlying spectral density $f_{uu}(\lambda)$. At the same time, it avoids the extremely large bias that is typical of the GPH estimator when a wide frequency band is employed. Therefore, it provides us with an alternate way of using a wider frequency band in log periodogram regression and a way to use more information, making the estimator more robust to various shapes in the short memory spectrum. In so doing, it can lead to substantial bias and MSE reductions when $f_{uu}(\lambda)$ has peaks that are close to the regression frequency band. On the other hand, it suffers from a mild bias increase when the error spectrum $f_{uu}(\lambda)$ changes monotonically, as it does in the case of AR(1) errors.

6.3 Empirical illustration

The methods were applied to US inflation series and stock returns. The inflation series constituted 624 observations of the monthly CPI inflation rate over the period 1947:1-1999:2; and the stock return series involved 3600 observations of the absolute value of returns on the daily S&P500 stock index from January 1979 to October 1992.[2] The first panel of Figure 6 graphs each series. The second panel of Figure 6 plots $\widehat{d}_{GPH}$ and $\widehat{d}_{pooled}$ for different values of m. (Specifically, $m = n^{0.5}, ..., n^{0.65}$ were used).

The value of the both estimator changes as m increases, although the pooled estimator appears to have a less sharp peak. The estimates of the memory parameter for inflation are in the region $(0.4, 0.8)$, indicating marginal nonstationarity of inflation. Those for stock return magnitudes are around 0.3, indicating stationary long range dependence.

The first panel of Figure 7 shows the residual fractionally differenced series $\widehat{u}_t = (1-L)^{\widehat{d}} X_t$, where $\widehat{d}$ is the pooled estimate with $m = n^{0.55}$. The spectral density estimates of $\widehat{u}_t$ are displayed in the second panel of Figure 7, using $\widehat{d}_{GPH}$ and $\widehat{d}_{pooled}$ estimates calculated with $m = n^{0.55}$.

[2] The inflation series were computed as $X_t = 100\Delta[\log(x_t)]$, where x_t is the US monthly CPI, over the period from January 1947 to February 1999. The stock return series were computed as $X_t = 100|\Delta(\log(x_t))|$, where x_t is the S&P 500 stock price index from January 1979 to October 1992.

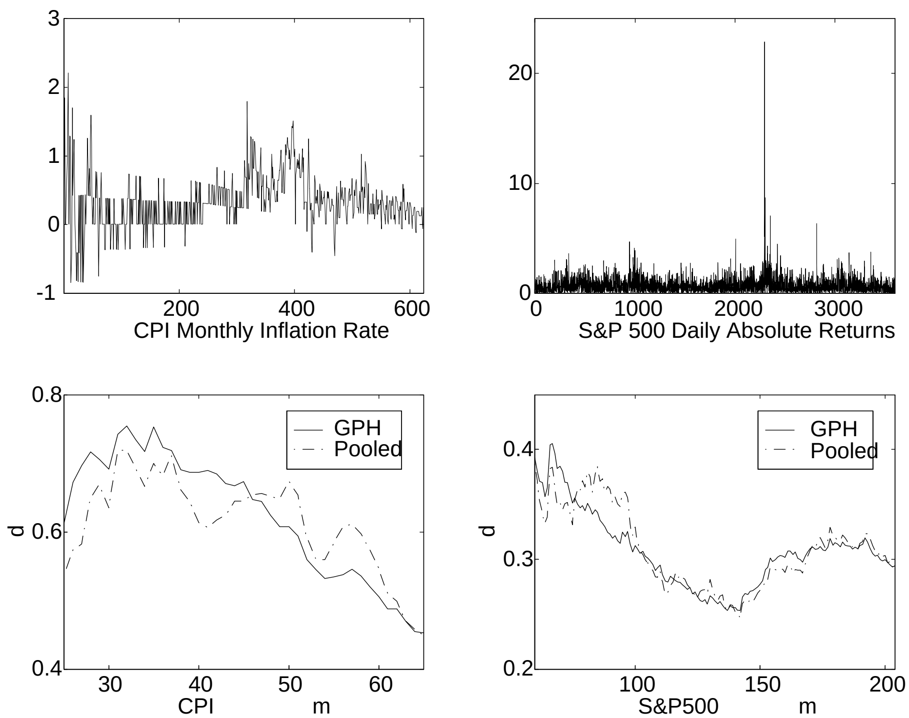

Figure 6: Inflation rate and stock return data and estimates of d

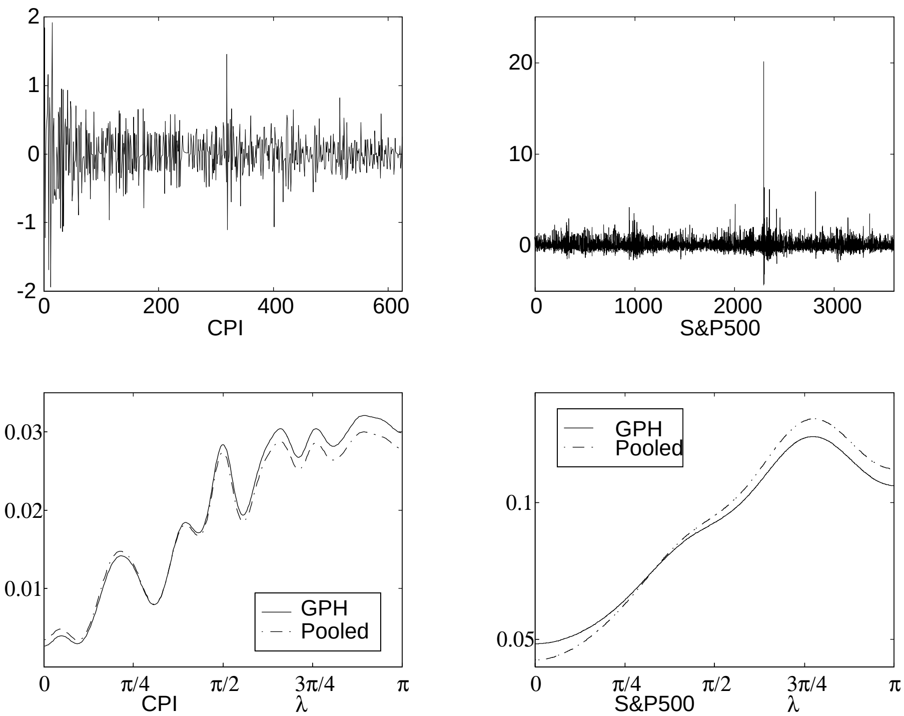

Figure 7: $(1-L)^{\widehat{d}} X_t$ and spectral density estimates

In both cases, the empirical estimates of the spectral density of u_t appear to have more power concentrated at higher frequencies. This is partly explained by the natural tendency of the GPH estimator to attribute power in the periodogram at lower frequencies to the long memory parameter d. The estimates of $f_{uu}(0)$ implied by the pooled estimator of d are higher than those obtained from the GPH estimator for the inflation series and lower for the stock return series. In other respects, the spectral densities estimates are very close.

7 Appendix: Proofs

7.1 Proof of Lemma 3.2

(a) $j = 0$

Assumptions 1 and 2 yield

$$\sum_{\{s:\lambda_s \in B_0\}} \left(X_{sj} - \overline{X}_{.j}\right)^2 = 4\sum_{s=1}^{m}\left(\ln\left|1-e^{i\lambda_s}\right| - \frac{1}{m}\sum_{s=1}^{m}\ln\left|1-e^{i\lambda_s}\right|\right)^2 = 4m + o\left(m\right),$$

as shown in lemma 1 of Hurvich and Beltrao (1994).

(b) $1 \leq j \leq \ell$

Taylor expansion gives, as in Hurvich and Beltrao (1994)

$$\left|1-e^{i\lambda_s}\right|^2 = 2\left(1-\cos\lambda_s\right) = \lambda_s^2\cos\xi_s, \quad 0<\xi_s<\lambda_s,$$

and

$$\begin{aligned} -\frac{1}{2}X_{sj} &= \ln\lambda_s + \frac{1}{2}\ln\cos\xi_s, \quad \lambda_s \in B_j \\ &= \ln\lambda_s - \frac{\xi_s^2}{4\cos^2\Omega_s}, \quad 0<\Omega_s<\xi_s \\ &= \ln\lambda_s + O\left(\xi_s^2\right) = \ln\lambda_s + O\left(\frac{j^2}{M^2}\right), \end{aligned}$$

for a sufficiently large n, because $\ell/M \to 0$ implies $\Omega_s \to 0$, and $\left(\cos^2\Omega_s\right)^{-1} = O(1)$ follows. Also, the order $O\left(j^2/M^2\right)$ is uniform in j.

Note that

$$\omega_j = \frac{(2j+1)\pi}{2M} = \frac{\pi m(2j+1)}{n}, \quad (\text{using } n = 2mM)$$

and it follows that

$$\begin{aligned} \lambda_s &\in B_j \Leftrightarrow -\frac{\pi}{2M} < \lambda_s - \omega_j \leq \frac{\pi}{2M} \\ &\Leftrightarrow -\frac{\pi m}{n} < \frac{2\pi s}{n} - \frac{(2j+1)\pi m}{n} \leq \frac{\pi m}{n} \\ &\Leftrightarrow -m < 2s - m(2j+1) \leq m \\ &\Leftrightarrow 0 < 2s - 2mj \leq 2m \\ &\Leftrightarrow 1 \leq s - mj \leq m. \end{aligned}$$

For $mj+1 \leq s, s' \leq mj+m$, by the mean value theorem

$$\begin{aligned} \ln s - \ln s' &= \frac{1}{\overline{s}}\left(s-s'\right) \quad \overline{s} \in \left[s, s'\right] \\ &= O(1/mj)\,O(m) = O(1/j), \end{aligned}$$

and

$$\ln s - \frac{1}{m}\sum_{s=mj+1}^{mj+m} \ln s = \frac{1}{m}\sum_{s'=mj+1}^{mj+m}\left(\ln s - \ln s'\right) = O(1/j).$$

It follows that

$$\begin{aligned} &\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj} - \overline{X}_{.j}\right)^2 \\ =\ & 4\sum_{\{s:\lambda_s\in B_j\}}\left[\ln\lambda_s - \frac{1}{m}\sum_{\{s:\lambda_s\in B_j\}}\ln\lambda_s + O\left(\frac{j^2}{M^2}\right)\right]^2 \end{aligned}$$

$$
\begin{aligned}
&= 4\sum_{\{s:\lambda_s\in B_j\}}\left[\ln s-\frac{1}{m}\sum_{\{s:\lambda_s\in B_j\}}\ln s+O\left(\frac{j^2}{M^2}\right)\right]^2\\
&= 4\sum_{s=mj+1}^{mj+m}\left(\ln s-\frac{1}{m}\sum_{s=mj+1}^{mj+m}\ln s\right)^2+\sum_{\{s:\lambda_s\in B_j\}}O\left(\frac{j}{M^2}\right)+O\left(\frac{mj^4}{M^4}\right)\\
&= 4\sum_{s/m=j+1/m}^{j+1}\left(\ln\frac{s}{m}-\frac{1}{m}\sum_{s/m=j+1/m}^{j+1}\ln\frac{s}{m}\right)^2+O\left(\frac{mj}{M^2}\right)+O\left(\frac{mj^4}{M^4}\right)\\
&\sim 4m\int_j^{j+1}(\ln x-c_j)^2\,dx+O\left(\frac{mj}{M^2}\right)+O\left(\frac{mj^4}{M^4}\right),
\end{aligned}
$$

where

$$
\begin{aligned}
c_j &= \int_j^{j+1}\ln x dx\\
&= [x(\ln x-1)]_j^{j+1}\\
&= (j+1)(\ln(j+1)-1)-j(\ln j-1)\\
&= (j+1)\ln(j+1)-j\ln j-1,\\
c_j+1 &= (j+1)\ln(j+1)-j\ln j,
\end{aligned}
$$

which gives

$$
\begin{aligned}
&\int_j^{j+1}(\ln x-c_j)^2\,dx\\
&= \left[x\left((\ln x-c_j-1)^2+1\right)\right]_j^{j+1}\\
&= (j+1)(\ln(j+1)-c_j-1)^2+j+1-j(\ln j-c_j-1)^2-j\\
&= (j+1)j^2(\ln(j+1)-\ln j)^2-j(j+1)^2(\ln(j+1)-\ln j)^2+1\\
&= -j(j+1)(\ln(j+1)-\ln j)^2+1.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\sum_{j=1}^{\ell}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{.j}\right)^2\\
&\sim 4m\sum_{j=1}^{\ell}\int_j^{j+1}(\ln x-c_j)^2\,dx+\sum_{j=1}^{\ell}\left(O\left(\frac{mj}{M^2}\right)+O\left(\frac{mj^4}{M^4}\right)\right)\\
&= 4m\sum_{j=1}^{\ell}\left[-j(j+1)(\ln(j+1)-\ln j)^2+1\right]+O\left(\frac{m\ell^2}{M^2}\right)+O\left(\frac{m\ell^5}{M^4}\right)\\
&= 4m\sum_{j=1}^{\ell}\left[-j(j+1)(\ln(j+1)-\ln j)^2+1\right]+o(m).
\end{aligned}
$$

There is no explicit numerical value for the quantity $\sum_{j=1}^{\ell}\left(-j\,(j+1)\,(\ln(j+1)-\ln j)^2+1\right)$ or its limit as $\ell\to\infty$. Nevertheless, the sum converges since

$$\begin{aligned}&\int_j^{j+1}(\ln x-c_j)^2\,dx\\ \leq\ &(\ln(j+1)-\ln j)^2\ \text{ (by the mean value theorem)}\\ \leq\ &1/j^2,\end{aligned}$$

which implies that $\sum_{j=1}^{\infty}\int_j^{j+1}(\ln x-c_j)^2\,dx<\infty$. Direct calculations using Mathematica produce the approximate numerical value $\sum_{j=1}^{\infty}\left(-j\,(j+1)\,(\ln(j+1)-\ln j)^2+1\right)\doteqdot 0.0803$.

7.2 Proof of Lemma 3.3

For $j=0$, by lemma 1 of Hurvich, Deo, and Brodsky (1998) (hereafter HDB), we have

$$\begin{aligned}\sum_{\{s:\lambda_s\in B_0\}}\eta_{sj}\left(X_{sj}-\overline{X}_{.j}\right)&=\sum_{\{s:\lambda_s\in B_0\}}(\ln f_{uu}(\lambda_s)-\ln f_{uu}(0))\left(X_{sj}-\overline{X}_{.j}\right)\\ =\sum_{\{s:\lambda_s\in B_0\}}\ln f_{uu}(\lambda_s)\left(X_{sj}-\overline{X}_{.j}\right)&=-\frac{8\pi^2}{9}\frac{f''_{uu}(0)}{f_{uu}(0)}\frac{m^3}{n^2}+o\left(\frac{m^3}{n^2}\right).\end{aligned}$$

For $j\geq 1$, we first collect together some useful technicalities in the following Sub-lemma.

7.2.1 Sub-lemma

1. $\frac{f'_{uu}(\omega_j)}{f_{uu}(\omega_j)}=\frac{f''_{uu}(0)\omega_j+O\left(\omega_j^2\right)}{f_{uu}(0)+O(\omega_j)}=\frac{f''_{uu}(0)}{f_{uu}(0)}\omega_j\,(1+O(\omega_j))=O(\omega_j)$ *uniformly in* j.
2. $O\left(\frac{M}{nj}\right)=O\left(\frac{1}{j^2}\right)$ *because* $j=o(m)$.
3. $\frac{\sin\omega_j}{1-\cos\omega_j}=\frac{\omega_j+O\left(\omega_j^2\right)}{0.5\omega_j^2+O\left(\omega_j^4\right)}=\frac{2}{\omega_j}\,(1+O(\omega_j))$ *uniformly in* j.

■

First, note that

$$\begin{aligned}\sum_{\{s:\lambda_s\in B_j\}}(\lambda_s-\omega_j)&=\sum_{s=mj+1}^{mj+m}\left(\frac{2\pi s}{n}-\frac{(2j+1)\,\pi m}{n}\right)=\frac{2\pi}{n}\sum_{s=mj+1}^{mj+m}(s-(j+1/2)\,m)\\ &=\frac{2\pi}{n}\sum_{s=mj+1}^{mj+m}(s-mj-m/2)=\frac{2\pi}{n}\sum_{k=1}^{m}(k-m/2)=\frac{2\pi}{n}\left(\frac{m\,(m+1)}{2}-\frac{m^2}{2}\right)\\ &=\frac{2\pi}{n}\frac{m}{2},\end{aligned}$$

because positive and negative terms cancel out as we sum over $s - jm$ from 1 to m. It follows that

$$\begin{aligned}
-\overline{X}_{.j} &= -\frac{1}{m}\sum_{\{s:\lambda_s\in B_j\}} X_{sj} = \frac{1}{m}\sum_{\{s:\lambda_s\in B_j\}} \ln\left(2-2\cos\lambda_s\right) \\
&= \frac{1}{m}\sum_{\{s:\lambda_s\in B_j\}}\left[\ln\left(2-2\cos\omega_j\right)+\frac{\sin\omega_j}{1-\cos\omega_j}\left(\lambda_s-\omega_j\right)+\frac{1}{2}\frac{1}{\cos\widetilde{\lambda}_s-1}\left(\lambda_s-\omega_j\right)^2\right] \\
&= \ln\left(2-2\cos\omega_j\right)+\frac{\sin\omega_j}{1-\cos\omega_j}\frac{1}{m}\sum_{\{s:\lambda_s\in B_j\}}\left(\lambda_s-\omega_j\right)+O\left(\frac{1}{\omega_j^2 M^2}\right) \\
&= \ln\left(2-2\cos\omega_j\right)+\frac{\sin\omega_j}{1-\cos\omega_j}\frac{\pi}{n}+O\left(\frac{1}{j^2}\right),
\end{aligned}$$

where $\widetilde{\lambda}_s \in (\omega_j, \lambda_s)$, and the term $O\left(1/j^2\right)$ is uniform in j.

By Taylor expansion

$$\begin{aligned}
\ln f_{uu}\left(\lambda_s\right) &= \ln f_{uu}\left(\omega_j\right)+\frac{f'_{uu}\left(\omega_j\right)}{f_{uu}\left(\omega_j\right)}\left(\lambda_s-\omega_j\right)+\left(\frac{f''_{uu}\left(\widetilde{\lambda}_s\right)}{f_{uu}\left(\widetilde{\lambda}_s\right)}-\left(\frac{f'_{uu}\left(\widetilde{\lambda}_s\right)}{f_{uu}\left(\widetilde{\lambda}_s\right)}\right)^2\right)\frac{\left(\lambda_s-\omega_j\right)^2}{2} \\
&= \ln f_{uu}\left(\omega_j\right)+\frac{f'_{uu}\left(\omega_j\right)}{f_{uu}\left(\omega_j\right)}\left(\lambda_s-\omega_j\right)+O\left(\frac{1}{M^2}\right).
\end{aligned}$$

Thus, in view of Sub-lemma 7.2.1,

$$\begin{aligned}
&\sum_{\{s:\lambda_s\in B_j;j\geq 1\}} \eta_{sj}\left(X_{sj}-\overline{X}_{.j}\right) \\
&= \sum_{\{s:\lambda_s\in B_j;j\geq 1\}}\left[\frac{f'_{uu}\left(\omega_j\right)}{f_{uu}\left(\omega_j\right)}\left(\lambda_s-\omega_j\right)+O\left(\frac{1}{M^2}\right)\right]\left[-\left(\frac{\sin\omega_j}{1-\cos\omega_j}\right)\left(\lambda_s-\omega_j-\frac{\pi}{n}\right)+O\left(\frac{1}{j^2}\right)\right] \\
&= \sum_{\{s:\lambda_s\in B_j;j\geq 1\}}\left[\frac{f'_{uu}\left(\omega_j\right)}{f_{uu}\left(\omega_j\right)}\left(\lambda_s-\omega_j\right)+O\left(\frac{1}{M^2}\right)\right] \\
&\quad\times\left[-\left(\frac{\sin\omega_j}{1-\cos\omega_j}\right)\left(\lambda_s-\omega_j\right)+O\left(\frac{M}{nj}\right)+O\left(\frac{1}{j^2}\right)\right] \\
&= -\left(\frac{f'_{uu}\left(\omega_j\right)}{f_{uu}\left(\omega_j\right)}\right)\left(\frac{\sin\omega_j}{1-\cos\omega_j}\right)\sum_{\{s:\lambda_s\in B_j;j\geq 1\}}\left(\lambda_s-\omega_j\right)^2 \\
&\quad+\left(\frac{f'_{uu}\left(\omega_j\right)}{f_{uu}\left(\omega_j\right)}\right)\sum_{\{s:\lambda_s\in B_j;j\geq 1\}}\left(\lambda_s-\omega_j\right)\cdot O\left(\frac{1}{j^2}\right) \\
&\quad-\left(\frac{\sin\omega_j}{1-\cos\omega_j}\right)\sum_{\{s:\lambda_s\in B_j;j\geq 1\}}\left(\lambda_s-\omega_j\right)\cdot O\left(\frac{1}{M^2}\right)+O\left(\frac{m}{M^2 j^2}\right) \\
&= -\left(\frac{f'_{uu}\left(\omega_j\right)}{f_{uu}\left(\omega_j\right)}\right)\left(\frac{\sin\omega_j}{1-\cos\omega_j}\right)\left(\frac{2\pi}{n}\right)^2\sum_{k=1}^{m}\left(k-\frac{m}{2}\right)^2
\end{aligned}$$

$$
\begin{aligned}
&+O\left(\frac{j}{M}\right)\sum_{\{s:\lambda_s\in B_j\}} O\left(\frac{1}{Mj^2}\right)-O\left(\frac{M}{j}\right)\sum_{\{s:\lambda_s\in B_j\}} O\left(\frac{1}{M^3}\right)+O\left(\frac{m}{M^2j^2}\right)\\
= & -\left(\frac{f'_{uu}(\omega_j)}{f_{uu}(\omega_j)}\right)\left(\frac{\sin\omega_j}{1-\cos\omega_j}\right)\frac{\pi^2}{m^2M^2}\left(\frac{m^3}{12}+O\left(m^2\right)\right)+O\left(\frac{m}{M^2j}\right)+O\left(\frac{m}{M^2j^2}\right)\\
= & -\left(\frac{f'_{uu}(\omega_j)}{f_{uu}(\omega_j)}\right)\left(\frac{\sin\omega_j}{1-\cos\omega_j}\right)\frac{\pi^2}{M^2}\frac{m}{12}+O\left(\frac{1}{M^2}\right)+O\left(\frac{m}{M^2j}\right)\\
= & -\left(\frac{f'_{uu}(\omega_j)}{f_{uu}(\omega_j)}\right)\left(\frac{\sin\omega_j}{1-\cos\omega_j}\right)\frac{\pi^2}{M^2}\frac{m}{12}+O\left(\frac{m}{M^2j}\right),
\end{aligned}
$$

where we use the fact that

$$
\sum_{k=1}^{m}\left(k-\frac{m}{2}\right)^2=\frac{m(m+1)(2m+1)}{6}-\frac{m^2(m+1)}{2}+\frac{m^3}{4}=\frac{m^3}{12}+O\left(m^2\right).
$$

Furthermore, from Sub-lemma 7.2.1 we have

$$
\left(\frac{f'_{uu}(\omega_j)}{f_{uu}(\omega_j)}\right)\left(\frac{\sin\omega_j}{1-\cos\omega_j}\right)=\frac{2f''_{uu}(0)}{f_{uu}(0)}\left(1+O\left(\omega_j\right)\right).
$$

Therefore,

$$
\begin{aligned}
&\sum_{j=1}^{\ell}\sum_{\{s:\lambda_s\in B_j\}}\eta_{sj}\left(X_{sj}-\overline{X}_{.j}\right)\\
= & -\frac{2f''_{uu}(0)}{f_{uu}(0)}\frac{\pi^2}{M^2}\frac{m}{12}+O\left(\sum_{j=1}^{\ell}\frac{j}{M}\frac{m}{M^2}\right)+O\left(\sum_{j=1}^{\ell}\frac{m}{M^2j}\right)\\
= & -\frac{2\pi^2}{3}\frac{f''_{uu}(0)}{f_{uu}(0)}\frac{m^3\ell}{n^2}+O\left(\sum_{j=1}^{\ell}\frac{m^4j}{n^3}\right)+O\left(\sum_{j=1}^{\ell}\frac{m^3}{n^2j}\right)\\
= & -\frac{2\pi^2}{3}\frac{f''_{uu}(0)}{f_{uu}(0)}\frac{m^3\ell}{n^2}+O\left(\frac{m^4\ell^2}{n^3}\right)+O\left(\frac{m^3}{n^2}\ln\ell\right)\\
= & -\frac{2\pi^2}{3}\frac{f''_{uu}(0)}{f_{uu}(0)}\frac{m^3\ell}{n^2}+o\left(\frac{m^3\ell}{n^2}\right),
\end{aligned}
$$

giving the required result. ■

7.3 Proof of Theorem 4.1

Again, it is helpful to start by collecting some useful preliminary results in the following Sub-lemmas.

7.3.1 Sub-Lemma

1. $\int_{-\pi}^{\pi}|D(\lambda)|\,d\lambda=O(\ln n)$

2. $K(\lambda) = O\left(\left(n\lambda^2\right)^{-1}\right), \quad 0 < |\lambda| \leq \pi$

3. $|D(\lambda)| \leq 2|\lambda|^{-1}, \quad 0 < |\lambda| \leq \pi.$

where $K(\lambda) = (2\pi n)^{-1}\left|\sum_{t=1}^{n}\sum_{s=1}^{n} e^{i(t-s)\lambda}\right| = (2\pi n)^{-1}|D(\lambda)|^2, \quad D(\lambda) = \sum_{t=1}^{n} e^{it\lambda}.$

Proof See Zygmund (1959) (p. 67 for 1, pp.89-90 for 2, pp.49-51 for 3). ■

7.3.2 Sub-Lemma

For some $C < \infty$ and for $\lambda \in [-\pi, 0) \cup (0, \pi]$,

1. $|f_{XX}(\lambda)| \leq C|\lambda|^{-2d}$
2. $|f_{XX}(\lambda)|^{-1} \leq C|\lambda|^{2d}$
3. $|f'_{XX}(\lambda)| \leq C|\lambda|^{-2d-1}$.

Proof First, the inequality $|x/2| \leq |\sin x| \leq |x|$ for $x \in [-\pi/2, \pi/2]$ implies that, for both positive and negative d, $\left|\sin^2(\lambda/2)\right|^{-d} \leq C_1|\lambda|^{-2d}$ holds for $\lambda \in [-\pi, 0) \cup (0, \pi]$ and for some $C_1 < \infty$. It follows that

$$\begin{aligned}
|f_{XX}(\lambda)| &= \left|4\sin^2(\lambda/2)\right|^{-d}|f_{uu}(\lambda)| \leq 4^{-d}C_1|\lambda|^{-2d}\sup_{\lambda\in(-\pi,\pi)}|f_{uu}(\lambda)| \leq C|\lambda|^{-2d}, \\
|f_{XX}(\lambda)|^{-1} &= \left|4\sin^2(\lambda/2)\right|^{d}|f_{uu}(\lambda)|^{-1} \leq 4^{d}C_1|\lambda|^{2d}\sup_{\lambda\in(-\pi,\pi)}|f_{uu}(\lambda)|^{-1} \leq C|\lambda|^{2d},
\end{aligned}$$

because $\sup_{\lambda\in(-\pi,\pi)}|f_{uu}(\lambda)| < \infty$ and $\inf_{\lambda\in(-\pi,\pi)}|f_{uu}(\lambda)| > 0$. The bound for $|f'_{XX}(\lambda)|$ uses the fact that $\sup_{\lambda\in(-\pi,\pi)}|f'_{uu}(\lambda)| < \infty$ and then

$$\begin{aligned}
&\left|f'_{XX}(\lambda)\right| \\
\leq\ &\left|4^{-d}(-d)\left(\sin^2(\lambda/2)\right)^{-d-1}\sin(\lambda/2)\cos(\lambda/2)\, f_{uu}(\lambda)\right| + \left|\left(4\sin^2(\lambda/2)\right)^{-d} f'_{uu}(\lambda)\right| \\
\leq\ &C_2|\sin(\lambda/2)|^{-2d-1} + C_3|\lambda|^{-2d} \leq C|\lambda|^{-2d-1}. \blacksquare
\end{aligned}$$

7.3.3 Sub-Lemma

For $0 < \lambda + \lambda_j < 2\pi$ and $0 \leq \lambda, \lambda_j \leq \pi$,

$$|\lambda - \lambda_j|\,|D(\lambda + \lambda_j)| \leq 2, \tag{15}$$

Proof Because $|D(\lambda)| = |D(2\pi - \lambda)|$ for $\pi \le \lambda \le 2\pi$, and in view of Sub lemma 7.3.1 (3),

$$\begin{aligned} & |\lambda - \lambda_j| \, |D(\lambda + \lambda_j)| \\ \le \; & 2\frac{|\lambda - \lambda_j|}{\lambda + \lambda_j}, \quad 0 < \lambda + \lambda_j \le \pi \\ \le \; & 2\frac{|\lambda - \lambda_j|}{2\pi - \lambda - \lambda_j}, \quad \pi < \lambda + \lambda_j < 2\pi, \end{aligned}$$

and

$$\begin{aligned} \frac{|\lambda - \lambda_j|}{\lambda + \lambda_j} &= \frac{(\lambda - \lambda_j)}{2\lambda_j + (\lambda - \lambda_j)} \le 1, \quad \lambda \ge \lambda_j \\ &= \frac{(\lambda_j - \lambda)}{2\lambda + (\lambda_j - \lambda)} \le 1, \quad \lambda < \lambda_j \\ \frac{|\lambda - \lambda_j|}{2\pi - \lambda - \lambda_j} &= \frac{(\lambda - \lambda_j)}{2\pi - 2\lambda + (\lambda - \lambda_j)} \le 1, \quad \lambda \ge \lambda_j \\ &= \frac{(\lambda_j - \lambda)}{2\pi - 2\lambda_j + (\lambda_j - \lambda)} \le 1. \quad \lambda < \lambda_j \; \blacksquare \end{aligned}$$

7.3.4 Sub-Lemma

If $\kappa_1 \lambda_j \le \lambda_j + \lambda \le 2\pi - \kappa_2 \lambda_j$ for $\kappa_1, \kappa_2 > 0$, $\sup \lambda_j |D(\lambda_j + \lambda)| < \infty$.

If $\kappa_1 \lambda_j \le \lambda_j - \lambda \le 2\pi - \kappa_2 \lambda_j$ for $\kappa_1, \kappa_2 > 0$, $\sup \lambda_j |D(\lambda_j - \lambda)| < \infty$.

Proof

$$\begin{aligned} \lambda_j |D(\lambda_j + \lambda)| \le \frac{2\lambda_j}{\lambda_j + \lambda} &\le \frac{2}{\kappa_1}, \quad \kappa_1 \lambda_j \le \lambda_j + \lambda \le \pi \\ \lambda_j |D(\lambda_j + \lambda)| \le \frac{2\lambda_j}{2\pi - \lambda_j - \lambda} &\le \frac{2}{\kappa_2}, \quad \pi < \lambda_j + \lambda \le 2\pi - \kappa_2 \lambda_j \end{aligned}$$

similarly for $\lambda_j |D(\lambda_j - \lambda)|$. $\blacksquare$

With these technicalities in hand, the proof of theorem 4.1 is almost identical to that in Robinson (1995). The main difference is that $f_{XX}(\lambda)$ is bounded by $|\lambda|^{-2d}$ over the full range of λ, and the evaluation of $|D(\lambda)|$ becomes complicated because λ_j is not restricted to a neighborhood of the origin. We proceed with each part in turn.

7.3.5 Proof of (a)

We start by showing

$$E[w(\lambda_j)\overline{w}(\lambda_j)] - f_{XX}(\lambda_j) = \int_{-\pi}^{\pi} \{f_{XX}(\lambda) - f_{XX}(\lambda_j)\} K(\lambda - \lambda_j) \, d\lambda = O\left(j^{-1}\lambda_j^{-2d} \ln n\right).$$

where $K(\lambda) = (2\pi n)^{-1} \left| \sum_{t=1}^{n} \sum_{s=1}^{n} e^{i(t-s)\lambda} \right|$ is Fejer's kernel. By sub-lemmas 7.3.1 and 7.3.2,

$$\begin{aligned}
\left| \int_{\lambda_j/2}^{\pi} \right| &\leq \left\{ \max_{\lambda_j/2 \leq \lambda \leq \pi} \left| f'_{XX}(\lambda) \right| \right\} \int_{\lambda_j/2}^{\pi} |\lambda - \lambda_j| \, K(\lambda - \lambda_j) \, d\lambda \\
&= O\left(\lambda_j^{-1-2d} n^{-1} \int_{\lambda_j/2}^{\pi} |D(\lambda - \lambda_j)| \, d\lambda \right), \quad \text{because } \lambda - \lambda_j \in (-\lambda_j/2, \pi - \lambda_j) \subseteq [-\pi, \pi] \\
&= O\left(\lambda_j^{-1-2d} n^{-1} \int_{-\pi}^{\pi} |D(\lambda)| \, d\lambda \right) \\
&= O\left(\lambda_j^{-1-2d} n^{-1} \ln n \right) = O\left(j^{-1} \lambda_j^{-2d} \ln n \right).
\end{aligned}$$

The symmetry of $f_{XX}(\lambda)$ and sub-lemma 7.3.3 (applicable since $0 < \lambda + \lambda_j < 2\pi$) yields

$$\begin{aligned}
\left| \int_{-\pi}^{-\lambda_j/2} \right| &= \left| \int_{\lambda_j/2}^{\pi} \{ f_{XX}(\lambda) - f_{XX}(\lambda_j) \} \, K(\lambda + \lambda_j) \, d\lambda \right| \\
&\leq \left\{ \max_{\lambda_j/2 \leq \lambda \leq \pi} \left| f'_{XX}(\lambda) \right| \right\} \int_{\lambda_j/2}^{\pi} |\lambda - \lambda_j| \, K(\lambda + \lambda_j) \, d\lambda \\
&= O\left(\left\{ \max_{\lambda_j/2 \leq \lambda \leq \pi} \left| f'_{XX}(\lambda) \right| \right\} n^{-1} \int_{\lambda_j/2}^{\pi} |\lambda - \lambda_j| \, |D(\lambda + \lambda_j)|^2 \, d\lambda \right) \\
&= O\left(\left\{ \max_{\lambda_j/2 \leq \lambda \leq \pi} \left| f'_{XX}(\lambda) \right| \right\} n^{-1} \int_{\lambda_j/2}^{\pi} |D(\lambda + \lambda_j)| \, d\lambda \right) \\
&= O\left(\lambda_j^{-1-2d} n^{-1} \ln n \right) = O\left(j^{-1} \lambda_j^{-2d} \ln n \right).
\end{aligned}$$

For the remaining part, as in Robinson (1995),

$$\begin{aligned}
\left| \int_{-\lambda_j/2}^{\lambda_j/2} \right| &\leq \max_{|\lambda| \leq \lambda_j/2} K(\lambda - \lambda_j) \int_{-\lambda_j/2}^{\lambda_j/2} \{ |f_{XX}(\lambda)| + |f_{XX}(\lambda_j)| \} \, d\lambda \qquad (16) \\
&= O\left(n^{-1} \lambda_j^{-1-2d} \right) = O\left(j^{-1} \lambda_j^{-2d} \right).
\end{aligned}$$

Finally, by sub-lemma 7.3.2,

$$E\left[w(\lambda_j) \overline{w}(\lambda_j) / f_{XX}(\lambda_j) \right] = 1 + O\left(j^{-1} \ln n \lambda_j^{-2d} / f_{XX}(\lambda_j) \right) = 1 + O\left(j^{-1} \ln n \right).$$

7.3.6 Proof of (b)

By Robinson (1995), we have

$$E\left[w(\lambda_j) \, w(\lambda_j) \right] = \int_{-\pi}^{\pi} \{ f_{XX}(\lambda) - f_{XX}(\lambda_j) \} \, E_{j,-j}(\lambda) \, d\lambda, \qquad (17)$$

for $0 < k < j < n/2$, where $E_{jk}(\lambda) = (2\pi n)^{-1} D(\lambda_j - \lambda) D(\lambda - \lambda_k)$. The calculations are similar to those given before, but we have to check the range of integration for each subinterval.

$$\begin{aligned}
\left| \int_{\lambda_j/2}^{\pi} \right| &= \left| \int_{\lambda_j/2}^{\pi} (2\pi n)^{-1} \{ f_{XX}(\lambda) - f_{XX}(\lambda_j) \} D(\lambda_j - \lambda) D(\lambda + \lambda_j) \, d\lambda \right| \\
&\leq \left\{ \max_{\lambda_j/2 \leq \lambda \leq \pi} \left| f'_{XX}(\lambda) \right| \right\} (2\pi n)^{-1} \int_{\lambda_j/2}^{\pi} |\lambda_j - \lambda| \, |D(\lambda_j - \lambda)| \, |D(\lambda + \lambda_j)| \, d\lambda \\
&= O\left(\lambda_j^{-1-2d} n^{-1} \int_{\lambda_j/2}^{\pi} |D(\lambda + \lambda_j)| \, d\lambda \right) \\
&= O\left(\lambda_j^{-1-2d} n^{-1} \ln n \right) = O\left(j^{-1} \lambda_j^{-2d} \ln n \right),
\end{aligned}$$

because $-\pi < -\pi + \lambda_j \leq \lambda_j - \lambda \leq \lambda_j/2 < \pi$. Similarly,

$$\begin{aligned}
& \left| \int_{-\pi}^{-\lambda_j/2} \{ f_{XX}(\lambda) - f_{XX}(\lambda_j) \} (2\pi n)^{-1} D(\lambda_j - \lambda) D(\lambda + \lambda_j) \, d\lambda \right| \\
&= \left| \int_{\lambda_j/2}^{\pi} \{ f_{XX}(\lambda) - f_{XX}(\lambda_j) \} (2\pi n)^{-1} D(\lambda_j - \lambda) D(\lambda + \lambda_j) \, d\lambda \right| \\
&= O\left(j^{-1} \lambda_j^{-2d} \ln n \right).
\end{aligned}$$

For the remaining part, $\lambda j/2 \leq \lambda_j - \lambda \leq 3\lambda_j/2 \leq 2\pi - \lambda_j/2$ and sub-lemma 7.3.4 yield

$$\begin{aligned}
& \int_{-\lambda_j/2}^{-\lambda_j/2} \{ f_{XX}(\lambda) - f_{XX}(\lambda_j) \} E_{j,-j}(\lambda) \, d\lambda \\
&= O\left(\max_{|\lambda| \leq \lambda_j/2} |E_{j,-j}(\lambda)| \int_0^{\lambda_j} \{ |f_{XX}(\lambda)| + |f_{XX}(\lambda_j)| \} \, d\lambda \right) \\
&= O\left(n^{-1} \lambda_j^{-2} \lambda_j^{1-2d} \right) = O\left(j^{-1} \lambda_j^{-2d} \right).
\end{aligned}$$

7.3.7 Proof of (c)

Similar to Robinson (1995), we expand the integral as follows, for $0 < k < j < n/2$,

$$\begin{aligned}
E[w(\lambda_j) \overline{w}(\lambda_k)] &= \int_{-\pi}^{\pi} f_{XX}(\lambda) E_{jk}(\lambda) \, d\lambda \\
&= \int_{(\lambda_j + \lambda_k)/2}^{\pi} \{ f_{XX}(\lambda) - f_{XX}(\lambda_j) \} E_{jk}(\lambda) \, d\lambda \qquad (18) \\
&\quad + \int_{\lambda_k/2}^{(\lambda_j + \lambda_k)/2} \{ f_{XX}(\lambda) - f_{XX}(\lambda_k) \} E_{jk}(\lambda) \, d\lambda \qquad (19) \\
&\quad - \{ f_{XX}(\lambda_j) - f_{XX}(\lambda_k) \} \int_{\lambda_k/2}^{(\lambda_j + \lambda_k)/2} E_{jk}(\lambda) \, d\lambda \qquad (20)
\end{aligned}$$

$$+\int_{-\pi}^{\lambda_k/2}\{f_{XX}(\lambda)-f_{XX}(\lambda_j)\}E_{jk}(\lambda)\,d\lambda. \qquad (21)$$

First, because $-\pi<\lambda_j-\pi\le\lambda_j-\lambda\le(\lambda_j-\lambda_k)/2<\pi$, (18) is bounded by

$$\begin{aligned}
&\left\{\max_{(\lambda_j+\lambda_k)/2\le\lambda\le\pi}\left|f'_{XX}(\lambda)\right|\right\}n^{-1}\int_{(\lambda_j+\lambda_k)/2}^{\pi}|\lambda_j-\lambda|\,|D(\lambda_j-\lambda)|\,|D(\lambda-\lambda_k)|\,d\lambda \\
=\;&\left\{\max_{(\lambda_j+\lambda_k)/2\le\lambda\le\pi}\left|f'_{XX}(\lambda)\right|\right\}n^{-1}\int_{(\lambda_j+\lambda_k)/2}^{\pi}|D(\lambda-\lambda_k)|\,d\lambda \qquad (22)\\
=\;&O\left(\lambda_j^{-1-2d}n^{-1}\ln n\right)\\
=\;&O\left(j^{-1}\lambda_j^{-2d}\ln n\right)\\
=\;&O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\left(\frac{k}{j}\right)^{1+d}\right)\\
=\;&O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\right).
\end{aligned}$$

Next, when $k\ge j/2$, (19) is bounded by,

$$\begin{aligned}
&\left\{\max_{\lambda_k/2\le\lambda\le(\lambda_j+\lambda_k)/2}\left|f'_{XX}(\lambda)\right|\right\}n^{-1}\int_{\lambda_k/2}^{(\lambda_j+\lambda_k)/2}|\lambda-\lambda_k|\,|D(\lambda_j-\lambda)|\,|D(\lambda-\lambda_k)|\,d\lambda \\
=\;&\left\{\max_{\lambda_k/2\le\lambda\le(\lambda_j+\lambda_k)/2}\left|f'_{XX}(\lambda)\right|\right\}n^{-1}\int_{\lambda_k/2}^{(\lambda_j+\lambda_k)/2}|D(\lambda_j-\lambda)|\,d\lambda \\
=\;&O\left(\lambda_k^{-1-2d}n^{-1}\ln n\right)\\
=\;&O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\left(\frac{k}{j}\right)^{-d}\right)\\
=\;&O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\right),
\end{aligned}$$

because $-\pi<-\lambda_k/2\le\lambda-\lambda_k\le(\lambda_j-\lambda_k)/2<\pi$, and, when $k<j/2$, by

$$\begin{aligned}
&\left\{\max_{\lambda_k/2\le\lambda\le\lambda_j}|f_{XX}(\lambda)|+|f_{XX}(\lambda_k)|\right\}\int_{\lambda_k/2}^{(\lambda_j+\lambda_k)/2}|E_{jk}(\lambda)|\,d\lambda\\
=\;&O\left(\left(\lambda_j^{-2d}+\lambda_k^{-2d}\right)(j-k)^{-1}\int_{\lambda_k/2}^{(\lambda_j+\lambda_k)/2}|D(\lambda-\lambda_k)|\,d\lambda\right)\\
=\;&O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\left(\frac{k}{j-k}\right)\left\{\left(\frac{\lambda_j}{\lambda_k}\right)^{-d}+\left(\frac{\lambda_k}{\lambda_j}\right)^{-d}\right\}\right)\\
=\;&O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\left(\frac{j}{j-k}\right)\left\{\left(\frac{k}{j}\right)^{1+d}+\left(\frac{k}{j}\right)^{1-d}\right\}\right)\\
=\;&O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\right),
\end{aligned}$$

because $0 < (\lambda_j - \lambda_k)/2 \le \lambda_j - \lambda \le \lambda_j < \pi$.

Similarly, (20) is bounded by

$$\begin{aligned}
& (\lambda_j - \lambda_k) \left\{ \max_{\lambda_k \le \lambda \le \lambda_j} \left| f'_{XX}(\lambda) \right| \right\} \int_{\lambda_k/2}^{(\lambda_j+\lambda_k)/2} \left| E_{jk}(\lambda) \right| d\lambda \\
= \;& O\left(\frac{j-k}{n} \lambda_k^{-1-2d} (j-k)^{-1} \int_{\lambda_k/2}^{(\lambda_j+\lambda_k)/2} \left| D(\lambda - \lambda_k) \right| d\lambda \right) \\
= \;& O\left(\lambda_k^{-1-2d} n^{-1} \ln n \right) = O\left(k^{-1} \lambda_j^{-d} \lambda_k^{-d} \ln n \right).
\end{aligned}$$

when $k \ge j/2$, and by

$$\begin{aligned}
& \left\{ \left| f_{XX}(\lambda_j) \right| + \left| f_{XX}(\lambda_k) \right| \right\} (j-k)^{-1} \int_{\lambda_k/2}^{(\lambda_j+\lambda_k)/2} \left| D(\lambda - \lambda_k) \right| d\lambda \\
= \;& O\left(\left(\lambda_j^{-2d} + \lambda_k^{-2d} \right) (j-k)^{-1} \int_{-\pi}^{\pi} \left| D(\lambda) \right| d\lambda \right) \\
= \;& O\left(k^{-1} \lambda_j^{-d} \lambda_k^{-d} \ln n \right).
\end{aligned}$$

when $k < j/2$.

Finally, (21) is calculated by dividing the interval into $[-\pi, -\lambda_j/2]$, $[-\lambda_j/2, -\lambda_k/2]$,and $[-\lambda_k/2, \lambda_k/2]$. First, for the integral on $[-\pi, -\lambda_j/2]$, we use $|D(-\lambda)| = |D(\lambda)|$ and sub-lemma 7.3.3 to derive

$$\begin{aligned}
& \int_{-\pi}^{-\lambda_j/2} (2\pi n)^{-1} \left\{ f_{XX}(\lambda) - f_{XX}(\lambda_j) \right\} D(\lambda_j - \lambda) D(\lambda - \lambda_k) d\lambda \\
\le \;& \int_{\lambda_j/2}^{\pi} (2\pi n)^{-1} \left\{ f_{XX}(\lambda) - f_{XX}(\lambda_j) \right\} \left| D(\lambda_j + \lambda) \right| \left| D(\lambda + \lambda_k) \right| d\lambda \\
= \;& O\left(\left\{ \max_{\lambda_j/2 \le \lambda \le \pi} \left| f'_{XX}(\lambda) \right| \right\} \int_{\lambda_j/2}^{\pi} n^{-1} \left| \lambda - \lambda_j \right| \left| D(\lambda_j + \lambda) \right| \left| D(\lambda + \lambda_k) \right| d\lambda \right) \\
= \;& O\left(\lambda_j^{-1-2d} n^{-1} \int_{\lambda_j/2}^{\pi} \left| D(\lambda + \lambda_k) \right| d\lambda \right) \\
= \;& O\left(\lambda_j^{-1-2d} n^{-1} \ln n \right) = O\left(k^{-1} \lambda_j^{-d} \lambda_k^{-d} \ln n \right).
\end{aligned}$$

For the integral on $[-\lambda_k/2, \lambda_k/2]$, note that both $\lambda_j \left| D(\lambda_j - \lambda) \right|$ and $\lambda_k \left| D(\lambda_k - \lambda) \right|$ are bounded (see sub-Lemma 7.3.4) because $\lambda_j/2 < \lambda_j - \lambda < 2\pi - \lambda_j/2$ and $\lambda_k/2 \le \lambda_k - \lambda < 2\pi - \lambda_k/2$, and then

$$\begin{aligned}
& \int_{-\lambda_k/2}^{\lambda_k/2} (2\pi n)^{-1} \left\{ f_{XX}(\lambda) - f_{XX}(\lambda_j) \right\} \left| D(\lambda_j - \lambda) \right| \left| D(\lambda - \lambda_k) \right| d\lambda \qquad (23) \\
= \;& \int_{-\lambda_k/2}^{\lambda_k/2} (2\pi n)^{-1} \left\{ f_{XX}(\lambda) - f_{XX}(\lambda_j) \right\} \left| D(\lambda_j - \lambda) \right| \left| D(\lambda_k - \lambda) \right| d\lambda
\end{aligned}$$

$$
\begin{aligned}
&= O\left(\frac{1}{n\lambda_j\lambda_k}\int_{-\lambda_k/2}^{\lambda_k/2}\left\{\left|f_{XX}(\lambda)\right|+\left|f_{XX}(\lambda_j)\right|\right\}d\lambda\right)\\
&= O\left(\frac{1}{n\lambda_j\lambda_k}\left(\lambda_k^{1-2d}+\lambda_j^{-2d}\lambda_k\right)\right)\\
&= O\left(n^{-1}\lambda_j^{-1}\lambda_k^{-2d}+n^{-1}\lambda_j^{-1-2d}\right)\\
&= O\left(k^{-1}\lambda_j^{-1}\lambda_k^{1-2d}+n^{-1}\lambda_j^{-1-2d}\right)\\
&= O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\left(\frac{\lambda_k}{\lambda_j}\right)^{1-d}+n^{-1}\lambda_j^{-1-2d}\right)=O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\right).
\end{aligned}
$$

In evaluating the integral on $[-\lambda_j/2, -\lambda_k/2]$, we use sub-lemma 7.3.4 and $\lambda_j < \lambda_j + \lambda < 2\pi - \lambda_j/2$ for $\lambda_k/2 \le \lambda \le \lambda_j/2$, giving

$$
\begin{aligned}
&\int_{-\lambda j/2}^{-\lambda_k/2}(2\pi n)^{-1}\left\{f_{XX}(\lambda)-f_{XX}(\lambda_j)\right\}D(\lambda_j-\lambda)\,D(\lambda-\lambda_k)\,d\lambda\\
&\le \int_{\lambda_k/2}^{\lambda_j/2}(2\pi n)^{-1}\left\{f_{XX}(\lambda)-f_{XX}(\lambda_j)\right\}\left|D(\lambda_j+\lambda)\right|\left|D(\lambda+\lambda_k)\right|d\lambda\\
&= O\left(\max_{\lambda_k/2\le\lambda\le\lambda_j}\left|f_{XX}(\lambda)\right|\int_{\lambda_k/2}^{\lambda_j/2}n^{-1}\left|D(\lambda_j+\lambda)\right|\left|D(\lambda+\lambda_k)\right|d\lambda\right) \qquad (24)\\
&= O\left(\left(\lambda_k^{-2d}+\lambda_j^{-2d}\right)\frac{\ln n}{j}\right)\\
&= O\left(n^{-1}\lambda_j^{-1}\lambda_k^{-2d}\ln n+n^{-1}\lambda_j^{-1-2d}\ln n\right)=O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\right).
\end{aligned}
$$

7.3.8 Proof of (d)

$$
\begin{aligned}
E\left[w(\lambda_j)\,w(\lambda_k)\right] &= \int_{-\pi}^{\pi}\left\{f_{XX}(\lambda)-f_{XX}(\lambda_j)\right\}E_{j,-k}(\lambda)\,d\lambda\\
&= \int_{-\pi}^{\pi}(2\pi n)^{-1}\left\{f_{XX}(\lambda)-f_{XX}(\lambda_j)\right\}D(\lambda_j-\lambda)\,D(\lambda+\lambda_k)\,d\lambda
\end{aligned}
$$

The evaluations are similar to those in the proof of (c). In particular,

$$
\begin{aligned}
&\int_{\lambda_j/2}^{\pi}(2\pi n)^{-1}\left\{f_{XX}(\lambda)-f_{XX}(\lambda_j)\right\}D(\lambda_j-\lambda)\,D(\lambda+\lambda_k)\,d\lambda\\
&= O\left(\max_{\lambda_j/2\le\lambda\le\pi}\left|f'_{XX}(\lambda)\right|n^{-1}\int_{\lambda_j/2}^{\pi}\left|D(\lambda+\lambda_k)\right|d\lambda\right)\\
&= O\left(\lambda_j^{-1-2d}n^{-1}\ln n\right)\\
&= O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\right),
\end{aligned}
$$

because $-\pi < \lambda_j - \pi \leq \lambda_j - \lambda \leq \lambda_j/2 < \pi$. Next,

$$\begin{aligned}
& \int_{-\pi}^{-\lambda_j/2} (2\pi n)^{-1} \{f_{XX}(\lambda) - f_{XX}(\lambda_j)\} D(\lambda_j - \lambda) D(\lambda + \lambda_k) d\lambda \\
\leq & \int_{\lambda_j/2}^{\pi} (2\pi n)^{-1} \{f_{XX}(\lambda) - f_{XX}(\lambda_j)\} |D(\lambda_j + \lambda)| |D(\lambda - \lambda_k)| d\lambda \\
= & O\left(\max_{\lambda_j/2 \leq \lambda \leq \pi} \left| f'_{XX}(\lambda) \right| n^{-1} \int_{\lambda_j/2}^{\pi} |D(\lambda - \lambda_k)| d\lambda \right) \\
= & O\left(\lambda_j^{-1-2d} n^{-1} \ln n \right) \\
= & O\left(k^{-1} \lambda_j^{-d} \lambda_k^{-d} \ln n \right),
\end{aligned}$$

by $0 < \lambda_j + \lambda < 2\pi$ and sub-lemma 7.3.3. The same argument as used in (24) yields

$$\begin{aligned}
& \int_{-\lambda_j/2}^{-\lambda_k/2} (2\pi n)^{-1} \{f_{XX}(\lambda) - f_{XX}(\lambda_j)\} D(\lambda_j - \lambda) D(\lambda + \lambda_k) d\lambda \\
\leq & \int_{\lambda_k/2}^{\lambda_j/2} (2\pi n)^{-1} \{f_{XX}(\lambda) - f_{XX}(\lambda_j)\} |D(\lambda_j + \lambda)| |D(\lambda - \lambda_k)| d\lambda \\
= & O\left(\max_{\lambda_k/2 \leq \lambda \leq \lambda_j} |f_{XX}(\lambda)| \int_{\lambda_k/2}^{\lambda_j/2} n^{-1} |D(\lambda_j + \lambda)| |D(\lambda - \lambda_k)| d\lambda \right) \\
= & O\left(\left(\lambda_k^{-2d} + \lambda_j^{-2d} \right) \frac{\ln n}{j} \right) = O\left(k^{-1} \lambda_j^{-d} \lambda_k^{-d} \ln n \right).
\end{aligned}$$

Further, the same argument as in (23) yields

$$\begin{aligned}
& \int_{-\lambda_k/2}^{\lambda_k/2} (2\pi n)^{-1} \{f_{XX}(\lambda) - f_{XX}(\lambda_j)\} D(\lambda_j - \lambda) D(\lambda + \lambda_k) d\lambda \\
= & O\left(\frac{1}{n \lambda_j \lambda_k} \int_{-\lambda_k/2}^{\lambda_k/2} \{|f_{XX}(\lambda)| + |f_{XX}(\lambda_j)|\} d\lambda \right) \\
= & O\left(k^{-1} \lambda_j^{-d} \lambda_k^{-d} \ln n \right).
\end{aligned}$$

Finally,

$$\begin{aligned}
& \int_{\lambda_k/2}^{\lambda_j/2} (2\pi n)^{-1} \{f_{XX}(\lambda) - f_{XX}(\lambda_j)\} D(\lambda_j - \lambda) D(\lambda + \lambda_k) d\lambda \\
= & O\left(\max_{\lambda_k/2 \leq \lambda \leq \lambda_j} |f_{XX}(\lambda)| \int_{\lambda_k/2}^{\lambda_j/2} n^{-1} |D(\lambda_j - \lambda)| |D(\lambda + \lambda_k)| d\lambda \right) \\
= & O\left(\left(\lambda_k^{-2d} + \lambda_j^{-2d} \right) \frac{\ln n}{j} \right) = O\left(k^{-1} \lambda_j^{-d} \lambda_k^{-d} \ln n \right),
\end{aligned}$$

by $0 < \lambda_j/2 < \lambda_j - \lambda < \pi$ and sub-lemma 7.3.4. ■

7.4 Proof of Lemma 4.3

The following elementary result is useful.

7.4.1 Sub-Lemma

1. $\ln M = O(\ln m)$,
2. $\ln n = O(\ln m)$,
3. $\ln^{-1} n = O\left(\ln^{-1} m\right)$.

Proof $n = 2mM$ and $M/m \to 0$ implies

$$\begin{aligned} \ln M &= O(\ln m), \\ \ln n &= \ln 2 + \ln m + \ln M = O(\ln m), \\ \frac{\ln m}{\ln n} &= \frac{\ln m}{\ln 2 + \ln m + \ln M} = O(1). \blacksquare \end{aligned}$$

From the proof of lemma 3.3, we have

$$X_{sj} - \overline{X}_{.j} = O(1/j),$$

uniformly in $1 \leq j < M$. Also, $E(\varepsilon_{sj}) = O(\ln n/s)$ uniformly for $m \leq s < 2/n$ because $\ln^2 n = o(m)$. It follows that

$$\sum_{\{s:\lambda_s \in B_j\}} E(\varepsilon_{sj})\left(X_{sj} - \overline{X}_{.j}\right) = O\left(\frac{1}{j}\sum_{s=mj+1}^{mj+m} \frac{\ln n}{s}\right) = O\left(\frac{\ln n}{j^2}\right),$$

because $1/s = O(1/mj)$ in $\{s : \lambda_s \in B_j\}$. Therefore,

$$\sum_{j=1}^{\ell} \sum_{\{s:\lambda_s \in B_j\}} E(\varepsilon_{sj})\left(X_{sj} - \overline{X}_{.j}\right) = O\left(\sum_{j=1}^{\ell} \frac{\ln n}{j^2}\right) = O(\ln n) = O(\ln m). \blacksquare$$

7.5 Proof of Lemma 4.4

$\mathrm{Var}(\varepsilon_{sj}) = \pi^2/6 + O(\ln n/s)$ and $\mathrm{Cov}(\varepsilon_{sj}, \varepsilon_{tk}) = O\left(\ln^2 n/t^2\right)$ uniformly for $m \leq t < s < n/2$ because $\ln^2 n = o(m)$. Hence

$$\begin{aligned} &\mathrm{Var}\left[\sum_{\{s:\lambda_s \in B_j\}} \varepsilon_{sj}\left(X_{sj} - \overline{X}_{.j}\right)\right] \\ = &\sum_{\{s:\lambda_s \in B_j\}} \left(X_{sj} - \overline{X}_{.j}\right)^2 \mathrm{Var}(\varepsilon_{sj}) + 2\sum_{t=mj+1}^{mj+m} \sum_{s=t+1}^{mj+m} \left(X_{sj} - \overline{X}_{.j}\right)\left(X_{tj} - \overline{X}_{.j}\right) \mathrm{cov}(\varepsilon_{sj}, \varepsilon_{tj}) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{\{s:\lambda_s\in B_j\}} \left(X_{sj}-\overline{X}_{.j}\right)^2 \left(\frac{\pi^2}{6}+O\left(\frac{\ln n}{s}\right)\right)+O\left(\frac{1}{j^2}\sum_{t=mj+1}^{mj+m}\sum_{s=t+1}^{mj+m}\frac{\ln^2 n}{t^2}\right)\\
&= \frac{\pi^2}{6}\sum_{\{s:\lambda_s\in B_j\}} \left(X_{sj}-\overline{X}_{.j}\right)^2+O\left(\frac{1}{j^2}\sum_{s=mj+1}^{mj+m}\frac{\ln n}{s}\right)+O\left(\frac{1}{j^2}\sum_{t=mj+1}^{mj+m}\frac{m\ln^2 n}{t^2}\right)\\
&= \frac{\pi^2}{6}\sum_{\{s:\lambda_s\in B_j\}} \left(X_{sj}-\overline{X}_{.j}\right)^2+O\left(\frac{\ln n}{j^3}\right)+O\left(\frac{1}{j^2}\frac{m\ln^2 n}{mj}\right)\\
&= \frac{\pi^2}{6}\sum_{\{s:\lambda_s\in B_j\}} \left(X_{sj}-\overline{X}_{.j}\right)^2+O\left(\frac{\ln^2 n}{j^3}\right)\\
&= \frac{\pi^2}{6}\sum_{\{s:\lambda_s\in B_j\}} \left(X_{sj}-\overline{X}_{.j}\right)^2+O\left(\frac{\ln^2 m}{j^3}\right),
\end{aligned}
$$

because $1/s=O(1/mj)$ and $\sum_{t=mj+1}^{mj+m}\frac{1}{t^2}\sim\int_{mj+1}^{mj+m}\frac{dt}{t^2}=\frac{1}{mj+1}-\frac{1}{mj+m}=O(1/mj)$.

For the covariance, we have for $j>k$

$$
\begin{aligned}
&\operatorname{Cov}\left[\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right),\sum_{\{t:\lambda_t\in B_k\}}\varepsilon_{tk}\left(X_{tk}-\overline{X}_{.k}\right)\right]\\
&= \sum_{\{s:\lambda_s\in B_j\}}\sum_{\{t:\lambda_t\in B_k\}}\left(X_{sj}-\overline{X}_{.j}\right)\left(X_{tk}-\overline{X}_{.k}\right)\operatorname{Cov}\left(\varepsilon_{sj},\varepsilon_{tk}\right)\\
&= \sum_{\{s:\lambda_s\in B_j\}}\sum_{\{t:\lambda_t\in B_k\}}O\left(\frac{1}{jk}\right)O\left(\frac{\ln^2 n}{t^2}\right)\\
&= O\left(\frac{1}{jk}\sum_{s=mj+1}^{mj+m}\sum_{t=mk+1}^{mk+m}\frac{\ln^2 n}{t^2}\right)\\
&= O\left(\frac{1}{jk}\frac{m\ln^2 n}{mk}\right)=O\left(\frac{\ln^2 n}{jk^2}\right)=O\left(\frac{\ln^2 m}{jk^2}\right).\ \blacksquare
\end{aligned}
$$

7.6 Proof of Lemma 4.5

$$
\begin{aligned}
&\operatorname{Cov}\left[\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right),\sum_{\{t:\lambda_t\in B_0\}}\varepsilon_{t0}\left(X_{t0}-\overline{X}_{.0}\right)\right]\\
&= \operatorname{Cov}\left[\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right),\sum_{t=1}^{\ln^6 n}\varepsilon_{t0}\left(X_{t0}-\overline{X}_{.0}\right)\right]\\
&\quad+\operatorname{Cov}\left[\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right),\sum_{t=\ln^6 n+1}^{m}\varepsilon_{t0}\left(X_{t0}-\overline{X}_{.0}\right)\right].
\end{aligned}
$$

First,

$$\begin{aligned}
& \mathrm{Cov}\left[\sum_{\{s:\lambda_s\in B_j\}} \varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right), \sum_{t=1}^{\ln^6 n} \varepsilon_{t0}\left(X_{t0}-\overline{X}_{.0}\right)\right] \\
\leq\ & Var\left[\left(\sum_{\{s:\lambda_s\in B_j\}} \varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right)\right)\right]^{1/2}\left[Var\left(\sum_{t=1}^{\ln^6 n} \varepsilon_{t0}\left(X_{t0}-\overline{X}_{.0}\right)\right)\right]^{1/2} \\
=\ & O\left(\frac{\sqrt{m}\ln^7 m}{j}\right),
\end{aligned}$$

because

$$Var\left(\sum_{t=1}^{\ln^6 n} \varepsilon_{t0}\left(X_{t0}-\overline{X}_{.0}\right)\right) = O\left(\ln^{14} m\right),$$

by HDB p.25, and as shown in the proof of lemma 4.3,

$$\begin{aligned}
Var\left(\sum_{\{s:\lambda_s\in B_j\}} \varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right)\right) &= \frac{\pi^2}{6}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{.j}\right)^2 + O\left(\frac{\ln^2 n}{j^3}\right) \\
&= \frac{\pi^2}{6}\sum_{\{s:\lambda_s\in B_j\}} O\left(\frac{1}{j^2}\right) + O\left(\frac{\ln^2 n}{j^3}\right) \\
&= O\left(\frac{m}{j^2}\right).
\end{aligned}$$

Next,

$$\begin{aligned}
& \mathrm{Cov}\left[\sum_{\{s:\lambda_s\in B_j\}} \varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right), \sum_{t=\ln^6 n+1}^{m} \varepsilon_{t0}\left(X_{t0}-\overline{X}_{.0}\right)\right] \\
=\ & \sum_{\{s:\lambda_s\in B_j\}}\sum_{t=\ln^6 n+1}^{m}\left(X_{sj}-\overline{X}_{.j}\right)\left(X_{t0}-\overline{X}_{.0}\right)\mathrm{Cov}\left(\varepsilon_{sj},\varepsilon_{t0}\right) \\
=\ & O\left(\frac{\ln m}{j}\sum_{t=\ln^6 n+1}^{m}\sum_{s=mj+1}^{mj+m}\frac{\ln^2 n}{t^2}\right) \\
=\ & O\left(\frac{\ln m}{j}\sum_{t=\ln^6 n+1}^{m}\frac{m\ln^2 n}{t^2}\right) \\
=\ & O\left(\frac{m\ln m}{j\ln^4 n}\right) = O\left(\frac{m}{j\ln^3 m}\right),
\end{aligned}$$

by $\sum_{t=\ln^6 n+1}^{m}\frac{1}{t^2} \sim \int_{\ln^6 n+1}^{m}\frac{dt}{t^2} = \frac{1}{\ln^6 n+1} - \frac{1}{m} = O\left(\frac{1}{\ln^6 n}\right)$. ■

7.7 Proof of Lemma 4.6

$\mathrm{Var}\Big[\sum_{j=0}^{\ell}\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right)\Big]$ can be decomposed into the following parts:

$$\begin{aligned}
&\mathrm{Var}\left[\sum_{j=0}^{\ell}\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right)\right]\\
=&\sum_{j=0}^{\ell}\mathrm{Var}\left[\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right)\right]\\
&+\sum_{j\neq k}^{\ell}\mathrm{Cov}\left[\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right),\sum_{\{t:\lambda_t\in B_k\}}\varepsilon_{tk}\left(X_{tk}-\overline{X}_{.k}\right)\right]\\
=&\sum_{j=0}^{\ell}\mathrm{Var}\left[\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right)\right]\\
&+2\sum_{j=1}^{\ell}\mathrm{Cov}\left[\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right),\sum_{\{t:\lambda_t\in B_0\}}\varepsilon_{t0}\left(X_{t0}-\overline{X}_{.0}\right)\right]\\
&+2\sum_{k=1}^{\ell}\sum_{j=k+1}^{\ell}\mathrm{Cov}\left[\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right),\sum_{\{t:\lambda_t\in B_k\}}\varepsilon_{tk}\left(X_{tk}-\overline{X}_{.k}\right)\right].
\end{aligned}$$

By theorem 1 of HDB,

$$\mathrm{Var}\left[\sum_{\{s:\lambda_s\in B_0\}}\varepsilon_{s0}\left(X_{s0}-\overline{X}_{.0}\right)\right]=\frac{4\pi^2 m}{6}+o\left(m\right).$$

For the remaining parts,

$$\begin{aligned}
\sum_{j=1}^{\ell}\mathrm{Var}\left[\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right)\right]&=\sum_{j=1}^{\ell}\left(\frac{\pi^2}{6}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{.j}\right)^2+O\left(\frac{\ln^2 m}{j^3}\right)\right)\\
&=\frac{\pi^2}{6}4m\Xi+o\left(m\right)+O\left(\ln^2 m\right)\\
&=\frac{\pi^2}{6}4m\Xi+o\left(m\right),
\end{aligned}$$

$$\begin{aligned}
&\sum_{j=1}^{\ell}\mathrm{Cov}\left[\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right),\sum_{\{t:\lambda_t\in B_0\}}\varepsilon_{t0}\left(X_{t0}-\overline{X}_{.0}\right)\right]\\
=&\ O\left(\sqrt{m}\ln^7 m\ln M\right)+O\left(\frac{m\ln M}{\ln^3 m}\right)=o\left(m\right),
\end{aligned}$$

and

$$\sum_{k=1}^{\ell}\sum_{j=k+1}^{\ell}\mathrm{Cov}\left[\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right),\sum_{\{t:\lambda_t\in B_k\}}\varepsilon_{tk}\left(X_{tk}-\overline{X}_{.k}\right)\right]$$

$$=\sum_{k=1}^{\ell}\sum_{j=k+1}^{\ell}O\left(\frac{\ln^2 m}{jk^2}\right)=\sum_{k=1}^{\ell}O\left(\frac{\ln^2 m\ln M}{k^2}\right)=O\left(\ln^3 m\right)=o\left(m\right).\ \blacksquare$$

7.8 Proof of Theorem 4.8

Recall

$$\widehat{d}-d=\frac{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\eta_{sj}\left(X_{sj}-\overline{X}_{.j}\right)}{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{.j}\right)^2}+\frac{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right)}{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{.j}\right)^2}.$$

By lemma 8 of HDB,

$$E\left(\sum_{\{s:\lambda_s\in B_0\}}\varepsilon_{s0}\left(X_{s0}-\overline{X}_{.0}\right)\right)=O\left(\ln^3 m\right).$$

Using this result and lemmas 3.2, 3.3, 4.3 and 4.6, we obtain

$$\begin{aligned}
E\left(\widehat{d}-d\right) &= \frac{-\frac{2\pi^2}{3}\frac{f_{uu}''(0)}{f_{uu}(0)}\frac{m^3L}{n^2}+o\left(\frac{m^3L}{n^2}\right)}{4m\left(1+\Xi\right)+o\left(m\right)}+\frac{O\left(\ln^3 m\right)+O\left(\ln m\right)}{4m\left(1+\Xi\right)+o\left(m\right)}\\
&= -\frac{\pi^2}{6\left(1+\Xi\right)}\frac{f_{uu}''\left(0\right)}{f_{uu}\left(0\right)}\frac{m^2L}{n^2}+o\left(\frac{m^2L}{n^2}\right)+O\left(\frac{\ln^3 m}{m}\right),\\
Var\left(\widehat{d}\right) &= \left[4m\left(1+\Xi\right)+o\left(m\right)\right]^{-2}\left[\frac{4\pi^2m}{6}\left(1+\Xi\right)+o\left(m\right)\right]\\
&= \frac{\pi^2}{24\left(1+\Xi\right)m}+o\left(\frac{1}{m}\right).
\end{aligned}$$

For the mean squared error, we have

$$\begin{aligned}
MSE\left(\widehat{d}\right) &= \left\{-\frac{\pi^2}{6\left(1+\Xi\right)}\frac{f_{uu}''\left(0\right)}{f_{uu}\left(0\right)}\frac{m^2L}{n^2}+o\left(\frac{m^2L}{n^2}\right)+O\left(\frac{\ln^3 m}{m}\right)\right\}^2\\
&\quad+\frac{\pi^2}{24\left(1+\Xi\right)m}+o\left(\frac{1}{m}\right)\\
&= \frac{\pi^4}{36\left(1+\Xi\right)^2}\left\{\frac{f_{uu}''\left(0\right)}{f_{uu}\left(0\right)}\right\}^2\frac{m^4L^2}{n^4}+\frac{\pi^2}{24\left(1+\Xi\right)m}\\
&\quad+o\left(\frac{m^2L}{n^2}\right)\left(\frac{m^2L}{n^2}\right)+O\left(\frac{\ln^3 m}{m}\right)\left(\frac{\ln^3 m}{m}\right)+O\left(\frac{m^2L}{n^2}\frac{\ln^3 m}{m}\right)+o\left(\frac{1}{m}\right)\\
&= \frac{\pi^4}{36\left(1+\Xi\right)^2}\left\{\frac{f_{uu}''\left(0\right)}{f_{uu}\left(0\right)}\right\}^2\frac{m^4L^2}{n^4}+\frac{\pi^2}{24\left(1+\Xi\right)m}\\
&\quad+o\left(\frac{m^4L^2}{n^4}\right)+O\left(\frac{\ln^6 m}{m^2}\right)+O\left(\frac{mL\ln^3 m}{n^2}\right)+o\left(\frac{1}{m}\right).
\end{aligned}$$

7.9 Proof of Lemma 5.1

First, we modify the propositions in Theorem 2 of Robinson (1995, p.1056). The assumptions $f'_{uu}(0) = 0$ and $|f''_{uu}(\omega)| < B_2 < \infty$ imply $\alpha = 2$. The univariate model in our case implies there is no β here. Therefore, (4.2) on p.1060 in Robinson changes to

$$\rho(\lambda_j) - C_g \lambda_j^{-\delta} = O\left[\left(\frac{j}{n}\right)^2 \lambda_j^{-\delta}\right]$$

and (a)-(d) in his Theorem 2 become

$$\begin{aligned}
(a') \quad E\left[v(\lambda_j)\overline{v}(\lambda_j)\right] &= 1 + O\left[\frac{\ln j}{j} + \left(\frac{j}{n}\right)^2\right] \\
(b') \quad E\left[v(\lambda_j) v(\lambda_j)\right] &= O\left(\frac{\ln j}{j}\right) \\
(c') \quad E\left[v(\lambda_j)\overline{v}(\lambda_k)\right] &= O\left(\frac{\ln j}{k}\right), \quad j > k \\
(d') \quad E\left[v(\lambda_j) v(\lambda_k)\right] &= O\left(\frac{\ln j}{k}\right), \quad j > k
\end{aligned}$$

where

$$\lambda_k = 2\pi k/n, \quad v(\lambda) = w(\lambda) / \left(C_g^{1/2}\lambda^{-d}\right), \quad w(\lambda) = (2\pi n)^{-1/2}\sum\nolimits_1^n X_t e^{it\lambda}.$$

In the following, we repeat Robinson's argument (1995, pp.1067-70) with corresponding modifications, although we try to keep the derivations here as self-contained as possible. Write $\chi_k = m^{-1/2} a_k U_k$. Fix an integer N. Then, $E\left(\sum_k \chi_k\right)^N$ is a sum of finitely many terms of the form

$$\sum_{k_1} \cdots \sum_{k_K} E\left(\prod_{i=1}^K \chi_{k_i}^{N_{k_i}}\right), \tag{25}$$

where $N_{k_1}, ..., N_{k_K}$ are all positive and sum to N and $1 \le K \le N$. Fix such a K and $N_{k_1}, ..., N_{k_K}$. Introduce the 2−vector $v_k^* = \left(v^R(\lambda_k), v^I(\lambda_k)\right)'$ and $2K \times 1$ vector $v^* = \left(v_{k_1}^{*\prime}, ..., v_{k_K}^{*\prime}\right)$, which is normally distributed with zero mean.

Theorem 2 of Robinson (1995) implies that

$$\begin{aligned}
E\left(v_j^* v_k^*\right) &= R + O\left(\frac{\ln j}{j} + \left(\frac{j}{n}\right)^2\right), \quad j = k, \quad R = \frac{1}{2} I_2, \\
&= O\left(\frac{\ln j}{k}\right), \quad j > k,
\end{aligned}$$

as $n \to \infty$. It follows that, $\Sigma^* = E\left(v^* v^{*\prime}\right)$ satisfies

$$\Sigma^* = I_K \otimes R + O\left(\frac{\ln \ell m}{m^{0.5+\delta}} + \left(\frac{\ell m}{n}\right)^2\right) = I_K \otimes R + o\left(m^{-1/2-\Delta}\right), \tag{26}$$

as $n \to \infty$.

For n sufficiently large, $\Psi = \Sigma^{*-1}$ exists by (26). Denote by Ψ_{ij} the (i,j) th 2×2 submatrix of Ψ and write

$$\widetilde{\Psi} = \begin{bmatrix} \Psi_{11} & & 0 \\ & \ddots & \\ 0 & & \Psi_{KK} \end{bmatrix}, \quad \overline{\Psi} = \Psi - \widetilde{\Psi}.$$

It follows that

$$\widetilde{\Psi} = I_K \otimes R^{-1} + o\left(m^{-1/2-\Delta}\right), \quad \overline{\Psi} = o\left(m^{-1/2-\Delta}\right) \quad \text{as} \quad n \to \infty. \tag{27}$$

Now denote by φ_p the density function of a p-dimensional standard normal variate. Then (25) is

$$\sum_{k_1} \cdots \sum_{k_K} |\Psi|^{1/2} \int \left(\prod_{i=1}^{K} \chi_{k_i}^{N_{k_i}}\right) \varphi_{2K}\left(\Psi^{1/2} v^*\right) dv^*, \tag{28}$$

for n sufficiently large. Consider

$$\sum_{k_1} \cdots \sum_{k_K} |\Psi|^{1/2} \prod_{i=1}^{K} \left\{\int \chi_{k_i}^{N_{k_i}} \varphi_2\left(\Psi_{ii}^{1/2} v_{k_i}^*\right) dv_{k_i}^*\right\}. \tag{29}$$

The difference between (28) and (29) is

$$\sum_{k_1} \cdots \sum_{k_K} |\Psi|^{1/2} \int \left(\prod_{i=1}^{K} \chi_{k_i}^{N_{k_i}}\right) \varphi_{2K}\left(\widetilde{\Psi}^{1/2} v^*\right) \left\{\exp\left(-\frac{1}{2} v^{*\prime} \overline{\Psi} v^*\right) - 1\right\} dv^*. \tag{30}$$

For any positive integer r, the mean value theorem indicates that $\left|e^u - \sum_{t=0}^{r-1} u^t/t!\right| \le |u|^r e^{|u|}/r!$, for all u. For all $\varepsilon > 0$ there exists $C_\varepsilon < \infty$ such that $|u|^r \le C_\varepsilon e^{|u|}$ for all u. Following (27), choose n so large that $\left\|\overline{\Psi}\right\| < \varepsilon$, where $\|\cdot\|$ is the Euclidean norm, that is, $\exp\left(\frac{1}{2}\left|v^{*\prime}\overline{\Psi}v^*\right|\right) \le \exp\left(\frac{1}{2}\varepsilon \|v^*\|^2\right)$. Again using (27), $\left|v^{*\prime}\overline{\Psi}v^*\right| = o\left(m^{-1/2-\Delta}\|v^*\|^2\right)$ uniformly in v^*. Thus

$$\left|\exp\left(-\frac{1}{2} v^{*\prime}\overline{\Psi}v^*\right) - \sum_{t=0}^{r-1} \frac{\left(-\frac{1}{2}v^{*\prime}\overline{\Psi}v^*\right)^t}{t!}\right| = o\left(m^{-r/2-r\Delta} \exp\left(\varepsilon \|v^*\|^2\right)\right),$$

as $n \to \infty$, uniformly in v^*. Thus the difference between (30) and

$$\sum_{k_1} \cdots \sum_{k_K} |\Psi|^{1/2} \int \left(\prod_{i=1}^{K} \chi_{k_i}^{N_{k_i}}\right) \varphi_{2K}\left(\widetilde{\Psi}^{1/2} v^*\right) \sum_{t=1}^{r-1} \frac{\left(-\frac{1}{2}v^{*\prime}\overline{\Psi}v^*\right)^t}{t!} dv^* \tag{31}$$

is

$$o\left(m^{-r/2-r\Delta} \sum_{k_1} \cdots \sum_{k_K} |\Psi|^{1/2} \int \prod_{i=1}^{K} \left|\chi_{k_i}^{N_{k_i}}\right| \exp\left\{-\frac{1}{2} v^{*\prime}\left(\widetilde{\Psi} - \varepsilon I_{2K}\right) v^*\right\} dv^*\right). \tag{32}$$

In view of (26), $|\Psi| = O(1)$, while $\frac{1}{2}v^{*\prime}\left(\widetilde{\Psi} - \varepsilon I_{2K}\right)v^* > \frac{1}{2}\eta\|v^*\|^2$ for some $\eta > 0$. Because $\|\chi_k\| \le m^{-1/2}|a_k|\|U_k\|$, we deduce from the finiteness of the moments of all orders of the log of a chi-squared variate that (32) is $o\left(\ln^K \ell \cdot m^{K-N/2-r/2-r\Delta}\right) \to 0$ on choosing $r = \max(2K - N, 1)$.

Now (31) makes a contribution only when such $r \ge 2$, which occurs only when $2K - N \ge 2$. Let D be the number of N_{k_i} which equal 1. Clearly $D \ge 2K - N$, that is, $D > t$ for $t = 1, ..., r-1 = 2K - N - 1$ in (31). Note that $v^{*\prime}\overline{\Psi}v^*$ is bilinear in the $v^*_{k_i}$ and for each $t = 1, ..., r-1$, hence $\left(v^{*\prime}\overline{\Psi}v^*\right)^t$ cannot involve more than t of the $v^*_{k_i}$. The corresponding t or fewer k_i can overlap with the Dk_i for which $N_{k_i} = 1$, but because $D > r - 1$, the $(k_1, ..., k_K, t)'$th summand in (31) can be written

$$|\Psi|^{1/2}\prod_{i=1}^{D-t}\left(\int \chi_{k_i}\varphi_2\left(\Psi_{ii}^{1/2}v^*_{k_i}\right)dv^*_{k_i}\right) \tag{33}$$

$$\times\int\frac{\left(-1/2v^{*\prime}\overline{\Psi}v^*\right)^t}{t!}\left(\prod_{i=D-t+1}^{K}\chi_{k_i}^{N_{k_i}}\varphi_2\left(\Psi_{ii}^{1/2}v^*_{k_i}\right)dv^*_{k_i}\right). \tag{34}$$

>From (27),

$$\varphi_2\left(\Psi_{ii}^{1/2}v^*_{k_i}\right) = \varphi_2\left(R^{-1/2}v^*_{k_i}\right)\left(1 + o\left(m^{-1/2-\Delta}\right)\left\|v^*_{k_i}\right\|^2\right), \tag{35}$$

uniformly in $v^*_{k_i}$ and $\int \chi_k\varphi_2\left(R^{-1/2}v^*_k\right)dv^*_k = 0$. For all positive p and q, uniformly in k,

$$\int\left\|m^{1/2}\chi_k\right\|^p\|v^*_k\|^q\varphi_2\left(R^{-1/2}v^*_k\right)dv^*_k = O\left(|a_k|^p\right).$$

Thus (33) is $o\left(m^{-(D-t)-\Delta(D-t)}\prod_{i=1}^{D-t}|a_{k_i}|\right)$ and (34) is $o\left(m^{-t-(N-D)/2-t\Delta}\prod_{i=D-t+1}^{K}|a_{k_i}|^{N_{k_i}}\right)$. It follows from the third part of (14) that (31) is $o\left(\ln^K \ell \cdot m^{K-N/2-D/2-\Delta D}\right) \to 0$ as $n \to \infty$. Thus, we have shown that (30) $\to 0$ as $n \to \infty$. Now from (27), $|\Psi| = |R|^{-K} + o\left(m^{-1/2-\Delta}\right)$ and

$$|R|^{-1/2}\int \chi_{k_i}^{N_{k_i}}\varphi_2\left(\Psi_{ii}^{1/2}v^*_{k_i}\right)dv^*_{k_i} = \mu_{k_i}^{(N_{k_i})}\left(1 + o\left(m^{-1/2-\Delta}\right)\right),$$

where $\mu_k^{(p)} = |R|^{-1/2}\int \chi_k^p\varphi_2\left(R^{-1/2}v^*_k\right)dv^*_k$. The difference between (29) and

$$\sum_{k_1}\cdots\sum_{k_K}\prod_{i=1}^{K}\mu_{k_i}^{(N_{k_i})} \tag{36}$$

is readily seen to be $o\left(\ln^K \ell \cdot m^{K-N/2-(1/2)\max(1,D)-\Delta}\right)$ using (35), and using $K - N/2 - D/2 \le 0$ when $D \ge 1$ and $K \le N/2$ when $D = 0$. However, (36) is

$$E\left[\sum_{k_1}\cdots\sum_{k_K}\prod_{i=1}^{K}\left(\frac{a_{k_i}W_{k_i}}{m^{1/2}}\right)^{N_{k_i}}\right].$$

Therefore, we have shown that the moments of $\sum_k \chi_k$ differ negligibly from those of the variate $m^{-1/2}\sum_k a_kW_k$, which converges in distribution to $N\left(0, \psi'(1)\right)$ upon applying (14), $W_k \sim iid\left(0, \psi'(1)\right)$ and the Lindeberg-Feller CLT. ■

7.10 Proof of Lemma 5.2

We adopt the argument from HDB, theorem 2. Let

$$m^{-1/2}\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right)\equiv T_1+T_2+T_3,$$

where

$$\begin{aligned}
T_1 &= \frac{1}{m^{1/2}}\sum_{s=1}^{\ln^8 m}\varepsilon_{s0}\left(X_{s0}-\overline{X}_{.0}\right),\\
T_2 &= \frac{1}{m^{1/2}}\sum_{s=1+\ln^8 m}^{m^{0.5+\delta}}\varepsilon_{s0}\left(X_{s0}-\overline{X}_{.0}\right),\\
T_3 &= \frac{1}{m^{1/2}}\sum_{s=1+m^{0.5+\delta}}^{m}\varepsilon_{s0}\left(X_{s0}-\overline{X}_{.0}\right)+\frac{1}{m^{1/2}}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right),
\end{aligned}$$

where $0<\delta<0.5$. Note that the condition required for theorem 2 of HDB, i.e., $m=o\left(n^{4/5}\right)$, is satisfied by our assumption. HDB show that $T_1=o_p(1)$ and $T_2=o_p(1)$.

We now prove that T_3 is asymptotically normal. Let

$$U_s=\ln I_s-\ln f_{uu}(0)-\psi'(1)+2d\ln\lambda_s,$$

as defined in Equation (2,4) of Robinson (1995). Then we have

$$U_s=\varepsilon_s+\ln\left\{\frac{f_{uu}(\lambda_s)}{f_{uu}(0)}\right\}-2d\ln\left\{\frac{|1-\exp(-i\lambda_s)|}{\lambda_s}\right\},$$

where $\varepsilon_s=\varepsilon_{s}.$ (we drop the second subscript of ε_{sj}). Hence,

$$T_3\equiv T_{31}+T_{32}+T_{33},$$

where

$$\begin{aligned}
T_{31} &= \frac{1}{m^{1/2}}\sum_{s=1+m^{0.5+\delta}}^{m}U_s\left(X_{s0}-\overline{X}_{.0}\right)+\frac{1}{m^{1/2}}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}}U_{sj}\left(X_{sj}-\overline{X}_{.j}\right),\\
T_{32} &= -\frac{1}{m^{1/2}}\sum_{s=1+m^{0.5+\delta}}^{m}\ln\left\{\frac{f_{uu}(\lambda_s)}{f_{uu}(0)}\right\}\left(X_{s0}-\overline{X}_{.0}\right)\\
&\quad -\frac{1}{m^{1/2}}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}}\ln\left\{\frac{f_{uu}(\lambda_s)}{f_{uu}(0)}\right\}\left(X_{sj}-\overline{X}_{.j}\right),\\
T_{33} &= \frac{2d}{m^{1/2}}\sum_{s=1+m^{0.5+\delta}}^{m}\ln\left\{\frac{|1-\exp(-i\lambda_s)|}{\lambda_s}\right\}\left(X_{s0}-\overline{X}_{.0}\right)\\
&\quad +\frac{2d}{m^{1/2}}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}}\ln\left\{\frac{|1-\exp(-i\lambda_s)|}{\lambda_s}\right\}\left(X_{sj}-\overline{X}_{.j}\right).
\end{aligned}$$

HDB show that the first term of T_{32} is $o(1)$ and $\ln\{f_{uu}(\lambda_s)/f_{uu}(0)\} = O\left(s^2/n^2\right)$ uniformly for $1 \le s \le m \ln M$. Hence the second term is, by $s = O(mj)$,

$$\begin{aligned} O\left(\frac{1}{m^{1/2}}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}}\frac{s^2}{n^2}\left(X_{sj}-\overline{X}_{.j}\right)\right) &= O\left(\frac{1}{m^{1/2}}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}}\frac{s^2}{n^2}\frac{1}{j}\right) \\ &= O\left(\frac{1}{m^{1/2}}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}}\frac{jm^2}{n^2}\right) \\ &= O\left(\frac{m^{2.5}L^2}{n^2}\right), \end{aligned}$$

which is $o(1)$ by $m^5 = O\left(n^{4-5\varepsilon}\right)$ and $L = O(\ln M)$. Hence, $T_{32} = o(1)$.

HDB show that the first term of T_{33} is $o(1)$. For the second term, from the proof of lemma 3.2(b), we have for $m \le s \le m(L+1)$

$$\ln\left|1-e^{i\lambda_s}\right| = -\frac{1}{2}X_{sj} = \ln\lambda_s + O\left(\frac{j^2}{M^2}\right),$$

uniformly in s. Hence,

$$\ln\left\{\frac{|1-\exp(-i\lambda_s)|}{\lambda_s}\right\} = O\left(\frac{j^2}{M^2}\right),$$

uniformly in s. Thus, the second term

$$\begin{aligned} &\frac{2d}{m^{1/2}}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}}\ln\left\{\frac{|1-\exp(-i\lambda_s)|}{\lambda_s}\right\}\left(X_{sj}-\overline{X}_{.j}\right) \\ &= O\left(\frac{1}{m^{1/2}}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}}\frac{j^2}{M^2}\frac{1}{j}\right) \\ &= O\left(\frac{mL^2}{m^{1/2}M^2}\right) \\ &= O\left(\frac{m^{2.5}L^2}{n^2}\right) = o(1). \end{aligned}$$

Thus, $T_{33} = o(1)$.

Finally, we need to prove that

$$\begin{aligned} T_{31} &= \frac{1}{m^{1/2}}\sum_{s=1+m^{0.5+\delta}}^{m}U_s\left(X_{s0}-\overline{X}_{.0}\right) + \frac{1}{m^{1/2}}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}}U_{sj}\left(X_{sj}-\overline{X}_{.j}\right) \\ &\xrightarrow{d} N\left(0, 4\pi^2(1+\Xi)/6\right). \end{aligned}$$

For $s = 1+m^{0.5+\delta}, ..., m$, HDB show that $\left(X_{sj}-\overline{X}_{.j}\right)/4$ satisfies the condition (14) of lemma 5.2. For $s = m+1, ..., m(L+1)$, we can use the argument in the proof of lemma 3.2 to obtain

$$\left|X_{sj}-\overline{X}_{.j}\right| = O\left(1/j\right) = o\left(m\right), \quad \frac{1}{4\Xi}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{.j}\right)^2 \sim m,$$

and

$$\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}}\left|X_{sj}-\overline{X}_{.j}\right|^p = \sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}} O\left(\frac{1}{j^p}\right) = O\left(\sum_{j=1}^{L}\frac{m}{j^p}\right) = O\left(m\ln L\right).$$

Hence $\left(X_{sj}-\overline{X}_{.j}\right)/4\left(1+\Xi\right),\ s = 1+m^{0.5+\delta},...,m\left(L+1\right),$ satisfies the condition (14) of lemma 5.2. Also $\ln M$ satisfies the condition for ℓ in Lemma 5.2, because $\ell^2 m^2 m^{0.5+\Delta}/n^2 = O\left(n^{-2.5\varepsilon}m^{\Delta}\ln^2 M\right) \to 0$ by a proper choice of Δ and $\ln^K \ln M/m^{\Delta} \to 0$ for any $K>0$.
Therefore,

$$\frac{1}{m^{1/2}}\sum_{s=1+m^{0.5+\delta}}^{m} U_s\left(X_{s0}-\overline{X}_{.0}\right)+\frac{1}{m^{1/2}}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}} U_{sj}\left(X_{sj}-\overline{X}_{.j}\right) \xrightarrow{d} N\left(0,4\pi^2\left(1+\Xi\right)/6\right). \blacksquare$$

7.11 Proof of Theorem 5.3

Let

$$\begin{aligned} m^{1/2}\left(\widehat{d}-d\right) &= \frac{m^{1/2}\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\eta_{sj}\left(X_{sj}-\overline{X}_{.j}\right)}{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{.j}\right)^2} \\ &\quad + \frac{m}{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{.j}\right)^2}\frac{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right)}{m^{1/2}} \\ &= V_1+V_2. \end{aligned}$$

Now

$$\begin{aligned} V_1 &= O\left(\frac{m^{1/2}}{m}\frac{m^3}{n^2}\right) = O\left(\frac{m^{5/2}}{n^2}\right) = o\left(1\right), \\ V_2 &= \frac{m}{4\left(1+\Xi\right)m+o\left(m\right)}\frac{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{.j}\right)\varepsilon_{sj}}{m^{1/2}} \xrightarrow{d} N\left(0,\frac{\pi^2}{24\left(1+\Xi\right)}\right), \end{aligned}$$

giving the required result. $\blacksquare$

References

[1] Davies, R. B., and D. S. Harte (1987) "Tests for Hurst Effect." *Biometrika* **74**: 95-101.

[2] Geweke, J., and S. Porter-Hudak (1983). "The Estimation and Application of Long Memory Time Series Models." *Journal of Time Series Analysis* **4**: 221-38.

[3] Gradshteyn, I. S., and I. M. Ryzhik (1965). *Tables of Integrals, Series and Products.* New York, Academic Press.

[4] Hurvich, C. M., and K. I. Beltrao (1994). "Automatic Semiparametric Estimation of the Memory Parameter of a Long-Memory Time Series." *Journal of Time Series Analysis* **15**: 285-302.

[5] Hurvich, C. M., R. Deo, and J. Brodsky (1998). "The Mean Squared Error of Geweke and Porter-Hudak's Estimator of the Memory Parameter of a Long-Memory Time Series." *Journal of Time Series Analysis* **19**: 19-46.

[6] Künsh, H. (1986). "Discrimination between Monotonic Trends and Long-Range Dependence." *Journal of Applied Probability* **23**: 1025-30.

[7] Robinson, P. M. (1995). "Log-Periodogram Regression of Time Series with Long Range Dependence." *Annals of Statistics* **23**: 1048-72.

[8] Zygmund, A. (1959). *Trigonometric Series.* Cambridge, Cambridge University Press.

8 Monte Carlo simulation results

Table 3. G $((a_1, a_2) = (-0.5, 0.0))$, $d = 0.3$

	BIAS	VAR	MSE	BIAS	VAR	MSE	BIAS	VAR	MSE
	$n = 200 \quad (m = 31)$			$n = 500 \quad (m = 56)$			$n = 1000 \quad (m = 89)$		
$\widehat{d}_{GPH}$	-0.0221	0.0156	0.0161	-0.0190	0.0073	0.0077	0.0252	0.0065	0.0071
$\widehat{d}_{pooled}$	-0.0580	0.0148	0.0182	-0.0400	0.0071	0.0087	0.0167	0.0068	0.0070
$\widehat{d}_{GPH}(3m)$	-0.3242	0.0051	0.1102	-0.1439	0.0020	0.0228	-0.0800	0.0014	0.0078
$\widehat{d}_{pooled}(m)$	0.0918	0.0647	0.0731	0.0246	0.0176	0.0182	0.0207	0.0092	0.0096

Table 4. H $((a_1, a_2) = (0.0, 0.0))$, $d = 0.3$

	BIAS	VAR	MSE	BIAS	VAR	MSE	BIAS	VAR	MSE
	$n = 200 \quad (m = 31)$			$n = 500 \quad (m = 56)$			$n = 1000 \quad (m = 89)$		
$\widehat{d}_{GPH}$	0.0026	0.0154	0.0154	-0.0035	0.0073	0.0073	0.0343	0.0065	0.0077
$\widehat{d}_{pooled}$	0.0019	0.0148	0.0148	-0.0046	0.0071	0.0071	0.0374	0.0066	0.0080
$\widehat{d}_{GPH}(3m)$	-0.0478	0.0050	0.0073	-0.0118	0.0021	0.0022	-0.0015	0.0014	0.0014
$\widehat{d}_{pooled}(m)$	0.1174	0.0708	0.0846	0.0323	0.0177	0.0187	0.0278	0.0096	0.0104

Table 5. I $((a_1, a_2) = (0.5, 0.0))$, $d = 0.3$

	BIAS	VAR	MSE	BIAS	VAR	MSE	BIAS	VAR	MSE
	$n = 200 \quad (m = 31)$			$n = 500 \quad (m = 56)$			$n = 1000 \quad (m = 89)$		
$\widehat{d}_{GPH}$	0.1697	0.0158	0.0446	0.0932	0.0075	0.0162	0.0992	0.0065	0.0163
$\widehat{d}_{pooled}$	0.1936	0.0152	0.0527	0.1203	0.0072	0.0217	0.1281	0.0066	0.0230
$\widehat{d}_{GPH}(3m)$	0.3160	0.0049	0.1047	0.2888	0.0021	0.0855	0.2461	0.0014	0.0620
$\widehat{d}_{pooled}(m)$	0.2059	0.0712	0.1136	0.0793	0.0174	0.0237	0.0593	0.0097	0.0132

Table 6. A $((a_1, a_2) = (-0.6, -0.6))$, $d = 0.3$

	BIAS	VAR	MSE	BIAS	VAR	MSE	BIAS	VAR	MSE
	$n = 200 \quad (m = 31)$			$n = 500 \quad (m = 56)$			$n = 1000 \quad (m = 89)$		
$\widehat{d}_{GPH}$	-0.0988	0.0151	0.0249	-0.0539	0.0073	0.0102	0.0063	0.0067	0.0067
$\widehat{d}_{pooled}$	-0.1629	0.0148	0.0413	-0.1240	0.0070	0.0224	-0.0376	0.0068	0.0083
$\widehat{d}_{GPH}(3m)$	-0.5242	0.0042	0.2789	-0.4990	0.0020	0.2510	-0.2886	0.0015	0.0848
$\widehat{d}_{pooled}(m)$	0.0352	0.0539	0.0552	0.0052	0.0170	0.0170	0.0096	0.0089	0.0090

Table 7. B $((a_1, a_2) = (0.6, -0.6))$, $d = 0.3$

	BIAS	VAR	MSE	BIAS	VAR	MSE	BIAS	VAR	MSE
	$n = 200 \quad (m = 31)$			$n = 500 \quad (m = 56)$			$n = 1000 \quad (m = 89)$		
$\widehat{d}_{GPH}$	-0.2376	0.0152	0.0717	-0.1079	0.0074	0.0190	-0.0261	0.0066	0.0073
$\widehat{d}_{pooled}$	-0.1286	0.0147	0.0313	-0.1003	0.0071	0.0171	-0.0384	0.0067	0.0082
$\widehat{d}_{GPH}(3m)$	0.1692	0.0042	0.0328	-0.0821	0.0021	0.0088	-0.2384	0.0014	0.0582
$\widehat{d}_{pooled}(m)$	0.0127	0.0731	0.0732	-0.0156	0.0181	0.0183	-0.0011	0.0094	0.0094

Table 8. C $((a_1, a_2) = (1.0, -0.6))$, $d = 0.3$

	BIAS	VAR	MSE	BIAS	VAR	MSE	BIAS	VAR	MSE
	$n = 200 \quad (m = 31)$			$n = 500 \quad (m = 56)$			$n = 1000 \quad (m = 89)$		
$\widehat{d}_{GPH}$	-0.2478	0.0153	0.0767	-0.1574	0.0074	0.0322	-0.0585	0.0066	0.0101
$\widehat{d}_{pooled}$	-0.1350	0.0148	0.0330	-0.0604	0.0072	0.0108	-0.0179	0.0066	0.0069
$\widehat{d}_{GPH}(3m)$	0.4783	0.0051	0.2339	0.2561	0.0021	0.0677	0.0635	0.0015	0.0055
$\widehat{d}_{pooled}(m)$	0.0018	0.0728	0.0728	-0.0353	0.0187	0.0200	-0.0155	0.0094	0.0096