Wavelet Transform for Log Periodogram Regression in Long Memory Stochastic Volatility Model

By JIN LEE

Department of Economics, National University of Singapore

Preliminary version

March, 2004

Address: Department of Economics
AS 2, Level 6
National University of Singapore
Singapore 117570

E-mail: ecsleej@nus.edu.sg
Fax: +65 6775-2646

Abstract. We consider semiparametric log periodogram regression estimation of memory parameter for the latent process in long memory stochastic volatility models. It is known that though widely used among researchers, the Geweke and Porter-Hudak (1983; GPH) LP estimator violates the Gaussian or Martingale assumption, which results in significant negative bias due to the existence of the spectrum of non-Gaussian noise. Through wavelet transform of the squared process, we effectively remove the noise spectrum around zero frequency, and obtain Gaussian-approximate spectral representation at zero frequency. We propose wavelet-based regression estimator, and derive the asymptotic mean squared error and the consistency in line with the asymptotic theory in the long memory literature. Simulation studies show that wavelet-based regression estimation is an effective way in reducing the bias, compared with the GPH estimator.

Keywords. long memory stochastic volatility, wavelet transform, log periodogram regression.

1. INTRODUCTION

We consider log periodogram (LP) regression estimation of memory parameter of the latent process in long memory stochastic volatility (LMSV) models. In LMSV models, the spectral density of the nonlinear processes such as squared or log squared process is the sum of the spectral density of Gaussian long memory process and that of non-Gaussian noise. Though the LP estimator of Geweke and Porter-Hudak (1983; GPH) is widely used in estimating memory parameter, it is well known that statistical inferences for LP estimator developed in the long memory literature are not directly applicable to LMSV model (Bollerslev and Wright (2000)). As clearly pointed out in Deo and Hurvich (2001), given the spectral representation of the squared processes, GPH estimator violates the Gaussian or Martingale assumption which the asymptotic theory is built upon in the long memory literature. As a result, GPH estimator suffers from significant negative bias mainly due to the existence of the spectrum of non-Gaussian noise. (Breidt et al (1998) and Deo and Hurvich (2001)).

In this paper, we introduce wavelet transform of the squared process to effectively remove the noise spectrum around zero frequency, and obtain Gaussian-approximate spectral representation. Thus, wavelet transformation can retrieve the Gaussianity, where the statistical inferences in the long memory context can be applied to LMSV model in straightforward manner. We derive the asymptotic mean squared error of the wavelet-based regression estimator, in line with Robinson (1995), Hurvich, Deo and Brodsky (1998) and Andrews and Guggenberger (2003). It is noted that the conditions for the consistency and for the convergence rate of the mean squared error on the growth rate of the fundamental frequency of GPH estimator depend on unknown memory parameter, which pose a limitation for practical use. Under Gaussianity through wavelet transformation, our proposed estimator is free from such problem.

2. THE MODEL

We consider a long memory stochastic volatility model for discrete time series $\{X_t, t = 1, 2, \cdots, n\}$

$$X_t = \sigma \exp(Z_t/2) e_t \qquad (1)$$

where $\{Z_t\}$ is a latent Gaussian long memory process with the memory parameter $d \in (0, 0.5)$, which is independent of mean zero i.i.d. process $\{e_t\}$. We assume that the spectral behavior of Z_t at zero frequency, which is standard in the long memory context.

Assumption 1: $f_Z(\lambda) = \lambda^{-2d} g(\lambda) \quad \text{as } \lambda \to 0,$

where $g(\lambda)$ is an even function on $[-\pi, \pi]$, and $0 < g(0) < \infty$.

The log squared process is written as a volatility measure

$$Y_t = \log(X_t^2) = \eta + Z_t + u_t \tag{2}$$

where $\eta = \log \sigma^2 + E(\log e_t^2)$ and $u_t = \log e_t^2 - E(\log e_t^2)$. Here, $\{u_t\}$ is mean zero i.i.d. with variance σ_u^2. Autocovariances $R(j)$ of $\{Y_t\}$ is identical to that of $\{Z_t\}$ for $j \neq 0$. Other nonlinear measures such as squared or absolute process can be similarly dealt with.

Given Assumption 1, the spectral density of Y_t is the sum of the spectral density of Gaussian long memory process and that of non-Gaussian noise,

$$f_Y(\lambda) = \lambda^{-2d} g(\lambda) + \frac{\sigma_u^2}{2\pi} = \lambda^{-2d}(g(\lambda) + \frac{\sigma_u^2}{2\pi}\lambda^{2d}), \qquad \text{as } \lambda \to 0. \tag{3}$$

It is clearly pointed out in Deo and Hurvich (2001) that given the spectral representation (3), the LP estimator of Geweke and Porter-Hudak (1983; GPH) violates the Gaussian or Martingale assumption which the asymptotic theory is built upon in the long memory context (See also Bollerslev and Wright (2000)). In particular, due to the existence of the spectrum of non-Gaussian noise, the dominant term of the bias of GPH estimator behaves at the order of λ^{2d}. Then, GPH estimator suffers from significant negative bias.

In this paper, we make use of wavelet transform of the squared process and obtain Gaussian-approximate spectral representation by effectively removing the noise spectrum around zero frequency. Define the wavelet transform for Y_t

$$w_{jq} = 2^{j/2} \sum_t Y_t \psi(2^j t - q), \tag{4}$$

where t is suitably re-indexed so that the support of the wavelet is fully covered. For example, if the support of ψ is $[0, 1]$, then we let $t = i/n$, for $i = 1, 2, \cdots, n$. The integer valued j and q are scale and translation parameter, respectively, where $j = 0, 1, \cdots, J$, $q = 0, 1, \cdots, 2^j - 1$. The finest (maximum) scale is set to $n = 2^J$. It can be seen that the transformed series w_{jq} is simply a linear combination of Y_t over a local interval which is determined by j and q. The function ψ is a wavelet, which is a well localized function. For reference, see Hernandez and Weiss (1996) and Daubechies (1992).

We explicitly introduce the properties of the wavelet functions.

ASSUMPTION 2
(a) $\psi : \mathbb{R} \rightarrow \mathbb{R}$*such that* $\int_{-\infty}^{\infty} \psi(x)dx = 0$, $\int_{-\infty}^{\infty} |\psi(x)|dx < \infty$,
and $\int_{-\infty}^{\infty}(1 + x^2)\psi(x)dx < \infty$.

(b) $|\hat{\psi}(\lambda)| = \lambda^v b(\lambda)$, with $b(t\lambda)/b(\lambda) = 1$ for all t, *as* $\lambda \rightarrow 0$,
with v integer, $0 < b(0) < \infty$,

where $\hat{\psi}(\lambda)$ is Fourier transform of ψ, $\hat{\psi}(\lambda) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} \psi(x)e^{-i\lambda x}dx$.

The assumption 2(a) describes the wavelet function. By Assumption 2(a), the spectral density function of w_{jq} is well defined (Kato and Masry(1999)). It is not necessary in our analysis that ψ forms an orthonormal basis for L_2, though it is often the case in the wavelet literature. Next, assumption 2(b) models the spectral behavior of $\hat{\psi}(\lambda)$ around $\lambda = 0$. Integer-valued v is the number of vanishing moment of ψ in the sense that $\int_{-\infty}^{\infty} x^r\psi(x)dx = 0$ *for* $r = 0, 1, \cdots, v - 1$. The v vanishing moment is equivalent to saying that the first v spectral derivatives are zero at zero frequency, $\frac{d^r}{d\lambda^r}\hat{\psi}(\lambda) = 0$ at $\lambda = 0$ for $r = 0, 1, \cdots, v-1$, from the relation, $\frac{d^r}{d\lambda^r}\hat{\psi}(\lambda)\Big|_{\lambda=0} = (-i)^r \int_{\mathbb{R}} x^r\psi(x)dx$. This assumption is satisfied if ψ has a compact support and belongs to $C^v(\mathbb{R})$, where $C^v(\mathbb{R})$ is the class of the functions f on the real line $\mathbb{R}$ such that all the derivatives up to the order v exist, and the v-th derivative $f^{(v)}$ is continuous on $\mathbb{R}$. For example, Haar wavelet, defined as

$$\psi(x) = \begin{cases} 1 & 0 \leq x \leq 0.5 \\ -1 & 0.5 < x \leq 1, \end{cases} \tag{5}$$

satisfies assumption 2 with $v = 1$. Further, $|\hat{\psi}(\lambda)| = (\lambda/4)[\sin^2(\lambda/4)/(\lambda/4)^2]$, where we have $b(0) = 1/4$. Another example includes a class of spline wavelets. The first order spline wavelet, often called Franklin wavelet, has $v = 2$, and the spline wavelet of order 2 has $v = 3$. In general, the spline wavelets of order n has $n - 1$ vanishing moment (Hernandez and Weiss (1996)). Also, $b(\lambda)$ is assumed to be a slowly varying function at zero frequency. We consider Haar wavelet for the analysis and for the simulation in our paper.

Write the wavelet transform of Y_t under Haar wavelet system (5),

$$w_{jq} = \alpha_{jq} + \beta_{jq} \tag{6}$$

where $\alpha_{jq} = 2^{j/2} \sum_{t=1/n}^{1} Z_t \psi(2^j t - q)$ and $\beta_{jq} = 2^{j/2} \sum_{t=1/n}^{1} U_t \psi(2^j t - q)$.

First, we show that the spectral density of β_{jq} becomes zero at zero frequency. Let $R_\beta(m) = E\beta_{jq}\beta_{jq+m}$ be the autocovariances of the transformed series β_{jq} at scale j, and $f_\beta^{(j)}$ be the spectral density at scale j. The wavelet transform β_{jq} is a linear combination of i.i.d. noise process U_t. Using Haar wavelet in (5), it is simply the difference of local sums of U_t over $t \in [2^{-j}q, 2^{-j}(q+0.5)]$ and over $t \in (2^{-j}q, 2^{-j}(q+0.5)]$. Moreover, β_{jq} becomes 1-dependent process. Thus, it behaves as MA(p) process of i.i.d. series, where the order p is determined by j. As j increases, the width of the interval decreases, and at j equals to the finest scale J, β_{jq} becomes MA(1) process, that is to say, $\beta_{jq} = 2^{J/2}(U_{2^{-J}q} - U_{2^{-J}(q+1)})$. If we define the autocovariance $R_\beta(m) = E\beta_{jq}\beta_{jq+m}$, then it is clear that when $j = J$,

$$f_\beta^{(J)}(0) = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} R_\beta(m) = 0, \quad (7)$$

where $R_\beta(0) = 2\sigma_U^2$, $R_\beta(1) = -\sigma_U^2$, and $R_\beta(m) = 0$ for $m > 1$.

Next, we let $f_\alpha^{(j)}(\lambda)$ be the spectral density function of α_{jq} at scale j. Given Assumption 1, we directly obtain $f_\alpha^{(j)}(\lambda)$ as follows. We write autocovariances of the wavelet transforms at scale j

$$\begin{aligned} E\alpha_{jq}\alpha_{j\tau} &= 2^j \sum_t \sum_s EZ_t Z_s \psi(2^j t - q)\psi(2^j s - \tau) \\ &= 2^j \sum_t \sum_s \left[\int_{-\pi}^{\pi} f_Z(\lambda) e^{i(t-s)\lambda} d\lambda\right] \psi(2^j t - q)\psi(2^j s - \tau) \\ &= 2^{-j} \int_{-\pi}^{\pi} f_Z(\lambda) |\hat{\psi}(2^{-j}\lambda)|^2 e^{i2^{-j}(q-\tau)\lambda} d\lambda, \end{aligned}$$

where the third line follows from discrete Fourier transform of ψ and the change of variables. It is then inferred that at $j = J$,

$$f_\alpha^{(J)}(\lambda) = 2^{-J} f_Z(\lambda) |\hat{\psi}(2^{-J}\lambda)|^2, \qquad \lambda \in [-\pi, \pi] \quad (8)$$

Since the scale parameter is restricted to $j = J$, we suppress J in the expression of spectral density as $f_w^{(J)}(\lambda) = f_w(\lambda)$. Combining (7) and (8), we have Gaussian-approximate spectral representation of the wavelet transform w_{jq} around zero frequency

$$f_w(\lambda) = C_J \lambda^{-2(d-v)} g(\lambda) h(\lambda) \qquad as\ \lambda \to 0, \text{for } d \in (0, 0.5) \quad (9)$$

where $C = 2^{-J(1+2v)}$ and $h(\lambda) = b^2(\lambda)$.

Spectral representation (9) provides a basis for semiparametric estimation of d. The spectral density function $f_w(\lambda)$ behaves as $\lambda^{-2(d-v)}$ around zero frequency, thus $f_w(\lambda) \sim$

$\lambda^{-2(d-1)}$ when Haar wavelet is used. The functions $g(\lambda)$ and $h(\lambda)$ arise from short-run dependence in Z_t and from wavelet transform, respectively. It is noted that both are even, continuous on $[-\pi, \pi]$, and bounded away from zero at zero frequency. The statistical properties of g and h are useful to derive the asymptotic bias of LP estimator.

3. WAVELET-BASED LOG PERIODOGRAM ESTIMATOR

In this section, we construct wavelet-based LP estimator, and show the consistency. We define a periodogram for wavelet transform at scale J,

$$I_k^{(J)} \equiv I_k = \frac{1}{2\pi n} \sum_{q=0}^{2^J-1} |w_{jq} \exp(i\lambda_k q)|^2, \qquad k = 1, 2, \cdots, m \tag{10}$$

where $\lambda_k = 2\pi k/n$. The periodogram can be simply computed by using the relation, $I_k = A_k^2 + B_k^2$, where $A_k = (2\pi n)^{-1/2} \sum_{q=0}^{2^J-1} w_{jq} \cos(\lambda_k q)$ and $B_k = (2\pi n)^{-1/2} \sum_{q=0}^{2^J-1} w_{jq} \sin(\lambda_k q)$.

As standard in the long memory literature, suitable conditions on the rate of growth m are imposed for the frequencies $\lambda_k = 2\pi k/n$, where $k = 1, 2, \cdots, m$.

ASSUMPTION 3 : $m = m(n) \to \infty$, *and* $m/n \to 0$ *as* $n \to \infty$.

The positive integer m is restricted to increase at slower rate than n.

Under the spectral representation in (9) with $s(\lambda) = g(\lambda)h(\lambda)$, we write the LP regression as

$$\log I_k = \alpha + \beta X_k + \log(s(\lambda_k)/s(0)) + \varepsilon_k, \qquad k = 1, 2, \cdots m \tag{11}$$

where $\alpha = (\log C_J + \log(s(0))$, $\beta = (d-1)$, $X_k = -2\log(\lambda_k)$, and $\varepsilon_k = \log(I_k/f_k)$.

The term $\log(s(\lambda_k)/s(0)) = \log(g(\lambda_k)/g(0)) + \log(h(\lambda_k)/h(0))$ is dominant for the asymptotic bias. To get the explicit form of the asymptotic bias, we have Taylor expansion for $\log(s(\lambda_k)/s(0))$ at $\lambda = 0$,

$$\log \frac{s(\lambda_k)}{s(0)} = \frac{1}{2} \frac{s''(0)}{s(0)} \lambda_k^2 + O(\lambda_k^4).$$

We obtain the asymptotic bias, and variance.

THEOREM 1 : *Suppose Assumptions 1- 3 hold. Then,*

$$(a)\ E\widehat{d} - d = -\frac{2\pi^2}{9} \frac{s''(0)}{s(0)} \frac{m^2}{n^2}(1 + o(1)) + O(m^4/n^4) + O(\frac{\log^3 m}{m}).$$

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

Theorem 1 shows that wavelet-based estimator $\widehat{d}$ is consistent for $d \in (0, 0.5)$ in the L_2 sense. The variance takes the same form as in the stationary Gaussian case. The proof is basically adapted from Hurvich, Deo and Brodsky (1998; HDB) and Andrews and Guggenberger (2003;AG), as well as Robinson(1995). Further, we obtain the form of MSE

$$MSE(\widehat{d}) = \left[\frac{2\pi^2}{9}\frac{s''(0)}{s(0)}\right]^2 \frac{m^4}{n^4}(1+o(1)) + O\left(\frac{m^3 \log^3 m}{n^4}\right) + \frac{\pi^2}{24m}(1+o(1)). \tag{12}$$

Given the expression of MSE, we directly obtain the optimal m^*

$$m^* = \left[0.4634 \cdot \left(\frac{s(0)}{s''(0)}\right)^{2/5} n^{4/5}\right] \tag{13}$$

where $[z]$ denotes the closest integer to z. Both $s(0)$ and $s''(0)$ are unknown, though the function h depends the known wavelet function. Thus, only the rate of the optimal m^* is available. It follows that we have $MSE(\widehat{d}) = O(n^{-4/5})$. This is the same convergence rate as that of GPH estimator in the stationary case, which is developed by HDB.

Given the optimal rate of m above, the asymptotic normality can be applied to the wavelet-based estimator.

COROLLARY 1 : *Suppose Assumption 1-3 hold, and* $m = o(n^{4/5})$, *then*

$$m^{1/2}(\widehat{d} - d) \rightarrow_d N(0, \frac{\pi^2}{24}) \quad \text{as } n \rightarrow \infty.$$

The proof, briefly stated in the Appendix, follows from Robinson (1995), HDB, and AG.

4. SIMULATION STUDIES

We compare the finite sample performance of the wavelet-based regression estimator and GPH estimator. In data generating process (1), we let $\sigma = 1$, and consider ARFIMA$(1, d, 0)$ process for $\{Z_t\}$,

$$(1-\phi)(1-L)^d Z_t = \varepsilon_t \tag{14}$$

where ϕ is the autoregressive parameter, and ε_t is i.i.d. with variance σ_ε^2. The $I(d)$ process $\{Z_t\}_{t=1}^n$ is generated through $Z_t = \sum_{k=0}^{t-1} \frac{(d)_k}{k!} u_{t-k}$, and $(d)_k = d(d+1)\cdots(d+k-1)$, where $u_t \sim i.i.d.N(0,1)$. Sample size is set to $n = 1024$. The value of σ_ε^2 is set to 0.37

as in Deo and Hurvich (2001) and Breidt et al (1998). We only consider the combination of $(d, \phi) = (0.2, 0), (0.2, 0.3)$ and $(0.3, 0.6)$. Other combinations show qualitatively similar results. Regarding the number of frequencies m in the regression ( ), we include the values of m from $m = [n^{0.3}]$ to $m = [n^{0.8}]$, where $[x]$ denotes the integer part of x. Conventional choice of m in practice is $m = [n^{0.5}]$, while the optimal rate of wavelet-based estimator grows at the rate of $n^{0.8}$. The range of m includes the neighboring values of these two choices. Then, we can see the pattern of the bias and mean squared error (MSE) over the different values of m. For each value of m, one thousand iterations are conducted.

For wavelet-based estimator, we use Haar wavelet. The integer-valued scale j is set to the finest scale J for the transformed periodogram. For $n = 1024 = 2^{10}$, we set $J = 10$, which generates the transformed series, $\{w_j(q), q = 0, 1, \cdots, 2^{10} - 1\}$. For GPH estimator, we do not truncate the low frequency components, which is known to perform better than the truncated version of GPH estimator.(Deo and Hurvich (2001)).

The Figures 1 to 3 plot the bias and MSE of regression estimators for different values of d and ϕ. In Figure 1 with $(d, \phi) = (0.2, 0)$, GPH estimator shows significant negative bias and the magnitude of the bias nearly remains unchanged for all values of m. On the other hand, wavelet estimator initially shows large positive bias for very small values of m, but beyond a certain level of m, the bias significantly decreases and reaches to nearly half the bias of GPH estimator. It is expected that the wavelet estimator performs poorly for very small values of m, as the optimal rate of wavelet estimator grows at the rate of $n^{0.8}$. Except for such small values of m, the bias of wavelet estimator is significantly reduced as m gradually increases. In particular, we observe that wavelets works the best for the bias over $m \in [0.25n^{0.8}, 0.5n^{0.8}]$. Owing to the reduced bias, wavelet estimator has smaller MSE than GPH except for small values of m.

Next, Figure 2 and 3 plot the results in the case of $(d, \phi) = (0.2, 0.3)$ and $(0.3, 0.6)$, respectively. Basic pattern of the bias and MSE remains unaffected by allowing the short-run dependence. Wavelet estimator has significantly smaller bias than GPH estimator for most of the values m under consideration. Compared to the case of $\phi = 0$ in Figure 1, it seems that the positive short-run dependence rather helps reduce the negative bias of both estimator.

In sum, GPH estimator shows significant negative bias, which is not improved by different choice of fundamental frequencies. Thus, our proposed wavelet-based estimator is an effective way to reduce the bias for memory parameter estimation in LMSV models.

REFERENCES

Andrews, D.W.K., and Guggenberger, P. (2003) A bias-reduced log periodogram regression estimator for the long-memory parameter. *Econometrica*, 71, 675-712.

Bollerslev, T., and Wright, J. (2000) Semiparametric estimation of long-memory volatility dependencies: The role of high-frequency data. *Journal of Econometrics*, 98, 81-106.

Breidt, F., Crato, N., and de Lima, P. (1998) The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics*, 83, 325-348.

Daubechies I. (1992) *Ten Lectures on Wavelets*. CBS-NSF regional conferences in applied mathematics 61, society for industrial and applied mathematics.

Deo, R. and Hurvich, C.M. (2001) On the log periodogram regression estimator of the memory parameter in long memory stochastic volatility models. *Econometric Theory*, 17, 686-710.

Geweke, J., and Porter-Hudak, S. (1983) The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4, 221-237.

Hernandez, E., and Weiss, G. (1996) *A First Course on Wavelets*. CRC Press.

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

Kato, T., and Masry, E. (1999) On the spectral density of the wavelet transform of fractional Brownian motion. *Journal of Time Series Analysis*, 20, 559-563.

Lee, J. (2004) Wavelet transform for regression estimation of non-stationary fractional time series. Manuscript, department of economics, National University of Singapore.

Masry, E. (1993) The wavelet transform of stochastic processes with stationary increments and its application to fractional Brownian motion. *IEEE Transactions on Information Theory*, 39, 260-264.

Moulin, P. (1994) Wavelet Thresholding Techniques for power spectrum estimation. *IEEE Transactions on Signal Processing*, 42, 3126-3136.

Phillips, P.C.B. (1999) Unit root log periodogram regression. discussion paper, Cowles foundation, Yale University.

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

—— (2001) The memory of stochastic volatility models. *Journal of Econometrics*, 101, 195-218.

Shimotsu, K., and Phillips, P.C.B. (2002) Exact local whittle estimation of fractional integration. discussion paper, Cowles foundation, Yale University.

Tanaka, K. (1999) The nonstationary fractional unit root. *Econometric Theory*, 15, 549-582.

Tewfik, A.H., and Kim, M. (1992) Correlation structure of the discrete wavelet coefficients of fractional Brownian motion. *IEEE Transactions on Information Theory*, 38, 904-909.

Velasco, C. (1999) Non-stationary log periodogram regression. *Journal of Econometrics*, 91, 325-371.

APPENDIX

PROOF OF THEOREM 1: Let I, X, R, and ε denote $m \times 1$ column vectors whose k-th elements are $\log I_k$, X_k, $\log(s(\lambda_k)/s(0))$, and ε_k, respectively. As in AG, we write the regression equation in matrix form as $\log I = (\log C_J + \log s(0))1_m + X\beta + R + \varepsilon$. Let $Z = X - 1_m\overline{X}$ with $\overline{X} = (X'1_m)/m$, we write

$$\log I = (\log C_J + \log s(0) + \overline{X}d)1_m + Z\beta + R + \varepsilon \tag{A1}$$

The bias term can be written as $E\widehat{d} - d = (Z'Z)^{-1}Z'(R + \varepsilon)$.

The proof consists of the three parts: (a) $Z'Z$, (b) $Z'R$, and (c) $Z'E(\varepsilon)$. First, note that $X_k = -2\log\lambda_k$, then from HDB (page22), we have $Z'Z = 4m(1 + o(1))$. Next, we write

$$Z'R = \frac{1}{2}\frac{s''(0)}{s(0)}Z'\lambda_k^2 + \textstyle\sum_{k=1}^{m}(X_k - \overline{X})O(\lambda_k^4). \tag{A2}$$

The first term in (A2) is written as

$$\begin{aligned}
\frac{1}{2}\frac{s''(0)}{s(0)}Z'\lambda_k^2 &= \frac{1}{2}\frac{s''(0)}{s(0)}Z'(\frac{k}{m})^2(\frac{2\pi m}{n})^2 \\
&= -\frac{2}{9}\frac{s''(0)}{s(0)}(\frac{2\pi m}{n})^2 m(1 + o(1)) \\
&= -\frac{8\pi^2}{9}\frac{s''(0)}{s(0)}\frac{m^3}{n^2}(1 + o(1)),
\end{aligned}$$

where the first line follows from $\lambda_k = 2\pi k/n$, and the second line from $Z'(k/m)^2 = -[4/9]m(1 + o(1))$ by Lemma 2(c) in AG.

The order of magnitude for the second term in (A2) follows from AG or HDB (page 38) that $\sum_{k=1}^{m}(X_k - \overline{X})O(\lambda_k^4) = O(m^5/n^4)$. Lastly, under the Gaussianity, we directly apply the Lemma 8 in HDB or Lemma 2(f) in AG. Then, we have $Z'E(\varepsilon) = O(\log^3 m)$. The proof of the variance term comes directly from HDB (proof of Theorem 1), then we omit it. This completes the proof.

PROOF OF COROLLARY 1: The proof of asymptotic normality directly follows from that of Theorem 2 in HDB or of Theorem 2 in AG, which are based on Robinson (1995). Below we only verify the Theorem 2 in Robinson (1995), which is essential to show the asymptotic normality.

Write discrete Fourier transform of transformed series $\{w_{Jq}\}$ for fixed J, and its normalized version as

$$u(\lambda_k) = (2\pi n)^{-1/2} \sum_{q=0}^{2^J} w_{jq} \exp(i\lambda_k q), \text{ and } v(\lambda_k) = u(\lambda_k)/f^{1/2}. \tag{A3}$$

where the normalization is made by using $f^{1/2}$ rather than $C_j \lambda^{-2d} |\hat{\psi}(\lambda)|^2$. It follows that

$$\begin{aligned} & E\{u(\lambda_k)\overline{u}(\lambda_k)\} \\ = \; & (2\pi n)^{-1} \sum_{q=0}^{2^J-1} \sum_{r=0}^{2^J-1} E(w_{Jq} w_{Jr}) \exp\{i(q-r)\lambda_k\} \\ = \; & \int_{-\pi}^{\pi} f(\lambda)(2\pi n)^{-1} \sum_{q=0}^{2^J-1} \sum_{r=0}^{2^J-1} \exp\{-i(q-r)\lambda\} \exp\{i(q-r)\lambda_k\} d\lambda \\ = \; & \int_{-\pi}^{\pi} f(\lambda) K(\lambda_k - \lambda) d\lambda. \end{aligned}$$

where $K(\lambda) = (2\pi n)^{-1} \sum_{q=0}^{2^J-1} \sum_{r=0}^{2^J-1} \exp\{i(q-r)\lambda\}$. Then, we obtain the same expression as that of (4.1) in Robinson (1995). Thus, the proof of Theorem 2 in Robinson (1995) is applied to have

$$E\{v(\lambda_k)\overline{v}(\lambda_k)\} = 1 + O(\frac{\log k}{k}).$$

By similar reasoning, we also obtain

$$\begin{aligned} E\{u(\lambda_k)u(\lambda_k)\} &= \int_{-\pi}^{\pi} f(\lambda) D(\lambda_k - \lambda)(\lambda + \lambda_k) d\lambda, \\ E\{u(\lambda_k)\overline{u}(\lambda_s)\} &= \int_{-\pi}^{\pi} f(\lambda) D(\lambda_k - \lambda) D(\lambda - \lambda_s) d\lambda, \\ E\{u(\lambda_k)u(\lambda_s)\} &= \int_{-\pi}^{\pi} f(\lambda) D(\lambda_k - \lambda) D(\lambda + \lambda_s) d\lambda, \end{aligned}$$

where $D(\lambda) = (2\pi n)^{-1} \sum_{q=0}^{2^J-1} \exp(iq\lambda)$. Then, again by the proof of Robinson (1995), we verify that $E\{u(\lambda_k)u(\lambda_k)\} = O(k/\log k)$, $E\{u(\lambda_k)\overline{u}(\lambda_s)\} = O(k/\log s)$, and $E\{u(\lambda_k)u(\lambda_s)\} = O(k/\log s)$.

Given the above results, the proof of Theorem 2 in HDB or of Theorem 2 in AG follows. This completes the proof.

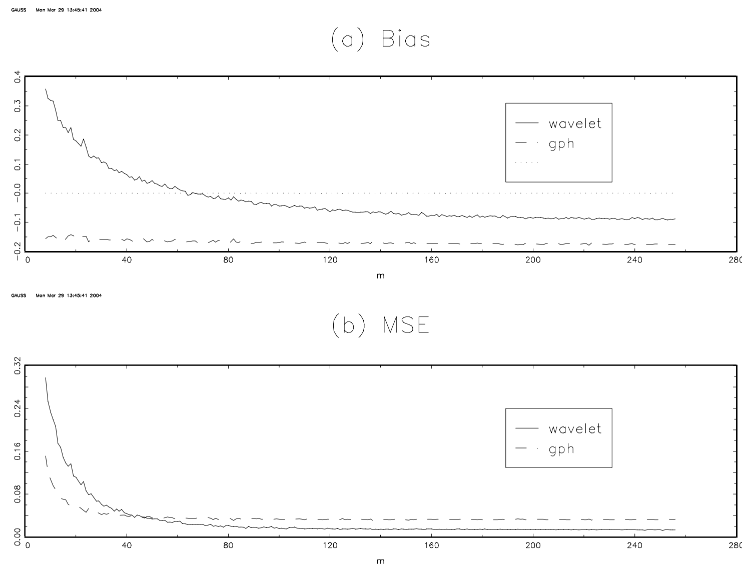

Figure 1:

Bias and MSE of Wavelet and GPH estimators:
$(d, \phi) = (0.2, 0), n = 1024.$

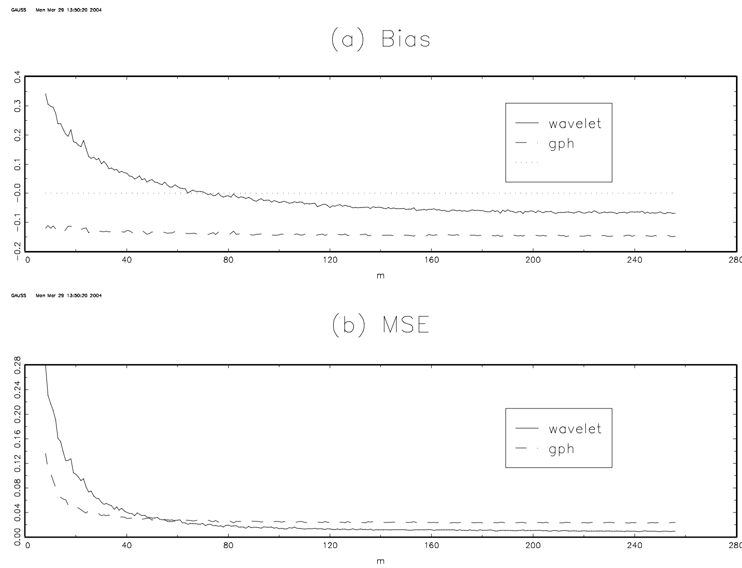

Figure 2:

Bias and MSE of Wavelet and GPH estimators:
$(d, \phi) = (0.2, 0.3), n = 1024,$

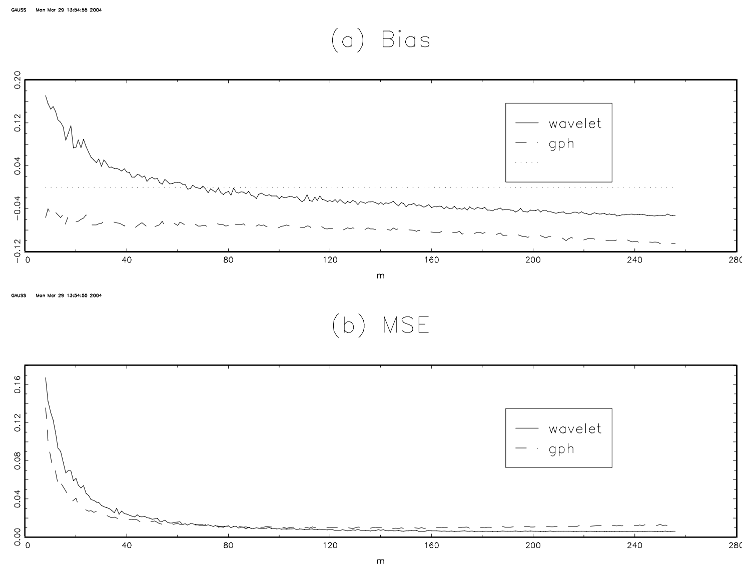

Figure 3:

Bias and MSE of Wavelet and GPH estimators:
$(d, \phi) = (0.3, 0.6), n = 1024$