On the Spectral Density of Fractional Ornstein-Uhlenbeck Processes*

Shuping Shi[†], Jun Yu[††], Chen Zhang[††]

[†] Macquarie University

[††]Singapore Management University

October 24, 2023

Abstract

This paper introduces a novel and easy-to-implement method for accurately approximating the spectral density of discretely sampled fractional Ornstein-Uhlenbeck (fOU) processes. The method offers a substantial reduction in approximation error, particularly within the rough region of the fractional parameter $H \in (0, 0.5)$. This approximate spectral density has the potential to enhance the performance of estimation methods and hypothesis testing that make use of spectral densities. We introduce the approximate Whittle maximum likelihood (AWML) method for discretely sampled fOU processes, utilising the approximate spectral density, and demonstrate that the AWML estimator exhibits properties of consistency and asymptotic normality when $H \in (0, 1)$, akin to the conventional Whittle maximum likelihood method. Through extensive simulation studies, we show that AWML outperforms existing methods in terms of both estimation accuracy and computational speed in finite samples. We then apply the AWML method to the log realized volatility of 40 financial assets. Our empirical findings reveal that the estimated Hurst parameters for these assets fall within the range of 0.10 to 0.23, indicating a rough volatility dynamic.

JEL classification: C13, C22, G10

Keywords: Fractional Brownian motion; Fractional Ornstein-Uhlenbeck process; Spectral density; Paxson approximation; Whittle maximum likelihood; Realized volatility

*We would like to thank Morten Nielsen and Tetsuya Takabatake for helpful comments. Shuping Shi, Department of Economics, Macquarie University, Sydney, Australia. Email: shuping.shi@mq.edu.au. Jun Yu, School of Economics and Lee Kong Chian School of Business, Singapore Management University, 90 Stamford Road, Singapore 178903. Email: yujun@smu.edu.sg. Chen Zhang, School of Economics, Singapore Management University, 90 Stamford Road, Singapore 178903. Email: chenzhang@smu.edu.sg. Yu and Zhang would like to acknowledge the research support provided by the Ministry of Education, Singapore, under its Academic Research Fund (AcRF) Tier 2 (Award Number MOE-T2EP402A20-0002).

1 Introduction

The fractional Ornstein-Uhlenbeck (fOU) process is a stochastic process widely applied across diverse fields such as in finance, physics, biology, and more to model diverse phenomena. An essential parameter of the fOU process is the Hurst exponent (H), which determines the degree of memory in the process. When $H = 0.5$, the fOU process becomes the standard Ornstein-Uhlenbeck (OU) process. For $H > 0.5$, the process is characterized by long memory, while for $H < 0.5$, the process exhibits rough dynamics. In finance, the fOU process was recently used to model the log volatility of financial assets. See, e.g., Gatheral et al. (2018), Wang et al. (2023), Fukasawa et al. (2022), Bennedsen et al. (2022), Bolko et al. (2023). Wang et al. (2023) have found that the fOU process yields more accurate forecasts than some popular discrete-time models.

Spectral density serves as a potent tool for scrutinizing the frequency characteristics inherent in time series data. Its usage extends to the detection of periodic patterns, trends, and cycles, as extensively demonstrated in previous studies (Burnside (1998), Messina et al. (2009), McElroy and Roy (2017), Angeletos et al. (2020)). Furthermore, spectral density plays a pivotal role in parameter estimation (Whittle (1954), Christiano and Vigfusson (2003)) and hypothesis testing (McElroy (2016)).

However, how to calculate the spectral density for discretely sampled fOU processes presents a challenge due to the requirement for calculating an infinite summation, which poses practical numerical difficulties.[1] Hult (2003) uses the truncation method to approximate the infinite summation in the spectral density and estimate the parameters using the truncation approximation of the spectral density. The order of approximation error of the truncation method is $O(K^{-2H})$, where K is cut-off point in the infinite summation. When H is greater than 0.5, which is the case considered in Hult (2003), with a moderately large K, the truncation works well. Unfortunately, when H takes a very small value (e.g., 0.2 or less), which is found empirically relevant in the rough volatility literature (Fukasawa et al., 2022, Bolko et al., 2023), the truncation method often suffers from diminished accuracy, even when a very large cut-off point K is used.

To surmount this issue, drawing inspiration from Paxson (1997) and Fukasawa and Takabatake (2019) where the spectral density of the fractional Gaussian noise (fGn) is approximated, we introduce a novel method for approximating the spectral density of the fOU process. Our proposed method, named the modified Paxson approximation, capitalizes on the eventual monotonicity of the summation terms to derive both lower and upper bounds for the spectral density. We then take the average of these two bounds to estimate the spectral density. The order of approximation error is shown to be $O(K^{-2H-1})$, which is lower than that of the truncation method. This improvement is especially important when H is close to zero because, as H converges to zero, the order of approximation error are $O(1)$ and $O(K^{-1})$ for the two methods, respectively. That is, when $H \to 0$, the approximation error of the truncation method is

[1] Although the fOU processes is related to an ARFIMA$(1,d,0)$ whose spectral density is trivial to calculate, the spectral density of fOU is different from that of ARFIMA$(1,d,0)$.

a constant regardless of the value of K, while the approximation error of the new method diminishes to zero as the K increases. Furthermore, the results imply that for a fixed H, the new method requires a much smaller K to achieve the same level of approximation accuracy and hence lead to improved computational efficiency. Since the modified Paxson method excels in both approximation accuracy and computational efficiency, it is well-suited for a wide range of practical applications, including the frequency-domain maximum likelihood (ML) estimation.

We demonstrate the practical value of the modified Paxson method for estimating parameters in fOU processes. Specifically, we introduce an approximate Whittle maximum likelihood (AWML) method. This method involves replacing the true spectral density of fOU with the approximate spectral density in the frequency-domain Whittle ML approach. We first demonstrate that the standard Whittle ML method exhibits consistency and asymptotic normality when applied to discrete-time observations generated from a stationary fOU process for $H \in (0,1)$, encompassing both long memory and anti-persistent behaviors. Then we show that the AWML estimator exhibits asymptotic equivalence to the Whittle ML estimator, inheriting all of its advantageous asymptotic properties.

Through simulation studies, we compare the finite sample performance of the AWML method with two existing methods: the maximum composite likelihooe (MCL) method (Bennedsen et al., 2022) and the change-of-frequency (CoF) approach (Wang et al., 2023).[2] The AWML method proves to be the most accurate of the three, with CoF being the least accurate. Furthermore, the AWML method is computationally more efficient than MCL. We apply the proposed estimation method for fOU to the log realized volatility of 40 financial assets. Our estimation results suggest that volatility always exhibits rough dynamics, with estimated H ranging from 0.1 to 0.25.

The rest of the paper is organized as follows. In Section 2, we present the fOU process and the spectral density of discretely sampled fOU. Section 3 introduces the truncation based approximation method for the spectral density, develops the modified Paxson method, and evaluates their approximation accuracy. Section 4 introduces the AWML method based on the modified Paxson approximate spectral density, along with MCL and CoF. We develop the asymptotic properties of the AWML method and examine its finite sample performance relative to alternative estimation methods. We present the empirical results in Section 5. The proofs of all the theoretical results are provided in the Appendix. Throughout the paper, the notation $A_K \preceq B_K$ signifies that A_K/B_K is either $O(1)$ or $o(1)$ as $K \to \infty$ (either $O_p(1)$ or $o_p(1)$ if A_K/B_K is random).

[2]Bennedsen et al. (2022) develop the MCL method to estimate not only the fOU process but also some other stationary processes.

2 Fractional OU Process

The fOU process $\{y_t : t \in \mathbb{R}\}$ is the stochastic process given by

$$dy_t = \kappa(\mu - y_t)\,dt + \sigma dB_t^H \text{ with } y_0 = O_p(1),$$

where $\sigma > 0$, μ is a constant, and B_s^H is an fBm process with $H \in (0,1)$. The fBm is a Gaussian process with mean zero and autocovariance given by

$$Cov\left(B_t^H, B_s^H\right) = \frac{1}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right), \ \ \forall t,s \in (-\infty, +\infty). \tag{1}$$

It includes the standard Brownian motion (Bm) process as a special case with $H = 0.5$. Just like the Bm, the fBm is nonstationary for all $H \in (0,1)$. However, the first difference or increment of fBm, which is fGn, is always stationary. The fBm has sample paths smoother (rougher) than the standard Bm when $H > 0.5$ ($H < 0.5$). The autocovariances of fGn are not summable when $H > 0.5$ (i.e. exhibiting long memory) but sum to zero when $H < 0.5$ (i.e. exhibiting anti-persistence).

The fOU process reduces to the traditional OU process when $H = 0.5$ and to the fBm when $\kappa = 0$. When $\kappa > 0$, fOU is stationary. When $\kappa < 0$, fOU is explosive. In this paper, we assume that $\kappa > 0$. The fOU process is (locally) Hölder continuous of order $H - \varepsilon$ (Gatheral et al., 2018, Gehringer and Li, 2020). The fOU process has a unique path-wise solution:

$$y_t = e^{-\kappa t} y_0 + \left(1 - e^{-\kappa t}\right)\mu + \int_{-\infty}^{t} \sigma e^{-\kappa(t-u)}\, \mathrm{d}B_u^H, \tag{2}$$

where $E(y_t) = \mu$ and $Var(y_t) = \sigma^2 \kappa^{-2H} H \Gamma(2H)$ with $\Gamma(.)$ being the Gamma function. See Hu and Nualart (2010a). When $\kappa > 0$, the autocovariance of fOU is given by

$$\gamma_k = \frac{\sigma^2}{2\kappa^{2H}} \left(\frac{1}{2}\int_{-\infty}^{\infty} e^{-|s|} |\kappa k + s|^{2H}\, \mathrm{d}s - |\kappa k|^{2H}\right), \tag{3}$$

and its spectral density is

$$f_y\left(\lambda; \kappa, H, \sigma^2\right) = \frac{\sigma^2}{2\pi} C(H) |\lambda|^{1-2H} (\kappa^2 + \lambda^2)^{-1} \text{ for } \lambda \in (-\infty, \infty), \tag{4}$$

where $C(H) = \Gamma(2H+1)\sin(\pi H)$. See Garnier and Sølna (2018) for the autocovariance and Hult (2003) for the spectral density.

Assumption 1 *The model parameters* $(\kappa, H, \sigma^2) \in \Theta$, *where* Θ *is a compact subset in* $\mathbb{R}^+ \times (0,1) \times \mathbb{R}^+$.

Lemma 2.1 *Under **Assumption** 1, the spectral density function* $f_y\left(\lambda; \kappa, H, \sigma^2\right)$ *is uniquely determined by the parameters* (κ, H, σ^2).

The proof of this lemma can be found in Appendix A. The identification result stands in stark contrast to the lack-of-identification problem observed in the ARFIMA$(1,d,0)$ model as discussed in Li et al. (2023).

Empirical observations are often sampled at a fixed sampling interval (denoted by Δ). The discrete-time observations from the fOU process are denoted by $\left\{y_{j\Delta}\right\}_{j=1}^{n}$, where n is the total number of observations. Hence, $T = n\Delta$ is the time span. The spectral density of discretely sampled fOU is provided by Hult (2003) and takes the form of

$$f_y^{\Delta}\left(\lambda;\kappa,H,\sigma^2\right) = \frac{\sigma^2}{2\pi}C(H)\Delta^{2H}\sum_{k=-\infty}^{\infty}\frac{|\lambda+2\pi k|^{1-2H}}{(\kappa\Delta)^2+(\lambda+2\pi k)^2} \text{ for } \lambda\in(0,2\pi). \tag{5}$$

Moreover, Hult (2003) shows that

$$E(y_{j\Delta}y_0) = \int_{-\infty}^{\infty} e^{ij\Delta\lambda} f_y\left(\lambda;\kappa,H,\sigma^2\right)\mathrm{d}\lambda = \int_0^{2\pi} e^{ij\lambda} f_y^{\Delta}\left(\lambda;\kappa,H,\sigma^2\right)\mathrm{d}\lambda, \tag{6}$$

which establishes the relationship between $f_y(\lambda;\theta)$ and $f_y^{\Delta}(\lambda;\theta)$. Therefore, by the Fourier inversion theorem, the uniqueness of $f_y(\lambda;\theta)$ implies the uniqueness of $f_y^{\Delta}(\lambda;\theta)$.

3 Approximation Methods

The spectral density of the discretely sampled fOU process (5) necessitates the computation of an infinite summation, and the truncation method emerges as a natural choice for its evaluation. In this section, we introduce the modified Paxson approximation for the spectral density and explore the approximation accuracy of both the truncation approach and the modified Paxson approximation.

Let $\beta = (\kappa, H)$ and $\theta = (\kappa, H, \sigma^2)$. The spectral density of discretely sampled fOU (5) involves the infinite summation

$$S(\lambda;\beta) \equiv \sum_{k=-\infty}^{\infty}\frac{|\lambda+2\pi k|^{1-2H}}{(\kappa\Delta)^2+(\lambda+2\pi k)^2},$$

which can be rewritten as

$$\begin{aligned} S(\lambda;\beta) &= \sum_{k=-\infty}^{-1}\frac{(-\lambda-2\pi k)^{1-2H}}{(\Delta\kappa)^2+(\lambda+2\pi k)^2} + \frac{\lambda^{1-2H}}{(\Delta\kappa)^2+\lambda^2} + \sum_{k=1}^{\infty}\frac{(\lambda+2\pi k)^{1-2H}}{(\Delta\kappa)^2+(\lambda+2\pi k)^2} \\ &= \sum_{k=1}^{\infty}\frac{(2\pi k-\lambda)^{1-2H}}{(\Delta\kappa)^2+(2\pi k-\lambda)^2} + \frac{\lambda^{1-2H}}{(\Delta\kappa)^2+\lambda^2} + \sum_{k=1}^{\infty}\frac{(\lambda+2\pi k)^{1-2H}}{(\Delta\kappa)^2+(\lambda+2\pi k)^2}, \\ &= \sum_{k=1}^{\infty} Q_{1,k}(\lambda) + Q_0(\lambda) + \sum_{k=1}^{\infty} Q_{2,k}(\lambda) \end{aligned}$$

$$= \sum_{k=1}^{K} Q_{1,k}(\lambda) + \sum_{k=K+1}^{\infty} Q_{1,k}(\lambda) + Q_0(\lambda) + \sum_{k=1}^{K} Q_{2,k}(\lambda) + \sum_{k=K+1}^{\infty} Q_{2,k}(\lambda) \quad (7)$$

where $Q_0(\lambda) \equiv \frac{\lambda^{1-2H}}{(\Delta\kappa)^2+\lambda^2}$, $Q_{1,k}(\lambda) \equiv \frac{(2\pi k-\lambda)^{1-2H}}{(\Delta\kappa)^2+(2\pi k-\lambda)^2}$, and $Q_{2,k}(\lambda) \equiv \frac{(\lambda+2\pi k)^{1-2H}}{(\Delta\kappa)^2+(\lambda+2\pi k)^2}$.

The approximate spectral density based on the truncation method, denoted by $\bar{f}_y^{\Delta}(\lambda;\theta)$, takes the form of

$$\bar{f}_y^{\Delta}(\lambda;\theta) = \frac{\sigma^2}{2\pi} C(H) \Delta^{2H} \bar{S}(\lambda;\beta), \quad (8)$$

with

$$\bar{S}(\lambda;\beta) = \sum_{k=1}^{K} Q_{1,k}(\lambda) + Q_0(\lambda) + \sum_{k=1}^{K} Q_{2,k}(\lambda).$$

Thus, the approximation eror of the truncation method is

$$\left| f_y^{\Delta}(\lambda;\theta) - \bar{f}_y^{\Delta}(\lambda;\theta) \right| = \frac{\sigma^2}{2\pi} C(H) \Delta^{2H} \left[\sum_{k=K+1}^{\infty} Q_{1,k}(\lambda) + \sum_{k=K+1}^{\infty} Q_{2,k}(\lambda) \right]. \quad (9)$$

Before we introduce the modified Paxson method, let us first briefly review the Paxson method that has been employed to approximate the spectral density of fGn. The spectral density of fGn, denoted by $f_x^{\Delta}(\lambda)$, takes the form of

$$f_x^{\Delta}(\lambda) = R(H,\lambda,\sigma) \left(|\lambda|^{-(2H+1)} + \sum_{k=1}^{K} Q_k(\lambda) + \sum_{k=K+1}^{\infty} Q_k(\lambda) \right),$$

where $R(H,\lambda,\sigma) = \frac{\sigma^2}{\pi}\Delta^{2H}(1-\cos\lambda)C(H)$, K is a pre-specified integer and $Q_k(\lambda) = (2\pi k+\lambda)^{-(2H+1)} + (2\pi k-\lambda)^{-(2H+1)}$. The monotonicity property of $Q_k(\lambda)$ in k ensures that $\sum_{k=K+1}^{\infty} Q_k(\lambda)$ has a upper bound $\int_K^{\infty} Q_k(\lambda)\,\mathrm{d}k$ and a lower bound $\int_{K+1}^{\infty} Q_k(\lambda)\,\mathrm{d}k$, that is,

$$q_K(\lambda) := \int_K^{\infty} Q_k(\lambda)\,\mathrm{d}k \geq \sum_{k=K+1}^{\infty} Q_k(\lambda) \geq \int_{K+1}^{\infty} Q_k(\lambda)\,\mathrm{d}k := q_{K+1}(\lambda).$$

Therefore, for any K, $f_x^{\Delta}(\lambda)$ is bounded from up and below, respectively, by

$$R(H,\lambda,\sigma) \left\{ |\lambda|^{-(2H+1)} + \sum_{k=1}^{K} Q_k(\lambda) + q_K(\lambda) \right\} \text{ and } R(H,\lambda,\sigma) \left\{ |\lambda|^{-(2H+1)} + \sum_{k=1}^{K} Q_k(\lambda) + q_{K+1}(\lambda) \right\}. \quad (10)$$

Paxson (1997) proposes an approximate spectral density by taking the average of the lower and upper bounds in (10). The Paxson approximate spectral density of fGn is denoted by $\tilde{f}_x^{\Delta}(\lambda)$ and given by

$$\tilde{f}_x^{\Delta}(\lambda) = 2(1-\cos\lambda)\frac{\sigma^2}{2\pi}C(H)\Delta^{2H}\left\{|\lambda|^{-(2H+1)} + \sum_{k=1}^{K} Q_k(\lambda) + \frac{1}{2}\left[q_K(\lambda) + q_{K+1}(\lambda)\right]\right\}. \tag{11}$$

By definition, the quantity

$$q_K(\lambda) = \frac{1}{4\pi H}\left[(2\pi K+\lambda)^{-2H} + (2\pi K-\lambda)^{-2H}\right].$$

Both the upper bound and the lower bound, therefore, possess closed-form expressions that greatly simplify the calculation. Furthermore, even when K is set to a small value, these bounds remain remarkably close to each other. As a result, the Paxson formula stands out as an exceptionally accurate method for approximating the spectral density of fGn while incurring low computational cost. Fukasawa and Takabatake (2019) employs this approximate spectral density to construct an approximate Whittle ML estimation method for H in fGn and find that the estimator has satisfactory finite sample performance.

For fOU processes, the Paxson method encounters two complications. Firstly, neither $Q_{1,k}(\lambda)$ nor $Q_{2,k}(\lambda)$ in (7) exhibits monotonic behavior with respect to k. Fortunately, we can identify a critical cut-off value, denoted by K^*, such that for all $k \geq K^*$, $Q_{1,k}(\lambda)$ and $Q_{2,k}(\lambda)$ are monotonic in k. Secondly, the upper and lower bounds for both $Q_{1,k}(\lambda)$ and $Q_{2,k}(\lambda)$ are represented as integrals, and regrettably, their analytical solutions remain unavailable. To facilitate the calculation, we estabilish alternative upper and lower bounds for each of these quantities. These results are reported in Lemma 3.1.

Lemma 3.1 *(1) The quantity $Q_{1,k}(\lambda)$ monotonically decreases in k when $k > \frac{\Delta\kappa+\lambda}{2\pi}$; (2) The quantity $Q_{2,k}(\lambda)$ monotonically decreases in k when $k > \frac{\Delta\kappa-\lambda}{2\pi}$; (3) For $K > \frac{\Delta\kappa+\lambda}{2\pi}$ and for any β and $\lambda \in (0, 2\pi)$,*

$$a_K(\lambda) \leq \int_{K+1}^{\infty} Q_{1,k}(\lambda)\,\mathrm{d}k \leq \sum_{k=K+1}^{\infty} Q_{1,k}(\lambda) \leq \int_{K}^{\infty} Q_{1,k}(\lambda)\,\mathrm{d}k \leq \frac{(2\pi K-\lambda)^{-2H}}{4\pi H}, \tag{12}$$

$$b_K(\lambda) \leq \int_{K+1}^{\infty} Q_{2,k}(\lambda)\,\mathrm{d}k \leq \sum_{k=K+1}^{\infty} Q_{2,k}(\lambda) \leq \int_{K}^{\infty} Q_{2,k}(\lambda)\,\mathrm{d}k \leq \frac{(2\pi K+\lambda)^{-2H}}{4\pi H}, \tag{13}$$

where

$$a_K(\lambda) = \frac{[2\pi(K+1)-\lambda]^{-2H}}{4\pi}\left\{\frac{1}{H} - \frac{(\Delta\kappa)^2}{(1+H)[2\pi(K+1)-\lambda]^2}\right\}, \tag{14}$$

$$b_K(\lambda) = \frac{[2\pi(K+1)+\lambda]^{-2H}}{4\pi}\left\{\frac{1}{H} - \frac{(\Delta\kappa)^2}{(1+H)[2\pi(K+1)+\lambda]^2}\right\}. \tag{15}$$

Remark 3.1 *According to Lemma 3.1(1)-(2), while $Q_{1,k}(\lambda)$ and $Q_{2,k}(\lambda)$ do not exhibit monotonic behavior with respect to k across the entire range of $k \in [1,\infty)$, they do become monotonic in k when*

$k > \frac{\Delta\kappa+\lambda}{2\pi}$. *This requirement can be easily satisfied in practical relevant cases. For example, if* $\lambda = 2\pi$ *(the upper bound for* λ*), the threshold is only* 2.59 *when* κ *is as large as* 2520 (120) *for daily (monthly) data. Since the estimated* κ *values reported in the volatility literature are typically much smaller than* 100*,* $Q_{1,k}(\lambda)$ *and* $Q_{2,k}(\lambda)$ *are monotonic in* k *for* $k > 2$*.*

Remark 3.2 *From Lemma 3.1(3), the natural upper and lower bounds for both* $Q_{1,k}(\lambda)$ *and* $Q_{2,k}(\lambda)$ *involve integrals and lack closed-form expressions. Conversely, the alternative bounds are in close forms and easy to compute, namely* $\left(a_K(\lambda), \frac{(2\pi K-\lambda)^{-2H}}{4\pi H}\right)$ *for* $\sum_{k=K+1}^{\infty} Q_{1,k}(\lambda)$ *and* $\left(b_K(\lambda), \frac{(2\pi K+\lambda)^{-2H}}{4\pi H}\right)$ *for* $\sum_{k=K+1}^{\infty} Q_{2,k}(\lambda)$*.*

Lemma 3.2 *The lower and upper bounds for the spectral density of the discretely sampled fOU are as follows:*

$$\frac{\sigma^2}{2\pi} C(H) \Delta^{2H} S_L(\lambda;\beta) \leq f_y^{\Delta}(\lambda) \leq \frac{\sigma^2}{2\pi} C(H) \Delta^{2H} S_U(\lambda;\beta). \tag{16}$$

with

$$S_L(\lambda;\beta) \equiv \sum_{k=1}^{K} Q_{1,k}(\lambda) + \sum_{k=1}^{K} Q_{2,k}(\lambda) + \frac{\lambda^{1-2H}}{(\Delta\kappa)^2+\lambda^2} + a_K(\lambda) + b_K(\lambda), \tag{17}$$

and

$$S_U(\lambda;\beta) \equiv \sum_{k=1}^{K} Q_{1,k}(\lambda) + \sum_{k=1}^{K} Q_{2,k}(\lambda) + \frac{\lambda^{1-2H}}{(\Delta\kappa)^2+\lambda^2} + \frac{1}{4\pi H}\left[(2\pi K-\lambda)^{-2H} + (2\pi K+\lambda)^{-2H}\right]. \tag{18}$$

Remark 3.3 *Although results in (16) remain valid for small values of* K *(as discussed in Remark 3.1), the gap between the lower and upper bounds can be substantial when* K *is small. To ensure the approximation accuracy of the modified Paxson method, which entails averaging the lower and upper bounds of the spectral density, a moderate* K *may be required.*

The modified Paxson approximation of the spectral density, denoted by $\tilde{f}_y^{\Delta}(\lambda)$, is obtained by averaging these bounds. The word 'modified' reflects the adjustment made to accommodate the complexities associated with fOU. The analytical expression of $\tilde{f}_y^{\Delta}(\lambda)$ and the order of its approximation error is provided in Theorem 3.1.

Theorem 3.1 *Let the modified Paxson approximate spectral density of the discretely sampled fOU be*

$$\begin{aligned} \tilde{f}_y^{\Delta}(\lambda) &= \frac{\sigma^2}{2\pi} C(H) \Delta^{2H} \left[\sum_{k=1}^{K} Q_{1,k}(\lambda) + \sum_{k=1}^{K} Q_{2,k}(\lambda) + \frac{\lambda^{1-2H}}{(\Delta\kappa)^2+\lambda^2} \right. \\ &\quad \left. + \frac{1}{2} a_K(\lambda) + \frac{1}{2} b_K(\lambda) + c_K(\lambda)\right], \text{ for } \lambda \in (0, 2\pi), \end{aligned} \tag{19}$$

where $K > \frac{\Delta\kappa+\lambda}{2\pi}$, *and*

$$c_K(\lambda) = \frac{(2\pi K-\lambda)^{-2H}}{8\pi H} + \frac{(2\pi K+\lambda)^{-2H}}{8\pi H}.$$

Under ***Assumption 1****, for any* $H \in (0,1)$ *and* $\lambda \in (0,2\pi)$*, the approximation errors of the truncation method and the modified Paxson method are, respectively,* $O\left(K^{-2H}\right)$ *and* $O\left(K^{-2H-1}\right)$*. That is, as* $K \to \infty$,

$$\sup_{\lambda\in(0,2\pi)} |f_y^{\Delta}(\lambda) - \bar{f}_y^{\Delta}(\lambda)| = O\left(K^{-2H}\right); \tag{20}$$

$$\sup_{\lambda\in(0,2\pi)} |f_y^{\Delta}(\lambda) - \tilde{f}_y^{\Delta}(\lambda)| = O\left(K^{-2H-1}\right). \tag{21}$$

Remark 3.4 *According to Theorem 3.1, the approximation accuracy of the truncation approach depends on both the cut-off point* K *and the fractional parameter* H*. When* H *assumes a relatively large value, the approximate error converges rapidly to zero as* K *becomes larger. This finding aligns with the results presented in Hult (2003), which highlight the efficacy of the truncation method when* $H > 0.5$*. However, in cases where the fractional parameter* H *is close to zero, a scenario often observed in empirical studies with fOU (Wang et al., 2023, Fukasawa et al., 2022, Bolko et al., 2023), the order of the approximation error becomes relatively large. Indeed, when* $H \to 0$*, the approximate error of the truncation method is* $O(1)$*. Consequently, the use of the truncation method in practice may lead to imprecise estimations by the AWML method.*

Remark 3.5 *Similar to the truncation method, the accuracy of the modified Paxson method's approximation deteriorates as* H *approaches zero. However, as* K *increases, the approximation error of the modified Paxson method converges to zero at a faster rate than the truncation approach. Specifically, the approximation errors of the truncation and modified Paxson approaches are* $O\left(K^{-2H}\right)$ *and* $O\left(K^{-2H-1}\right)$*, respectively. This improvement is crucial, especially when* H *is close to zero. Under such circumstances, the approximation errors of the modified Paxson method and the truncation approach exhibit orders of* $O(K^{-1})$ *and* $O(1)$*, respectively. The former diminishes as* K *increases to infinity, while the latter is a constant regardless of the setting of* K*.*

Remark 3.6 *While the approximation accuracy enhances as* K *increases for both approximation methods, computing the quantities* $\sum_{k=1}^{K} Q_{1,k}$ *and* $\sum_{k=1}^{K} Q_{2,k}$ *becomes computationally costly when* K *is very large. Hence, there exists a trade-off between numerical accuracy (requiring a large* K*) and computational efficiency (benefiting from a small* K*).*

Remark 3.7 *It may seem surprising that the approximation error exhibits* uniform *convergence when* $H > 0.5$*, given that in this scenario,* $f_y^{\Delta}(\lambda)$ *behaves as* $\lambda^{1-2H} \to \infty$ *as* $\lambda \to 0$*. The reason behind this uniform bound is that both* $f_y^{\Delta}(\lambda)$ *and* $\tilde{f}_y^{\Delta}(\lambda)$ *contain the same term,* $\frac{\lambda^{1-2H}}{(\Delta\kappa)^2+\lambda^2}$*, which effectively cancels out in the approximation error. Uniform convergence plays a crucial role in establishing the asymptotic equivalence between the Whittle ML estimator and the AWML estimator.*

Approximation Accuracy

We evaluate the accuracy of two approximation methods for the spectral density of discretely sampled fOU: the truncation method (5) and the modified Paxson approximation method (8). We examine three different values of the Hurst parameter ($H = 0.1, 0.2, 0.7$) and three configurations for K (namely, $K = 50, 100, 200$). The remaining model parameters are held constant, with κ set to 25 and a fixed sampling interval of $\Delta = 1/250$, which corresponds to daily observations over a one-year period.[3] Based on the path-wise solution (2), we may obtain the exact discrete time representation of fOU as:

$$y_{j\Delta} = e^{-\kappa\Delta} y_{(j-1)\Delta} + \left(1 - e^{-\kappa\Delta}\right)\mu + \varepsilon_{j\Delta} \text{ with } \varepsilon_{j\Delta} = \sigma \int_{(j-1)\Delta}^{j\Delta} e^{-\kappa(j\Delta-u)} \, \mathrm{d} dB_u^H.$$

The values of κ and Δ leads to an autoregressive coefficient of 0.90. The monotonically of the $Q_{1,k}(\lambda)$ and $Q_{2,k}(\lambda)$ is achieved when $k > \frac{\Delta\kappa + 2\pi}{2\pi} = 1.016$.

Figure 1 plots the ratio of the lower and upper bounds of $f_y^{\Delta}(\lambda)$ to the modified Paxson approximation for $\lambda \in (0, \pi)$.[4] Several observations emerge from Figure 1. First and foremost, for all values of K and all values of H, both the lower and upper bounds closely align with the modified Paxson approximation. This indicates that the modified Paxson approximation, which averages the two bounds, offers highly accurate approximations to the true spectral density. Secondly, for any fixed H, increasing K enhances the precision of the approximation, consistent with the predictions of Theorem 3.1. For example, when $H = 0.1$, both bounds are in the neighbourhood of 0.06% relative errors if $K = 50$, which decrease to 0.015% for $K = 200$. Thirdly, for any fixed K, the larger H is, the higher the precision of the approximation. For example, when $K = 200$, both bounds are in the neighbourhood of 0.015% relative errors if $H = 0.1$ and decrease to 0.006% (0.00003%) if $H = 0.2$ ($H = 0.7$), consistent with the prediction of Theorem 3.1. The conclusion of this exercise is that the modified Paxson method offers reliable approximations to the true spectral density. For the rest of the paper, we treat the modified Paxson approximation with $K = 200$ as the true spectral density.

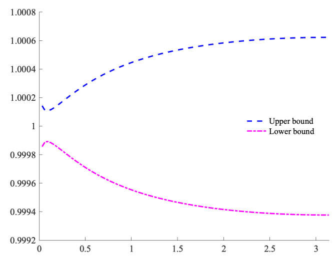
To compare the performance of the modified Paxson method and the truncation method, we calculate the ratio of the two approximate spectral densities $\bar{f}_y^{\Delta}(\lambda)/\tilde{f}_y^{\Delta}(\lambda)$, the truncation method to the modified Paxson method. Three values of K are considered for the truncation method, 2000, 5000, 20000. For the modified Paxson method, we fix K to 200. Figure 2 presents the ratios for $H = 0.1, 0.2, 0.3$. Evidently, unlike the modified Paxson approach, the truncation method is less reliable, especially when H is small. Even with large values of K, the truncation method falls short in accurately approximating the spectral density. It works well for $H \geq 0.5$, which explains its usage in previous studies (e.g., Hult, 2003). However, given recent empirical findings of small H values, using the truncation method for spectral density approximation is not recommended. Indeed, as shown in Theorem 3.1, the approximate error

[3] Results are qualitatively unchanged for $\kappa = 1, 5, 250$.

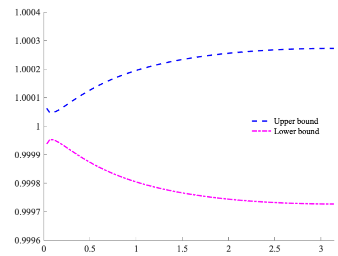
[4] We focus on this range for the spectral density as $f_y^{\Delta}(\lambda) = f_y^{\Delta}(2\pi - \lambda)$.

Figure 1: Ratio of the upper and lower bounds of the spectral density to the modified Paxson approximation $\tilde{f}_y^{\Delta}(\lambda)$ for $\lambda \in (0,\pi)$ with $\kappa = 25$, $\Delta = 1/250$.

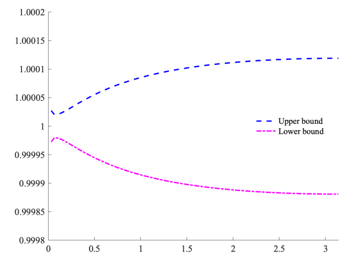
(a) $K = 50, H = 0.1$



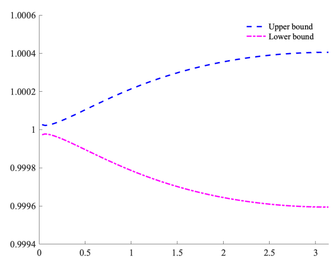
(b) $K = 100, H = 0.1$



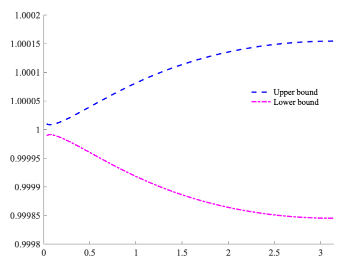
(c) $K = 200, H = 0.1$



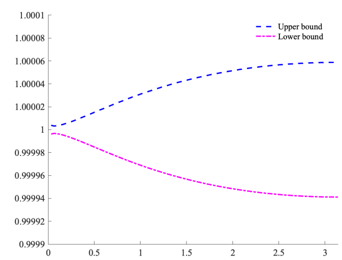
(d) $K = 50, H = 0.2$



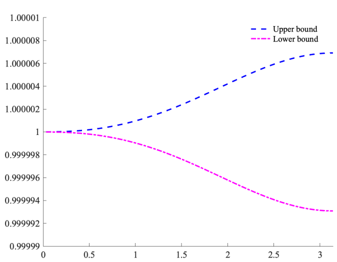
(e) $K = 100, H = 0.2$



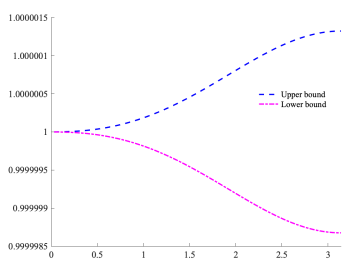
(f) $K = 200, H = 0.2$



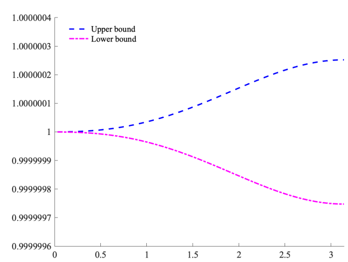
(g) $K = 50, H = 0.7$



(h) $K = 100, H = 0.7$



(i) $K = 200, H = 0.7$



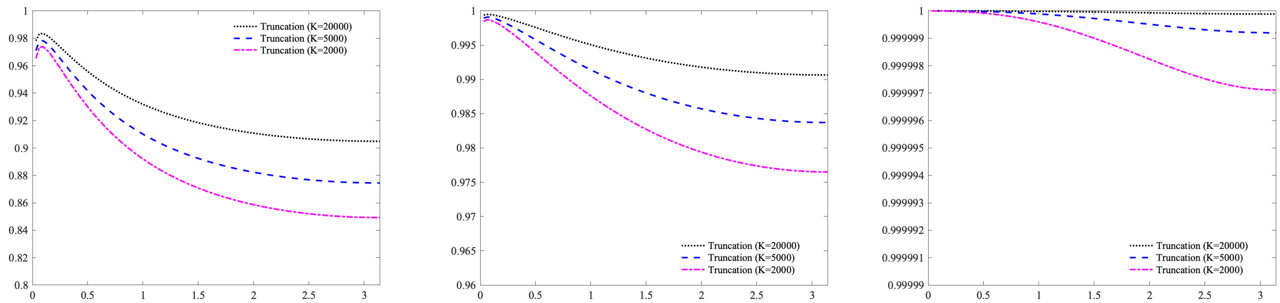
Note: The lower and upper bounds are given in Lemma 3.2, while the modified Paxson approximation is given in Theorem 3.1.

is $O(1)$ when H approaches zero, regardless of the choice of K. Furthermore, for the specific values of H, although the approximation error reduces when K increases, the truncation method becomes exceedingly computationally intensive when K is very large, thereby negatively impacting the computational efficiency of the AWML method.[5]

[5] We compute the approximate spectral density for a set of 1000 frequencies ranging between 0.001π to π, using $K = 200$ for the modified Paxson method and $K = 20,000$ for the truncation method. On a standard laptop computer (MacBook Pro with an M2 chip, 8-core CPU, and 10-core GPU), the modified Paxson method takes 0.05 seconds, while the truncation method requires 71 times of CPU time compared to the modified Paxson method.

Figure 2: Ratios of the truncation method with $K = 2000, 5000, 20000$ to the modified Paxson method with $K = 200$: $\kappa = 25$, $\Delta = 1/250$

(a) $H = 0.1$ (b) $H = 0.2$ (c) $H = 0.7$



Note: The approximate spectral density based on the truncation method is given in equation (8), while the spectral density based on the modified Paxson approximation is presented in Theorem 3.1.

4 Model Estimation

In this section, we introduce the AWML method for estimating the parameters of the fOU process, based on the Paxson approximate spectral density. We compare the performance of AWML with two existing methods: the MCL approach and the CoF approach. Our main focus is on the Hurst parameter H, which is a key parameter in the fOU process. We demonstrate that AWML provides an effective and efficient alternative to existing methods for estimating H.

4.1 Whittle ML and Approximate Whittle ML Methods

Consider a sample $\{y_{j\Delta}\}_{j=1}^{n}$. Let $\eta(\lambda;\beta) = \sigma^{-2} f_y^{\Delta}(\lambda;\beta,\sigma^2)$. The normalized spectral density, denoted by $h(\lambda;\beta)$, is defined as $h(\lambda;\beta) = \eta(\lambda;\beta) g(\beta)$ with $g(\beta) = \exp\left(-\frac{1}{2\pi}\int_0^{2\pi} \log \eta(\lambda;\beta)\, d\lambda\right)$. This normalized spectral density satisfies the property that

$$\int_0^{2\pi} \log h(\lambda;\beta)\, d\lambda = 0. \tag{22}$$

The log Whittle likelihood function (Whittle, 1951) is specified as the following:

$$\log L_W(\beta) = \frac{1}{m} \sum_{s=1}^{m} \frac{I(\lambda_s)}{h(\lambda_s;\beta)}, \tag{23}$$

where $m = \lfloor n/2 \rfloor$ and $\lambda_s = 2\pi s/n$ (with $s = 1, \ldots, m$) is the Fourier frequency. The periodogram $I(\lambda_s)$ is defined as

$$I(\lambda_s) = \frac{1}{2\pi n} \left| \sum_{j=1}^{n} y_{j\Delta} \exp(-ij\lambda_s) \right|^2. \tag{24}$$

The objective function of the Whittle ML method (23) is derived from the time-domain ML with the approximations: $\Sigma_z^{-1} \approx \left[a_{jk}\right]_{j,k=1}^{T}$ with $a_{jk} \approx (2\pi)^{-2} \int_0^{2\pi} f_y^{\Delta}\left(\lambda;\beta,\sigma^2\right)^{-1} e^{i(j-k)\lambda}\,\mathrm{d}\lambda$ and $\log|\Sigma_z| \approx T\,(2\pi)^{-1} \int_0^{2\pi} \log f_y^{\Delta}\left(\lambda;\beta,\sigma^2\right)\mathrm{d}\lambda$.

The Whittle ML estimator (denoted by $\hat{\beta}_W$) is defined as

$$\hat{\beta}_W = \arg\min_{\beta} \log L_W(\beta)$$

and the parameter σ^2 is estimated separately as the averaged ratio between $I(\lambda_s)$ and the original spectral density $\eta(\lambda_s;\hat{\beta}_W)$, i.e.,

$$\hat{\sigma}_W^2 = \frac{1}{m}\sum_{s=1}^{m} \frac{I(\lambda_s)}{\eta(\lambda_s;\hat{\beta}_W)}.$$

The Whittle ML method generally offers significantly improvements in terms of computational efficiency compared to the time-domain ML (TDML) method, particularly when dealing with a large sample size n. This efficiency gain arises because implementing the TDML method necessitates the calculation of the variance-covariance matrix Σ_z and its determinant and inverse ($|\Sigma_z|$ and Σ_z^{-1}), for which closed-form expressions are unavailable. However, computing the spectral density of fOU, which are essential for the implementation of the Whittle ML method, entails infinite summations and poses a significant challenge. As shown in Section 3, when the value of H is close to zero, achieving satisfactory computational accuracy with the truncation method is challenging, even with exceedingly large values of K. In contrast, the modified Paxson approximate spectral density $\tilde{f}_y^{\Delta}(\lambda)$ is much more computational efficient and accurate across all values of H.

Consequently, we propose to replace the true spectral density $f_y^{\Delta}(\lambda;\beta)$ with the modified Paxson approximate $\tilde{f}_y^{\Delta}(\lambda;\beta) = \sigma^2\tilde{\eta}(\lambda;\beta)$, resulting in the AWML method, denoted by $\hat{\theta}_{AW} = \left(\hat{\beta}_{AW}, \hat{\sigma}_{AW}^2\right)$ and defined by

$$\hat{\beta}_{AW} = \arg\min_{\beta} \frac{1}{m}\sum_{s=1}^{m} \frac{I(\lambda_s)}{\tilde{h}(\lambda_s;\beta)} \ \textit{and}\ \hat{\sigma}_{AW}^2 = \frac{1}{m}\sum_{s=1}^{m} \frac{I(\lambda_s)}{\tilde{\eta}(\lambda_s;\hat{\beta}_{AW})},$$

where $\tilde{h}(\lambda_s;\beta) = \tilde{\eta}(\lambda;\beta)\tilde{g}(\beta)$ and $\tilde{g}(\beta) = \exp\left(-\frac{1}{2\pi}\int_0^{2\pi} \log\tilde{\eta}(\lambda;\beta)\,\mathrm{d}\lambda\right)$.

Like the TDML method, the Whittle ML estimator is shown to be consistent, asymptotically normal and efficient in the sense of Fisher by Fox and Taqqu (1986) for strongly dependent processes whose spectral density satisfies $f_y(\lambda;\beta) \sim |\lambda|^{-2d} L_\beta(\lambda)$ as $\lambda \to 0$, where $d \in (0, 1/2)$ and $L_\beta(\lambda)$ varies slowly at 0. Velasco and Robinson (2000) extend the asymptotic results to include nonstationary $d \in [0.5, 1)$ or antipersistent $d \in (-0.5, 0)$, while Lieberman et al. (2009) provide the asymptotic results for $d < 0.5$.[6] Hult (2003) checks the validity of the asymptotic results of the Whittle ML estimator for the discretely

[6] Fukasawa and Takabatake (2019) study the asymptotic properties of the Whittle estimator for the case $d \in (-0.5, 0.5)$ under a double asymptotic scheme (i.e., the sampling interval shrinks to zero and the time span goes to infinity), which is different from the long span scheme that we adopt here.

sampled fOU process when $H \in [0.5, 1)$. Our aim here is to establish the asymptotic properties of the Whittle ML and AWML estimators for the discretely sampled fOU process with $H \in (0, 1)$. We first derive the asymptotic results of the Whittle ML estimator under the given data generating process.

Theorem 4.1 *Assume that the true parameters* $(\kappa_0, H_0, \sigma_0^2) \in Interior(\Theta)$ *where* Θ *is a compact set in* $\mathbb{R}^+ \times (0,1) \times \mathbb{R}^+$. *As* $n \to \infty$, *the Whittle ML estimator is consistent and has asymptotic normality:*

$$\sqrt{n}\left(\hat{\beta}_W - \beta_0\right) \to_d N\left(0, 4\pi\Sigma_0^{-1}\right),$$

and $\hat{\sigma}_W^2 \to_p \sigma_0^2$ *where*

$$\Sigma_0 = \int_0^{2\pi} \left\{\frac{\partial}{\partial\beta} \log\eta\left(\lambda;\beta_0\right)\right\} \left\{\frac{\partial}{\partial\beta'} \log\eta\left(\lambda;\beta_0\right)\right\} \mathrm{d}\lambda.$$

Theorem 4.2 below shows that this AWML estimator is asymptotically equivalent to the Whittle ML estimator and hence consistent and asymptotically normal.

Theorem 4.2 *Assume that the true parameters* $(\kappa_0, H_0, \sigma_0^2) \in Interior(\Theta)$ *where* Θ *is a compact set in* $\mathbb{R}^+ \times (0,1) \times \mathbb{R}^+$. *As* $n \to \infty$ *and* $K \to \infty$, *the AWML estimator is asymptotically equivalent to the Whittle ML estimator such that*

$$\begin{aligned} \sqrt{n}\left(\hat{\beta}_{AW} - \beta_0\right) &= \sqrt{n}\left(\hat{\beta}_W - \beta_0\right) + o_p(1), \\ \hat{\sigma}_{AW}^2 &= \hat{\sigma}_W^2\left[1 + o_p(1)\right]. \end{aligned}$$

4.2 Other estimation methods

We now briefly review two existing methods for estimating parameters in the fOU process: the MCL method of Bennedsen et al. (2022) and the CoF method Wang et al. (2023). Both methods are designed to improve the computational speed of TDML.

4.2.1 TDML and MCL

Let $y = (y_{1\Delta}, y_{2\Delta}, \cdots, y_{n\Delta})'$, $\theta = (\kappa, H, \sigma)$, and Σ_y be the covariance matrix of y.[7] Since $y - \hat{\mu} \sim N(0, \Sigma_y)$, the log likelihood function of fOU is

$$\log L(\theta) \propto -\frac{1}{2}\log|\Sigma_y| - \frac{1}{2}(y - \hat{\mu})'\Sigma_y^{-1}(y - \hat{\mu}). \tag{25}$$

[7] Following Lieberman et al. (2009), we assume there exists a consistent estimate (denoted by $\hat{\mu}$) of μ.

The TDML estimator is defined as

$$\hat{\theta}_{ML} = \arg\max_{\theta\in\Theta} \log L(\theta).$$

The asymptotic theory of $\hat{\theta}_{ML}$ was studied in Dahlhaus (1989) when $H \in (0.5, 1)$ and Lieberman et al. (2012) when $H \in (0, 0.5)$. Under both settings, the TDML estimator is consistent, asymptotically normal, and asymptotically efficient in the sense of Fisher.

However, the TDML method can be computationally intensive, which makes it unsuitable for a sample with a large sample size. To address this difficulty, Bennedsen et al. (2022) propose the MCL method, which aims to improve the computational speed. The objective function of the MCL is given by:

$$\log L^c(\theta) = \sum_{k=1}^{M} \sum_{i=1}^{n-k} \log \omega\left(y_{(i+k)\Delta}, y_{i\Delta}; \theta\right),$$

where $M \in \mathbb{N}^+$ is fixed, $\omega\left(y_{(i+k)\Delta}, y_{i\Delta}; \theta\right)$ is the pairwise joint probability density function (pdf) of $\left(y_{(i+k)\Delta}, y_{i\Delta}\right)$. Since $\left(y_{(i+k)\Delta}, y_{i\Delta}\right)$ follows a bivariate normal distribution with correlation coefficient ρ_k, the pdf of $\left(y_{(i+k)\Delta}, y_{i\Delta}\right)$ takes the form of

$$\omega\left(y_{(i+k)\Delta}, y_{i\Delta}; \theta\right) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho_k^2}} \exp\left[-\frac{1}{2\sigma^2\left(1-\rho_k^2\right)}\left(y_{(i+k)\Delta}^2 + y_{i\Delta}^2 - 2y_{i\Delta}y_{(i+k)\Delta}\rho_k\right)\right].$$

The computation of the correlation coefficient ρ_k is based on (3). The MCL estimator is defined as

$$\hat{\theta}_{MCL} = \arg\max_{\theta\in\Theta} \log L^c(\theta).$$

Unfortunately, the asymptotic theory for the MCL estimator, including the rate of convergence and the asymptotic distribution, critically depends on the true value of H (i.e., whether $H < 0.75$ or $H > 0.75$). Moreover, the asymptotics of the MCL estimator for the case of $H = 0.75$ is yet to be developed in the literature. See Bennedsen et al. (2022) for details.

4.2.2 CoF method

The CoF approach only relies on the first two moments of $y_{i\Delta}$ which makes it very easy to implement. The Hurst parameter H is estimated as

$$\hat{H}_{CoF} = \frac{1}{2}\log_2\left(\frac{\sum_{j=1}^{n-4}\left(y_{(j+4)\Delta} - 2y_{(j+2)\Delta} + y_{j\Delta}\right)^2}{\sum_{j=1}^{n-2}\left(y_{(j+2)\Delta} - 2y_{(j+1)\Delta} + y_{j\Delta}\right)^2}\right),$$

where $\log_2$ is 2-based logarithm. The numerator and the denominator are the second order difference of $y_{j\Delta}$ at different frequencies and hence the name of the method. The remaining two parameters σ and κ are estimated in a second stage such that

$$\hat{\sigma}^2_{CoF} = \frac{\sum_{j=1}^{n-2}\left(y_{(j+2)\Delta} - 2y_{(j+1)\Delta} + y_{j\Delta}\right)^2}{n\left(4-2^{2\hat{H}}\right)\Delta^{2\hat{H}}} \text{ and } \hat{\kappa}_{CoF} = \left(\frac{n\sum_{j=1}^{n} y_{j\Delta}^2 - \left(\sum_{j=1}^{n} y_{j\Delta}\right)^2}{n^2\hat{\sigma}^2\hat{H}\Gamma\left(2\hat{H}\right)}\right)^{-\frac{1}{2\hat{H}}}.$$

Both $\hat{H}_{CoF}$ and $\hat{\sigma}^2_{CoF}$ exhibit consistency and asymptotic normality when $T\Delta \to 0$ and $n = T/\Delta \to \infty$. The asymptotic variance of $\hat{H}$ increases as H moves closer to zero. While $\hat{H}_{CoF}$ achieves $\sqrt{n}$-consistency, the convergent rate of $\hat{\sigma}^2_{CoF}$ is slower, specifically at a rate of $\sqrt{n}/(\log(1/\Delta))$. The consistency of $\hat{\kappa}_{CoF}$ requires $T \to \infty$ and $\Delta \to 0$. The asymptotic theory for $\hat{\kappa}_{CoF}$, including the rate of convergence, the asymptotic variance, and the asymptotic distribution, heavily depends on the true value of H; for more details, refer to Wang et al. (2023).

In summary, both MCL and CoF yield the asymptotic theory that is discontinuous in H, making statistical inferences difficult to implement. In contrast, the asymptotic theory for the AWML method remains the same across all values of H within the range of $(0,1)$, thereby simplifying statistical inferences.

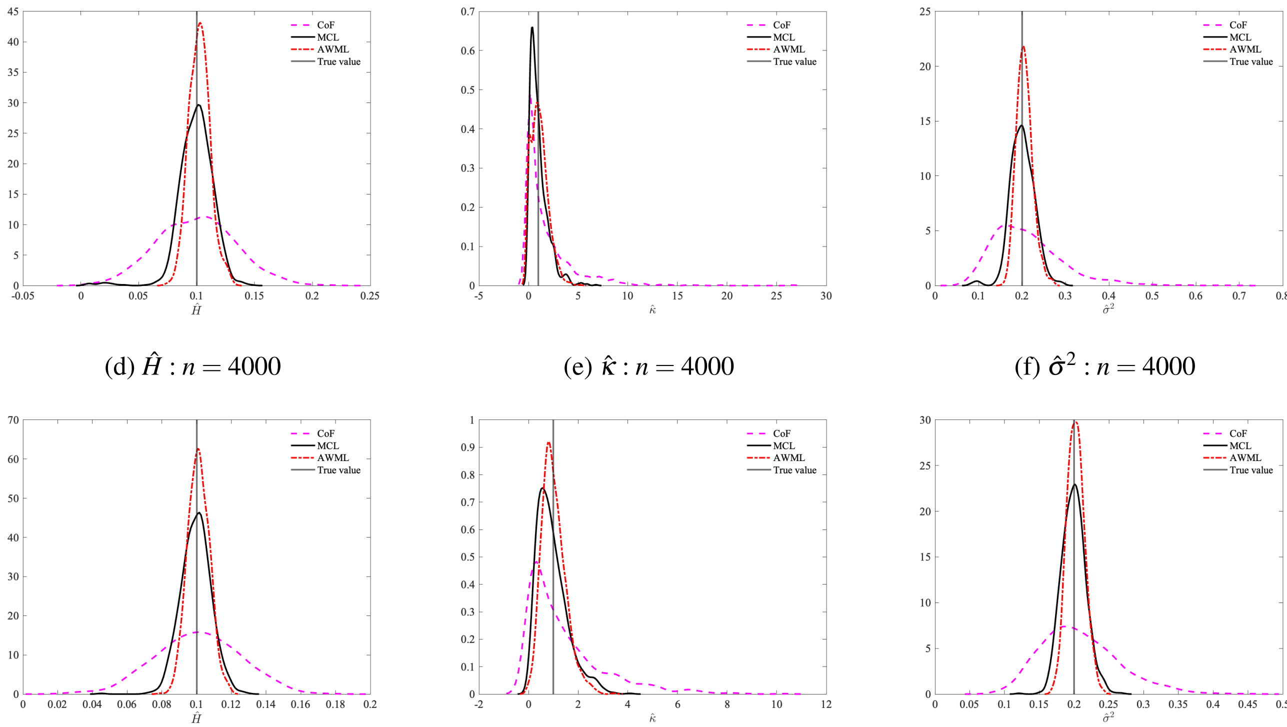
4.3 Simulation studies

The data generating process is fOU defined in (2) with two parameter settings: $\left\{\kappa, H, \sigma^2\right\} = \{1, 0.1, 0.2\}$ and $\left\{\kappa, H, \sigma^2\right\} = \{5, 0.2, 1\}$, which are close to the estimated coefficients in the empirical applications. We set the sample size to $n = 2000, 4000$ and conduct 1000 replications to compare the estimation accuracy of AWML, MCL, and CoF. For AWML, we use $K = 200$. We use the CoF estimate as an initial value.

Figures 3 and 4 depict the kernel density of the three estimates from each of the three estimation methods under the two parameter settings. The top row corresponds to a sample size of $n = 2000$, while the bottom row corresponds to $n = 4000$. It is evident that the AWML method delivers the most precise estimation results, with values clustering closely around the true value and a narrower distribution, followed by MCL. The advantage of AWML is more pronounced in Figure 3 (with a smaller value of H) than in Figure 4.

Table 1 reports the bias, standard error (std), and root mean square error (RMSE) of the three estimates from each of the three estimation methods under the two parameter settings. Evidently, the AWML method performs better than MCL, which in turn performs much better than CoF. For example, when $H = 0.1$ and $n = 4000$, the RMSE of CoF estimator for H is about four times greater than that of AWML. The benefit is also visible when we contrast the AWML method with MCL. For instance, when $H = 0.1$, the AWML method demonstrates a 50% (37%) reduction in RMSE for H in comparison to MCL for

Figure 3: Kernel densities of the estimated model parameters: $H = 0.1, \kappa = 1, \sigma^2 = 0.2$



(a) $\hat{H} : n = 2000$ (b) $\hat{\kappa} : n = 2000$ (c) $\hat{\sigma}^2 : n = 2000$

(d) $\hat{H} : n = 4000$ (e) $\hat{\kappa} : n = 4000$ (f) $\hat{\sigma}^2 : n = 4000$

Note: The kernel densities are computed from the 1000 estimates of the model parameters. The x-axis is the estimated coefficients and y-axis displays the kernel density.

$n = 2000$ ($n = 4000$). The advantage of AWML method remains sizeable when H increases to 0.2. Furthermore, the AWML method is computationally less expensive than MCL. On a standard laptop computer, it takes AWML and MCL, respectively, approximately 1.19 and 4.32 seconds to complete one iteration for 2000 observations.[8]

5 Financial Volatility Dynamics

Volatility is a critical factor in many financial decisions, and the availability of intraday data allows for the construction of realized volatility measures. In this study, we obtain the daily realized volatility (RV) of several assets, including the S&P 500 index exchange-traded fund (ETF), ETFs tracking various industry indices, and 30 Dow Jones industrial average stocks, from the Risk Lab.[9] The sample period runs from 2012 to 2019, with the number of observations ranging between 1909 and 2005. The choice of the sample period allow us to avoid the global financial crisis in 2008 and the stock turbulence in March

[8]The CPU time is based on MacBook Pro M2 chip, 8-core CPU and 10-core GPU.

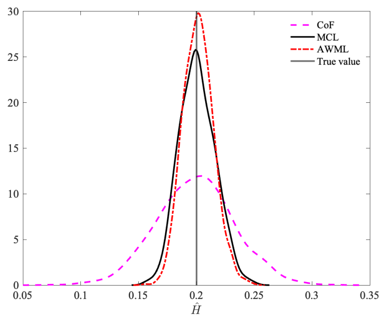
[9]https://dachxiu.chicagobooth.edu/#risklab.

Table 1: Bias, standard error and RMSE of alternative estimators.

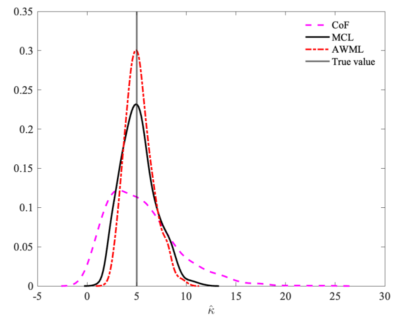
		CoF	MCL	AWML	CoF	MCL	AWML
		$\Delta = 1/250, n = 2000$			$\Delta = 1/250, n = 4000$		
DGP 1: $H = 0.1, \kappa = 1, \sigma^2 = 0.2$							
H	Bias	-0.0004	-0.0004	0.0024	0.0007	-0.0004	0.0011
	Std	0.0335	0.0151	0.0091	0.0243	0.0086	0.0062
	RMSE	0.0335	0.0151	0.0094	0.0243	0.0086	0.0063
κ	Bias	0.8561	0.0203	0.0976	0.4440	-0.0458	0.0174
	Std	2.6400	0.9263	0.8465	1.5629	0.6184	0.4793
	RMSE	2.7754	0.9266	0.8521	1.6247	0.6201	0.4796
σ^2	Bias	0.0125	0.0011	0.0059	0.0081	-0.0005	0.0023
	Std	0.0817	0.0278	0.0185	0.0553	0.0171	0.0124
	RMSE	0.0826	0.0279	0.0194	0.0559	0.0171	0.0126
DGP 2: $H = 0.2, \kappa = 5, \sigma^2 = 21$							
H	Bias	-0.0009	0.0007	0.0016	-0.0017	-0.0003	0.0002
	Std	0.0332	0.0154	0.0131	0.0241	0.0095	0.0085
	RMSE	0.0332	0.0154	0.0132	0.0242	0.0095	0.0085
κ	Bias	0.5272	0.1855	0.2636	0.1661	0.0328	0.0802
	Std	3.4972	1.7727	1.3945	2.4975	1.1861	0.9816
	RMSE	3.5367	1.7824	1.4192	2.5030	1.1865	0.9849
σ^2	Bias	0.0560	0.0184	0.0263	0.0128	-0.0030	0.0029
	Std	0.4050	0.1648	0.1393	0.2776	0.0976	0.0867
	RMSE	0.4089	0.1658	0.1417	0.2779	0.0976	0.0867

Figure 4: Kernel densities of the model parameter estimates: $H = 0.2, \kappa = 5, \sigma^2 = 1$

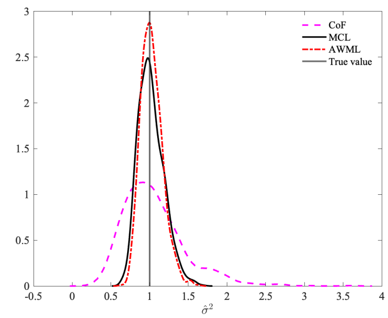
(a) $n = 2000, H = 0.2$



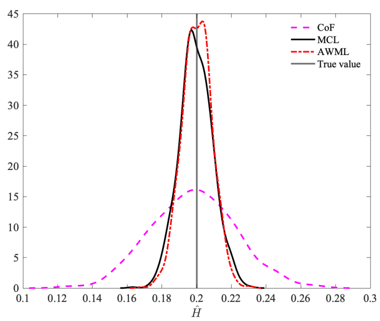
(b) $n = 2000, \kappa = 5$



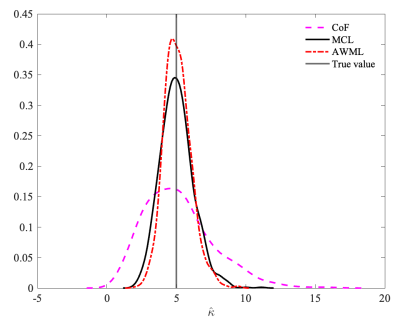
(c) $n = 2000, \sigma^2 = 1$



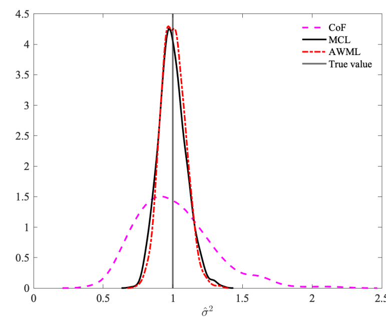
(d) $n = 4000, H = 0.2$



(e) $n = 4000, \kappa = 5$



(f) $n = 4000, \sigma^2 = 1$



Note: The kernel densities are computed from the 1000 estimates of the model parameters. The x-axis is the estimated coefficients and y-axis displays the kernel density.

2020 at the onset of Covid-19, as in Shi et al. (2022). The logarithm of RV is modeled as the fOU process. To estimate the model parameters, we employ three different approaches: AWML, MCL, and CoF. The sampling interval Δ is set to $1/252$. We use the estimates from CoF as the initial values for AWML and MCL. As in our simulations, we set $K = 200$ for AWML. The computing times for MCL and AWML are approximately 20 and 8 seconds, respectively.

Table 2 reports the estimation results of the fOU process for the ten index ETFs (top panel) and the 30 Dow Jones industrial average stocks (bottom panel). The results indicate that AWML and MCL provide very similar estimates for all three model parameters, with the estimated H ranging between 0.12 and 0.23. In contrast, the CoF estimates of H are consistently higher than those of AWML and MCL. The estimated κ by AWML and MCL falls between 1.5 and 13.23, which implies a very persistent process with the autoregressive coefficient $e^{-\kappa\Delta}$ between 0.95 and 0.994. The CoF estimates of κ and σ^2 are significantly higher than those of the ML methods, suggesting that the former may overestimate the volatility and the degree of mean reversion.

Table 2: Estimation results of fOU for the index ETFs and the 30 Dow Jones Industrial Average stocks

	CoF			MCL			AWML		
	H	κ	σ^2	H	κ	σ^2	H	κ	σ^2
	Index ETFs								
SPY	0.29	11.85	1.64	0.23	5.79	0.91	0.23	4.79	0.90
XLB	0.25	12.08	0.97	0.16	2.41	0.37	0.16	2.17	0.37
XLE	0.34	15.64	1.72	0.19	1.62	0.33	0.20	1.87	0.37
XLF	0.24	15.36	0.97	0.19	7.18	0.54	0.19	5.51	0.53
XLI	0.24	10.73	0.91	0.19	5.32	0.56	0.19	3.67	0.52
XLK	0.26	14.79	1.38	0.21	7.51	0.79	0.20	4.57	0.71
XLP	0.17	4.56	0.42	0.16	4.06	0.40	0.15	2.78	0.36
XLU	0.17	5.22	0.34	0.16	3.56	0.28	0.14	1.56	0.23
XLV	0.22	7.11	0.74	0.19	3.67	0.5	0.18	2.59	0.47
XLY	0.23	9.77	0.93	0.19	4.41	0.56	0.18	3.55	0.55
	Dow Jones 30 stocks								
AAPL	0.36	34.44	3.72	0.21	7.81	0.71	0.21	5.05	0.66
ALD	0.24	18.27	1.04	0.18	7.41	0.52	0.17	4.47	0.46
AMGN	0.22	10.89	0.71	0.15	2.80	0.34	0.15	2.08	0.34
AXP	0.23	23.13	0.96	0.16	7.95	0.42	0.15	5.44	0.38
BA	0.31	32.66	2.23	0.17	6.46	0.50	0.16	2.90	0.43
BEL	0.19	15.70	0.50	0.15	7.84	0.32	0.14	4.62	0.28
CAT	0.26	16.88	1.08	0.14	1.80	0.29	0.13	1.30	0.27
CHV	0.22	4.30	0.52	0.20	2.84	0.41	0.19	1.98	0.37
CRM	0.31	25.50	2.11	0.19	6.06	0.57	0.18	2.68	0.49
CSCO	0.28	21.86	1.16	0.19	6.93	0.44	0.18	3.95	0.39
DIS	0.25	22.61	1.03	0.18	9.08	0.48	0.17	5.52	0.43
GS	0.30	22.57	1.27	0.22	9.59	0.54	0.21	5.90	0.49
HD	0.25	21.32	1.00	0.17	5.93	0.39	0.16	4.45	0.37
IBM	0.29	31.91	1.56	0.17	7.01	0.40	0.16	4.90	0.37
INTC	0.26	17.18	0.91	0.21	9.05	0.53	0.18	3.90	0.40
JNJ	0.12	1.54	0.26	0.14	2.78	0.31	0.13	2.13	0.29
JPM	0.26	15.65	0.98	0.22	8.86	0.60	0.20	5.34	0.52
KO	0.21	12.73	0.58	0.17	6.10	0.36	0.15	3.08	0.29
MCD	0.29	37.41	1.58	0.14	4.75	0.30	0.13	2.62	0.27
MMM	0.26	15.73	1.17	0.17	4.26	0.47	0.16	1.77	0.39
MRK	0.21	15.81	0.66	0.15	5.83	0.35	0.14	3.78	0.32
MSFT	0.29	19.13	1.31	0.21	7.19	0.54	0.21	5.03	0.51
NIKE	0.27	30.84	1.17	0.17	9.78	0.41	0.16	5.47	0.34
PG	0.18	8.02	0.42	0.16	4.99	0.32	0.15	3.54	0.29
SPC	0.17	13.04	0.49	0.12	4.47	0.28	0.12	2.29	0.26
UNH	0.21	14.82	0.68	0.14	4.91	0.35	0.13	2.26	0.29
V	0.23	15.12	0.79	0.16	4.89	0.38	0.15	3.05	0.36
WAG	0.27	37.57	1.67	0.18	13.23	0.57	0.15	5.30	0.43
WMT	0.26	33.99	1.26	0.14	6.32	0.33	0.13	3.53	0.30
XOM	0.25	9.81	0.73	0.19	3.14	0.35	0.18	2.43	0.34

6 Conclusion

Computing the spectral density of the discretely sampled fOU process presents challenges due to the infinite summation in the formula, rendering it impractical for widespread application. Particularly when the Hurst parameter is low, e.g., 0.1, the commonly employed truncation method falls short. In our paper, we introduce an improved Paxson approximation for the spectral density. Our novel approach not only enhances accuracy but also boasts superior computational efficiency compared to the truncation method.

Furthermore, we demonstrate how the approximate spectral density of fOU can serve as a valuable tool for parameter estimation via the Whittle maximum likelihood method. Specifically, we propose substituting the true spectral density with its approximate counterpart in the implementation of the Whittle maximum likelihood technique, which we refer to as the AWML method. We show that for data sampled from the stationary fOU processes with $H \in (0, 1)$, the standard Whittle estimator exhibits consistency, asymptotic normality, and efficiency. Moreover, we demonstrate that the AWML estimator is asymptotically equivalent to the Whittle ML estimator. In other words, the asymptotic distribution of AWML is continuous in $H \in (0, 1)$ whereas it critically depends on the true value of H for the existing methods (e.g., MCL and CoF). This property facilitates statistical inference when using our method. Simulation results show that our proposed method outperforms existing estimation methods in finite samples.

We apply the proposed estimation method to analyze the log realized volatility of the S&P 500 index ETF, nine industry index ETFs, and 30 Dow Jones industrial average stocks from 2012 to 2019. Our estimated Hurst parameters range from 0.1 to 0.23, indicating rough volatility dynamics. Our study adds to the ongoing research on fOU processes and provides new insights into the use of the modified Paxson approximate spectral density in estimation.

Although we only demonstrate the usefulness of the approximate spectral density in constructing the ML estimate in the present paper, it can be applied in various other contexts. One potential applications is nonnested model comparison (McElroy, 2016), which is beyond the scope of the present paper and will be explored in a separate paper.

References

Angeletos, G.-M., F. Collard, and H. Dellas (2020). Business-cycle anatomy. *American Economic Review 110*(10), 3030–70. 2

Bennedsen, M., K. Christensen, and P. Christensen (2022). Likelihood-based estimation of rough stochastic volatility models. *Working paper*. 2, 3, 14, 15

Bolko, A. E., K. Christensen, M. S. Pakkanen, and B. Veliyev (2023). A GMM approach to estimate the roughness of stochastic volatility. *Journal of Econometrics 235*. 2, 9

Burnside, C. (1998). Detrending and business cycle facts: A comment. *Journal of Monetary Economics 41*(3), 513–532. 2

Christiano, L. J. and R. J. Vigfusson (2003). Maximum likelihood in the frequency domain: the importance of time-to-plan. *Journal of Monetary Economics 50*(4), 789–815. 2

Dahlhaus, R. (1989). Efficient parameter estimation for self-similar processes. *The Annals of Statistics*, 1749–1766. 15

Fox, R. and M. S. Taqqu (1986). Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *The Annals of Statistics*, 517–532. 13

Fukasawa, M. and T. Takabatake (2019). Asymptotically efficient estimators for self-similar stationary Gaussian noises under high frequency observations. *Bernoulli 25*(3), 1870–1900. 2, 7, 13

Fukasawa, M., T. Takabatake, and R. Westphal (2022). Consistent estimation for fractional stochastic volatility model under high-frequency asymptotics. *Mathematical Finance 32*(4), 1086–1132. 2, 9

Garnier, J. and K. Sølna (2018). Option pricing under fast-varying and rough stochastic volatility. *Annals of Finance 14*(4), 489–516. 4

Gatheral, J., T. Jaisson, and M. Rosenbaum (2018). Volatility is rough. *Quantitative Finance 18*(6), 933–949. 2, 4

Gehringer, J. and X.-M. Li (2020). Functional limit theorems for the fractional ornstein–uhlenbeck process. *Journal of Theoretical Probability*, 1–31. 4

Giraitis, L., H. L. Koul, and D. Surgailis (2012). *Large Sample Inference for Long Memory Processes*. IMPERIAL COLLEGE PRESS. 35

Hu, Y. and D. Nualart (2010a). Parameter estimation for fractional Ornstein–Uhlenbeck processes. *Statistics & Probability letters 80*(11-12), 1030–1038. 4

Hu, Y. and D. Nualart (2010b). Parameter estimation for fractional ornstein–uhlenbeck processes. *Statistics & probability letters 80*(11-12), 1030–1038. 35

Hult, H. (2003). *Topics on fractional Brownian motion and regular variation for stochastic processes*. Ph. D. thesis, Matematik. 2, 4, 5, 9, 10, 13

Li, J., P. C. Phillips, S. Shi, and J. Yu (2023). Weak identification of long memory with implications for volatility modeling. *Available at SSRN*. 5

Lieberman, O., R. Rosemarin, and J. Rousseau (2009). Asymptotic theory for maximum likelihood estimation in stationary fractional Gaussian processes, under short, long and intermediate memory. *preprint*. 13, 14, 28, 29

Lieberman, O., R. Rosemarin, and J. Rousseau (2012). Asymptotic theory for maximum likelihood estimation of the memory parameter in stationary Gaussian processes. *Econometric Theory 28*(2), 457–470. 15

McElroy, T. (2016). Nonnested model comparisons for time series. *Biometrika 103*(4), 905–914. 2, 21

McElroy, T. and A. Roy (2017). Detection of seasonality in the frequency domain. *Statistics*, 01. 2

Messina, J., C. Strozzi, and J. Turunen (2009). Real wages over the business cycle: OECD evidence from the time and frequency domains. *Journal of Economic Dynamics and Control 33*(6), 1183–1200. 2

Paxson, V. (1997). Fast, approximate synthesis of fractional Gaussian noise for generating self-similar network traffic. *ACM SIGCOMM Computer Communication Review 27*(5), 5–18. 2, 6

Shi, S., J. Yu, and C. Zhang (2022). Finite sample comparison of alternative estimators for fractional Gaussian noise. 19

Velasco, C. and P. M. Robinson (2000). Whittle pseudo-maximum likelihood estimation for nonstationary time series. *Journal of the American Statistical Association 95*(452), 1229–1243. 13

Wang, X., W. Xiao, and J. Yu (2023). Modeling and forecasting realized volatility with the fractional Ornstein–Uhlenbeck process. *Journal of Econometrics 234*, 28–52. 2, 3, 9, 14, 16

Whittle, P. (1951). *Hypothesis testing in time series analysis*, Volume 4. Almqvist & Wiksells boktr. 12

Whittle, P. (1954). On stationary processes in the plane. *Biometrika 41*, 434–449. 2

A Spectral Densities

Proof of Lemma 2.1. Let $\theta = \left(\kappa, H, \sigma^2\right)$. The spectral density of fOU is given by

$$f_y(\lambda;\theta) = \frac{\sigma^2}{2\pi}\Gamma(2H+1)\sin(\pi H)\frac{|\lambda|^{1-2H}}{\kappa^2+\lambda^2}, \text{ with } \lambda \in (-\infty,\infty).$$

We aim to show that if $\theta_1 \neq \theta_2$, there must be at least one value of λ for which $f_y(\lambda;\theta_1) \neq f_y(\lambda;\theta_2)$ under the assumption that $\theta \in \Theta$ which is a compact set in $\mathbb{R}^+ \times (0,1) \times \mathbb{R}^+$. Without loss of generality, we assume that $\sigma_1 = \sigma_2$.

We prove this uniqueness by contradiction. Suppose that $f_y(\lambda;\theta_1) = f_y(\lambda;\theta_2)$ for all λ. By definition, it implies that

$$\frac{\kappa_1^2+\lambda^2}{\kappa_2^2+\lambda^2} = \frac{\Gamma(2H_1+1)\sin(\pi H_1)}{\Gamma(2H_2+1)\sin(\pi H_2)}|\lambda|^{2(H_2-H_1)}. \tag{26}$$

Now, let us consider different cases:

- If $H_1 = H_2$, then $\kappa_1 \neq \kappa_2$, given that $\theta_1 \neq \theta_2$. The right-hand side of the equation is constant for all λ, while the left-hand side of (26) depends on λ. Therefore, the only way for equality to hold is if $\kappa_1 = \kappa_2$, which contradicts with the initial assumption of $\theta_1 \neq \theta_2$.

- If $H_1 > H_2$, we have $H_2 - H_1 < 0$ and as $\lambda \to 0$, $|\lambda|^{2(H_2-H_1)} \to \infty$. Consequently, the right-hand side of the equation diverges to ∞. However, the left-hand side remains bounded as long as κ_1 and κ_2 are bounded from zero and infinity, as assumed due to the compactness of κ in $\mathbb{R}^+$. Therefore, the equality does not hold.

- If $H_1 < H_2$, we have $H_2 - H_1 > 0$ and as $\lambda \to 0$, $|\lambda|^{2(H_2-H_1)} \to 0$. Consequently, the right-hand side of the equation converges to zero. However, the left-hand side remains bounded away from zero under the assumption that κ_1 and κ_2 are bounded from zero and infinity. Thus, the equality fails in this case as well.

This concludes the proof.

■

Proof of Lemma 3.1. (1) Taking derivative of $Q_{1,k}(\lambda)$ with respect to k, we have

$$\frac{\partial Q_{1,k}(\lambda)}{\partial k} = \frac{2\pi(1-2H)(2\pi k-\lambda)^{-2H}\left[(\Delta\kappa)^2+(2\pi k-\lambda)^2\right] - 2\pi(2\pi k-\lambda)^{1-2H}2(2\pi k-\lambda)}{\left[(\Delta\kappa)^2+(2\pi k-\lambda)^2\right]^2}.$$

To check the sign of $\partial Q_{1,k}(\lambda)/\partial k$, we can focus on the numerator as the denominator is always positive. Since $H > 0$, the numerator is

$$\begin{aligned}
&2\pi(1-2H)(2\pi k-\lambda)^{-2H}\left[(\Delta\kappa)^2+(2\pi k-\lambda)^2\right]-2\pi(2\pi k-\lambda)^{1-2H}2(2\pi k-\lambda)\\
=&2\pi(2\pi k-\lambda)^{-2H}\left[(1-2H)(\Delta\kappa)^2-(1+2H)(2\pi k-\lambda)^2\right]\\
<&2\pi(2\pi k-\lambda)^{-2H}\left[(\Delta\kappa)^2-(2\pi k-\lambda)^2\right].
\end{aligned}$$

When $k > \frac{\Delta\kappa+\lambda}{2\pi}$, $\partial Q_{1,k}(\lambda)/\partial k < 0$. Hence, for any given $\Delta\kappa$, we can always find K such that for any $k > K$, the function $Q_{1,k}(\lambda)$ monotonically decreases in k.

(2) Similarly, the first order derivative of $Q_{2,k}(\lambda)$ is

$$\frac{\partial Q_{2,k}(\lambda)}{\partial k}=\frac{(-2H+1)(\lambda+2\pi k)^{-2H}2\pi[(\Delta\kappa)^2+(\lambda+2\pi k)^2]-(\lambda+2\pi k)^{-2H+1}2(\lambda+2\pi k)2\pi}{[(\Delta\kappa)^2+(\lambda+2\pi k)^2]^2}.$$

Since $H > 0$, the numerator can be simplified as

$$2\pi(\lambda+2\pi k)^{-2H}[(1-2H)(\Delta\kappa)^2-(1+2H)(\lambda+2\pi k)^2]<2\pi(\lambda+2\pi k)^{-2H}[(\Delta\kappa)^2-(\lambda+2\pi k)^2].$$

It follows that when $k > \frac{\Delta\kappa-\lambda}{2\pi}$, $\partial Q_{2,k}(\lambda)/\partial k < 0$. That is, for any given $\Delta\kappa$, we can always find K such that for any $k > K$, the function $Q_{2,k}(\lambda)$ monotonically decreases in k.

(3) By combining the above results, when $k > \frac{\Delta\kappa+\lambda}{2\pi}$, we have both $Q_{1,k}(\lambda)$ and $Q_{2,k}(\lambda)$ monotonically decrease in k. Based on a properties of the harmonic sum, we have

$$\begin{aligned}
\int_{K+1}^{\infty} Q_{1,k}(\lambda)\,\mathrm{d}k &\le \sum_{k=K+1}^{\infty} Q_{1,k}(\lambda)\le\int_{K}^{\infty} Q_{1,k}(\lambda)\,\mathrm{d}k,\\
\int_{K+1}^{\infty} Q_{2,k}(\lambda)\,\mathrm{d}k &\le \sum_{k=K+1}^{\infty} Q_{2,k}(\lambda)\le\int_{K}^{\infty} Q_{2,k}(\lambda)\,\mathrm{d}k,
\end{aligned}$$

in which the two left Riemann integrals are the lower bounds and the two right Riemann integrals are the upper bounds.

Since the two sets of integrals are not available in closed-form, the upper bounds can be further bounded by

$$\int_K^\infty Q_{1,k}(\lambda)\,\mathrm{d}k=\int_K^\infty\frac{(2\pi k-\lambda)^{1-2H}}{(\Delta\kappa)^2+(2\pi k-\lambda)^2}\,\mathrm{d}k\le\int_K^\infty(2\pi k-\lambda)^{-1-2H}dk=\frac{(2\pi K-\lambda)^{-2H}}{4\pi H},$$

and

$$\int_K^\infty Q_{2,k}(\lambda)\,\mathrm{d}k=\int_K^\infty\frac{(2\pi k+\lambda)^{1-2H}}{(\Delta\kappa)^2+(2\pi k+\lambda)^2}\,\mathrm{d}k\le\int_K^\infty(2\pi k+\lambda)^{-1-2H}\,\mathrm{d}k=\frac{(2\pi K+\lambda)^{-2H}}{4\pi H}.$$

The lower bounds can be further bounded by

$$\begin{aligned}
\int_{K+1}^{\infty} Q_{1,k}(\lambda)\,\mathrm{d}k &= \int_{K+1}^{\infty} \frac{(2\pi k-\lambda)^{1-2H}}{(\Delta\kappa)^2+(2\pi k-\lambda)^2}\,\mathrm{d}k \\
&= \int_{K+1}^{\infty} \frac{(2\pi k-\lambda)^2}{(\Delta\kappa)^2+(2\pi k-\lambda)^2}\frac{(2\pi k-\lambda)^{1-2H}}{(2\pi k-\lambda)^2}\,\mathrm{d}k \\
&= \int_{K+1}^{\infty} \left[1-\frac{(\Delta\kappa)^2}{(\Delta\kappa)^2+(2\pi k-\lambda)^2}\right]\frac{(2\pi k-\lambda)^{1-2H}}{(2\pi k-\lambda)^2}\,\mathrm{d}k \\
&= \int_{K+1}^{\infty} \frac{(2\pi k-\lambda)^{1-2H}}{(2\pi k-\lambda)^2}dk - \int_{K+1}^{\infty} \frac{(\Delta\kappa)^2}{(\Delta\kappa)^2+(2\pi k-\lambda)^2}\frac{(2\pi k-\lambda)^{1-2H}}{(2\pi k-\lambda)^2}\,\mathrm{d}k \\
&\geq \int_{K+1}^{\infty} (2\pi k-\lambda)^{-1-2H}dk - (\Delta\kappa)^2\int_{K+1}^{\infty}(2\pi k-\lambda)^{-3-2H}\,\mathrm{d}k \\
&= \frac{1}{4\pi H}\left[2\pi(K+1)-\lambda\right]^{-2H} - (\Delta\kappa)^2\frac{1}{4\pi(1+H)}\left[2\pi(K+1)-\lambda\right]^{-2-2H} \\
&= \frac{1}{4\pi}\left[2\pi(K+1)-\lambda\right]^{-2H}\left\{\frac{1}{H}-\frac{(\Delta\kappa)^2}{(1+H)\left[2\pi(K+1)-\lambda\right]^2}\right\},
\end{aligned}$$

and

$$\begin{aligned}
\int_{K+1}^{\infty} Q_{2,k}(\lambda)\,\mathrm{d}k &= \int_{K+1}^{\infty} \frac{(2\pi k+\lambda)^{1-2H}}{(\Delta\kappa)^2+(2\pi k+\lambda)^2}\,\mathrm{d}k \\
&= \int_{K+1}^{\infty} \frac{(2\pi k+\lambda)^2}{(\Delta\kappa)^2+(2\pi k+\lambda)^2}\frac{(2\pi k+\lambda)^{1-2H}}{(2\pi k+\lambda)^2}\,\mathrm{d}k \\
&= \int_{K+1}^{\infty} \left[1-\frac{(\Delta\kappa)^2}{(\Delta\kappa)^2+(2\pi k+\lambda)^2}\right]\frac{(2\pi k+\lambda)^{1-2H}}{(2\pi k+\lambda)^2}\,\mathrm{d}k \\
&\geq \frac{1}{4\pi H}\left[2\pi(K+1)+\lambda\right]^{-2H} - (\Delta\kappa)^2\frac{1}{4\pi(1+H)}\left[2\pi(K+1)+\lambda\right]^{-2-2H} \\
&= \frac{1}{4\pi}\left[2\pi(K+1)+\lambda\right]^{-2H}\left\{\frac{1}{H}-\frac{(\Delta\kappa)^2}{(1+H)\left[2\pi(K+1)+\lambda\right]^2}\right\}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\frac{\left[2\pi(K+1)-\lambda\right]^{-2H}}{4\pi}\left\{\frac{1}{H}-\frac{(\Delta\kappa)^2}{(1+H)\left[2\pi(K+1)-\lambda\right]^2}\right\} &\leq \sum_{k=K+1}^{\infty} Q_{1,k}(\lambda) \leq \frac{(2\pi K-\lambda)^{-2H}}{4\pi H}, \\
\frac{\left[2\pi(K+1)+\lambda\right]^{-2H}}{4\pi}\left\{\frac{1}{H}-\frac{(\Delta\kappa)^2}{(1+H)\left[2\pi(K+1)+\lambda\right]^2}\right\} &\leq \sum_{k=K+1}^{\infty} Q_{2,k}(\lambda) \leq \frac{(2\pi K+\lambda)^{-2H}}{4\pi H}.
\end{aligned}$$

■

Proof of Lemma 3.2. (1) The discrete time spectral density of fOU is

$$f_y^{\Delta}(\lambda) = \frac{\sigma^2}{2\pi} C(H)\Delta^{2H} S(\lambda;\beta).$$

The quantity $S(\lambda;\beta)$ can be rewritten as

$$S(\lambda;\beta) = \sum_{k=1}^{K} Q_{1,k}(\lambda) + \sum_{k=K+1}^{\infty} Q_{1,k}(\lambda) + Q_0(\lambda) + \sum_{k=1}^{K} Q_{2,k}(\lambda) + \sum_{k=K+1}^{\infty} Q_{2,k}(\lambda).$$

It follows directly from Lemma 3.1(3) that

$$\begin{aligned} S(\lambda;\beta) \geq S_L(\lambda;\beta) \equiv & \sum_{k=1}^{K} Q_{1,k}(\lambda) + \sum_{k=1}^{K} Q_{2,k}(\lambda) + \frac{\lambda^{1-2H}}{(\Delta\kappa)^2 + \lambda^2} \\ & + \frac{1}{4\pi}\left[2\pi(K+1) - \lambda\right]^{-2H} \left\{ \frac{1}{H} - \frac{(\Delta\kappa)^2}{(1+H)\left[2\pi(K+1)-\lambda\right]^2} \right\} \\ & + \frac{1}{4\pi}\left[2\pi(K+1) + \lambda\right]^{-2H} \left\{ \frac{1}{H} - \frac{(\Delta\kappa)^2}{(1+H)\left[2\pi(K+1)+\lambda\right]^2} \right\}, \end{aligned} \tag{27}$$

and

$$\begin{aligned} S(\lambda;\beta) \leq \quad S_U(\lambda;\beta) \equiv & \sum_{k=1}^{K} Q_{1,k}(\lambda) + \sum_{k=1}^{K} Q_{2,k}(\lambda) + \frac{\lambda^{1-2H}}{(\Delta\kappa)^2 + \lambda^2} \\ & + \frac{1}{4\pi H}(2\pi K - \lambda)^{-2H} + \frac{1}{4\pi H}(2\pi K + \lambda)^{-2H}. \end{aligned} \tag{28}$$

Therefore, the spectral density $f_y^{\Delta}(\lambda)$ satisfies the following inequality:

$$\frac{\sigma^2}{2\pi} C(H)\Delta^{2H} S_L(\lambda;\beta) \leq f_y^{\Delta}(\lambda) \leq \frac{\sigma^2}{2\pi} C(H)\Delta^{2H} S_U(\lambda;\beta). \tag{29}$$

The modified Paxson approximation of the fOU spectral density is defined as the average of the lower and upper bounds such that

$$\begin{aligned} \tilde{f}_y^{\Delta}(\lambda) &= \frac{\sigma^2}{2\pi} C(H)\Delta^{2H} \frac{1}{2}\left[S_L(\lambda;\beta) + S_U(\lambda;\beta)\right] \\ &= \frac{\sigma^2}{2\pi} C(H)\Delta^{2H} \left\{ \sum_{k=1}^{K} Q_{1,k}(\lambda) + \sum_{k=1}^{K} Q_{2,k}(\lambda) + \frac{\lambda^{1-2H}}{(\Delta\kappa)^2 + \lambda^2} \right. \\ & \quad + \frac{1}{8\pi}\left[2\pi(K+1) - \lambda\right]^{-2H} \left\{ \frac{1}{H} - \frac{(\Delta\kappa)^2}{(1+H)\left[2\pi(K+1)-\lambda\right]^2} \right\} \\ & \quad + \frac{1}{8\pi}\left[2\pi(K+1) + \lambda\right]^{-2H} \left\{ \frac{1}{H} - \frac{(\Delta\kappa)^2}{(1+H)\left[2\pi(K+1)+\lambda\right]^2} \right\} \end{aligned}$$

$$+\frac{1}{8\pi H}(2\pi K-\lambda)^{-2H}+\frac{1}{8\pi H}(2\pi K+\lambda)^{-2H}\Bigg\}.$$

■

Proof of Theorem 3.1. (1) From Lemma 3.1, the lower bound of the quantity is given by

$$\begin{aligned}\sum_{k=K+1}^{\infty}Q_{1,k}(\lambda)+\sum_{k=K+1}^{\infty}Q_{2,k}(\lambda) &\geq \frac{[2\pi(K+1)-\lambda]^{-2H}}{4\pi}\left\{\frac{1}{H}-\frac{(\Delta\kappa)^2}{(1+H)[2\pi(K+1)-\lambda]^2}\right\}\\ &\quad+\frac{[2\pi(K+1)+\lambda]^{-2H}}{4\pi}\left\{\frac{1}{H}-\frac{(\Delta\kappa)^2}{(1+H)[2\pi(K+1)+\lambda]^2}\right\}\\ &= \frac{1}{H}\left\{\frac{[2\pi(K+1)-\lambda]^{-2H}}{4\pi}+\frac{[2\pi(K+1)+\lambda]^{-2H}}{4\pi}\right\}[1+o(1)]\\ &\geq \frac{1}{H}(2\pi)^{-2H-1}(K+2)^{-2H}[1+o(1)]=O\left(K^{-2H}\right)\end{aligned}$$

and the upper bound of the quantity is

$$\begin{aligned}\sum_{k=K+1}^{\infty}Q_{1,k}(\lambda)+\sum_{k=K+1}^{\infty}Q_{2,k}(\lambda) &\leq \frac{1}{4\pi H}\left[(2\pi K-\lambda)^{-2H}+(2\pi K+\lambda)^{-2H}\right]\\ &\leq \frac{1}{H}(2\pi)^{-2H-1}(K-1)^{-2H}=O\left(K^{-2H}\right).\end{aligned}$$

This implies that

$$\sum_{k=K+1}^{\infty}Q_{1,k}(\lambda)+\sum_{k=K+1}^{\infty}Q_{2,k}(\lambda)=O\left(K^{-2H}\right).$$

Therefore, for a fixed H and Δ, as $K\to\infty$, the approximation error of the truncation method

$$\left|f_y^{\Delta}(\lambda;\theta)-\bar{f}_y^{\Delta}(\lambda;\theta)\right|=\frac{\sigma^2}{2\pi}C(H)\Delta^{2H}\left[\sum_{k=K+1}^{\infty}Q_{1,k}(\lambda)+\sum_{k=K+1}^{\infty}Q_{2,k}(\lambda)\right]=O\left(K^{-2H}\right)$$

for any $\lambda\in(0,2\pi)$.

(2) Let

$$f_y^{\Delta,L}(\lambda)=\frac{\sigma^2}{2\pi}C(H)\Delta^{2H}S_L(\lambda;\beta)\ \text{and}\ f_y^{\Delta,U}(\lambda)=\frac{\sigma^2}{2\pi}C(H)\Delta^{2H}S_U(\lambda;\beta).$$

The modified Paxson approximate spectral density is

$$\tilde{f}_y^{\Delta}(\lambda)=\frac{1}{2}\left(f_y^{\Delta,L}(\lambda)+f_y^{\Delta,U}(\lambda)\right).$$

The approximation error is

$$\left|\tilde{f}_y^{\Delta}(\lambda) - f_y^{\Delta}(\lambda)\right| = \frac{1}{2}\left|f_y^{\Delta,L}(\lambda) - f_y^{\Delta}(\lambda) + f_y^{\Delta,U}(\lambda) - f_y^{\Delta}(\lambda)\right| \leq \left|f_y^{\Delta,U}(\lambda) - f_y^{\Delta,L}(\lambda)\right|.$$

By definition

$$f_y^{\Delta,U}(\lambda) - f_y^{\Delta,L}(\lambda) = \frac{\sigma^2}{2\pi}C(H)\Delta^{2H}\left[S_U(\lambda;\beta) - S_L(\lambda;\beta)\right].$$

Hence, by the mean value theorem, we have, uniformly in $\lambda \in (0, 2\pi)$,

$$\begin{aligned}
S_U(\lambda;\beta) - S_L(\lambda;\beta) &= \frac{(2\pi K - \lambda)^{-2H} + (2\pi K + \lambda)^{-2H} - [2\pi(K+1) - \lambda]^{-2H} - [2\pi(K+1) + \lambda]^{-2H}}{4\pi H} \\
&\quad + \frac{1}{4\pi}\frac{(\Delta\kappa)^2}{1+H}\left\{[2\pi(K+1) - \lambda]^{-2H-2} + [2\pi(K+1) + \lambda]^{-2H-2}\right\} \\
&= \xi^{-2H-1} + \xi'^{-2H-1} + \frac{(\Delta\kappa)^2}{\pi}\lambda\xi''^{-2H-3} \\
&\leq \xi^{-2H-1} + \xi'^{-2H-1} + (\Delta\kappa)^2\xi''^{-2H-3} \\
&\leq 2(2\pi)^{-2H-1}(K-1)^{-2H-1} + (\Delta\kappa)^2(2\pi)^{-2H-3}K^{-2H-3} = O\left(K^{-2H-1}\right),
\end{aligned}$$

where $\xi \in [2\pi K - \lambda, 2\pi(K+1) - \lambda]$, $\xi' \in [2\pi K + \lambda, 2\pi(K+1) + \lambda]$, $\xi'' \in [2\pi(K+1) - \lambda, 2\pi(K+1) + \lambda]$. It follows that

$$\sup_{\lambda\in(0,2\pi)}\left|f_y^{\Delta,U}(\lambda) - f_y^{\Delta,L}(\lambda)\right| = \sup_{\lambda\in(0,2\pi)}\left|\frac{\sigma^2}{2\pi}C(H)\Delta^{2H}\left[S_U(\lambda;\beta) - S_L(\lambda;\beta)\right]\right| = O\left(K^{-2H-1}\right).$$

Consequently, we have $\left|\tilde{f}_y^{\Delta}(\lambda) - f_y^{\Delta}(\lambda)\right| = O\left(K^{-2H-1}\right)$. ■

B Whittle ML and Approximate Whittle ML Estimators

Proof of Theorem 4.1. Let $q_\phi(\lambda)$ represent a general spectral density that belongs to the parametric family $\left\{q_\phi : \phi \in \Theta \subseteq \mathbb{R}^p\right\}$ such that for all $\phi \in \Theta$,

$$q_\phi(\lambda) \sim |\lambda|^{-\alpha(\phi)}L_\phi(\lambda) \text{ as } \lambda \to 0,$$

where $\alpha(\phi) < 1$, which implies that the underlying process is stationary. Lieberman et al. (2009) show that under **Assumptions** 2-3 listed below, the Whittle ML estimator is consistent and asymptotically normal.

Assumption 2 *The true parameter value $\phi_0 \in Interior(\Theta) \subset \mathbb{R}^p$ where Θ is compact. If ϕ_1 and ϕ_2 are distinct elements in Θ, then the set $\left\{\lambda \in [0, 2\pi] : q_{\phi_1}(\lambda) \neq q_{\phi_2}(\lambda)\right\}$ has a positive Lebesgue measure.*

Assumption 3 *There exists a continuous function $\alpha(\phi) : \Theta \to (-\infty, 1)$ such that for each $\delta > 0$, the following conditions hold for every $(\phi, \lambda) \in \Theta \times (0, 2\pi)$:*

(1) $q_\phi(\lambda)$, $q_\phi^{-1}(\lambda)$ *and* $\frac{\partial}{\partial\lambda}q_\phi(\lambda)$ *are continuous at all* (λ,ϕ), $\lambda\in(0,2\pi)$ *and*

$$q_\phi(\lambda)=O\left(\lambda^{-\alpha(\phi)-\delta}\right);\ q_\phi^{-1}(\lambda)=O\left(\lambda^{\alpha(\phi)-\delta}\right);\ \frac{\partial}{\partial\lambda}q_\phi(\lambda)=O\left(\lambda^{-\alpha(\phi)-1-\delta}\right);$$

(2) $\partial q_\phi(\lambda)/\partial\phi_j$ *and* $\partial^2 q_\phi(\lambda)/\partial\phi_i\partial\phi_j$ *are continuous at all* (λ,θ), $\lambda\in(0,2\pi)$ *and for any* $l\in\{1,2,3\}$ *and* $k_i\in\{1,2,\ldots,p\}$,

$$\frac{\partial^l}{\partial\phi_{k_1}\cdots\partial\phi_{k_l}}q_\theta(\lambda)=O\left(\lambda^{-\alpha(\phi)-\delta}\right);$$

(3) $\partial^2 q_\phi(\lambda)/\partial\lambda\partial\phi_i$ *are continuous at all* (λ,θ), $\lambda\in(0,2\pi)$ *and*

$$\frac{\partial^2}{\partial\lambda\partial\phi_i}q_\phi(\lambda)=O\left(\lambda^{-\alpha(\phi)-1-\delta}\right),1\le i\le p.$$

Assumption 4 *The constants appearing in the* $O(\cdot)$ *terms above can be chosen independently of* ϕ.

We aim to establish the asymptotic theory for the Whittle ML estimator in the discretely sampled fOU process. Given the results in Lieberman et al. (2009), what we need to do is to check if **Assumptions** 2-4 are satisfied in our setting.

The spectral density of the discretely sampled fOU process is given by:

$$f_y^\Delta(\lambda)=\sigma^2\eta(\lambda;\beta)\ \text{with}\ \eta(\lambda;\beta)=\frac{1}{2\pi}C(H)\Delta^{2H}S(\lambda;\beta)\ \ \text{for}\ \lambda\in[0,2\pi], \tag{30}$$

and

$$S(\lambda;\beta)\equiv\sum_{k=1}^{\infty}Q_{1,k}(\lambda)+Q_0(\lambda)+\sum_{k=1}^{\infty}Q_{2,k}(\lambda).$$

Let $\alpha(\beta)$ be a continuous function of β that takes the form of:

$$\alpha(\beta)=\begin{cases}2H-1, & \text{if } H>1/2,\\ 0, & \text{if } H\le 1/2.\end{cases}$$

We show in Lemma 2.1 that for the discretely sampled fOU, under **Assumption** 1, the spectral densities $f_y(\lambda)$ and $\tilde{f}_y(\lambda)$ are uniquely identified by θ and hence **Assumption** 2 is satisfied. We now verify the continuity property in **Assumption** 3. Since σ^2, $C(H)$ and Δ^{2H} are continuous, it suffices to verify the continuity property for $S(\lambda;\beta)$.

- $S(\lambda;\beta)$: for any k, we have

$$\begin{aligned}Q_{1,k}(\lambda) &= \frac{(2\pi k-\lambda)^{1-2H}}{(\Delta\kappa)^2+(2\pi k-\lambda)^2}<(2k\pi-\lambda)^{-1-2H}<\pi^{-1-2H}(2k)^{-1-2H},\\ Q_{2,k}(\lambda) &= \frac{(2k\pi+\lambda)^{1-2H}}{(\Delta\kappa)^2+|2k\pi+\lambda|^2}<(2k\pi+\lambda)^{-1-2H}<\pi^{-1-2H}(2k)^{-1-2H}.\end{aligned} \tag{31}$$

It follows that

$$\sum_{k=1}^{\infty} Q_{1,k}(\lambda) + \sum_{k=1}^{\infty} Q_{2,k}(\lambda) < 2\,(2\pi)^{-1-2H} \sum_{k=1}^{\infty} k^{-1-2H} < \infty. \tag{32}$$

By Weierstrass's M test, $\sum_{k=1}^{\infty} Q_{1,k}(\lambda) + \sum_{k=1}^{\infty} Q_{2,k}(\lambda)$ converges uniformly across all (λ,β). Consequently, we can find an N such that

$$\left| \sum_{k=N+1}^{\infty} Q_{1,k}(\lambda_1) + \sum_{k=N+1}^{\infty} Q_{2,k}(\lambda_1) \right| + \left| \sum_{k=N+1}^{\infty} Q_{1,k}(\lambda_2) + \sum_{k=N+1}^{\infty} Q_{2,k}(\lambda_2) \right| < \varepsilon/2.$$

Then, for any λ_1 and λ_2 with $|\lambda_1 - \lambda_2| < \varepsilon$, we have

$$\begin{aligned}
& |S(\lambda_1;\beta) - S(\lambda_2;\beta)| \\
\leq\ & \sum_{k=1}^{N} \left| Q_{1,k}(\lambda_1) - Q_{1,k}(\lambda_2) \right| + \sum_{k=1}^{N} \left| Q_{2,k}(\lambda_1) - Q_{2,k}(\lambda_2) \right| + |Q_0(\lambda_1) - Q_0(\lambda_2)| \\
+\ & \left| \sum_{k=N+1}^{\infty} Q_{1,k}(\lambda_1) + \sum_{k=N+1}^{\infty} Q_{2,k}(\lambda_1) \right| + \left| \sum_{k=N+1}^{\infty} Q_{1,k}(\lambda_2) + \sum_{k=N+1}^{\infty} Q_{2,k}(\lambda_2) \right| < \varepsilon.
\end{aligned}$$

The first three term can be made arbitrarily small because $Q_{1,k}(\lambda)$ and $Q_{2,k}(\lambda)$ are continuous with respect to (λ,β) for each k and $Q_0(\lambda)$ is continuous with respect to λ except $\lambda = 0$. Similarly, one can show the continuity of $S(\lambda;\beta)$ with respect to β. Therefore, $S(\lambda;\beta)$ is continuous at all (λ,β) with $\lambda \in (0,2\pi)$.

- $S(\lambda;\beta)^{-1}$: For any λ_1 and λ_2 with $|\lambda_1 - \lambda_2| < \varepsilon$, we have

$$\left| S(\lambda_1;\beta)^{-1} - S(\lambda_2;\beta)^{-1} \right| = \frac{|S(\lambda_1;\beta) - S(\lambda_2;\beta)|}{S(\lambda_1;\beta)\, S(\lambda_2;\beta)} < \varepsilon,$$

since $S(\lambda;\beta)$ is continuous and bounded below and away from zero at all (λ,β) with $\lambda \in (0,2\pi)$. Similarly, one can show the continuity of $S(\lambda;\beta)^{-1}$ with respect to β.

- $\frac{\partial}{\partial\lambda} S(\lambda;\beta)$: By definition,

$$\frac{\partial}{\partial\lambda} S(\lambda;\beta) = \sum_{k=1}^{\infty} \frac{\partial}{\partial\lambda} Q_{1,k}(\lambda) + \sum_{k=1}^{\infty} \frac{\partial}{\partial\lambda} Q_{2,k}(\lambda) + \frac{\partial}{\partial\lambda} Q_0(\lambda),$$

where

$$\frac{\partial}{\partial\lambda} Q_0(\lambda) = \frac{(1-2H)\,\lambda^{-2H}}{(\Delta\kappa)^2 + \lambda^2} - \frac{2\lambda^{2-2H}}{[(\Delta\kappa)^2 + \lambda^2]^2},$$

$$\frac{\partial}{\partial\lambda} Q_{1,k}(\lambda) = \frac{(2H-1)\,(2\pi k - \lambda)^{-2H}}{(\Delta\kappa)^2 + (2\pi k - \lambda)^2} + \frac{2(2\pi k - \lambda)^{2(1-H)}}{[(\Delta\kappa)^2 + (2\pi k - \lambda)^2]^2}$$

$$
\begin{aligned}
&\leq \frac{2H(2\pi k-\lambda)^{-2H}}{(2\pi k-\lambda)^2}+\frac{2(2\pi k-\lambda)^{2(1-H)}}{(2\pi k-\lambda)^4}=2\,(H+1)\,(2\pi k-\lambda)^{-2H-2} \\
&< 2\,(H+1)\,(2\pi)^{-2H-2}\,k^{-2H-2},
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial}{\partial\lambda}Q_{2,k}(\lambda) &= \frac{(1-2H)\,(2\pi k+\lambda)^{-2H}}{(\Delta\kappa)^2+(2\pi k+\lambda)^2}-\frac{2(2\pi k+\lambda)^{2(1-H)}}{\left[(\Delta\kappa)^2+(2\pi k+\lambda)^2\right]^2} \\
&\leq (1-2H)\,(2\pi k+\lambda)^{-2H-2}<(1-2H)\,(2\pi)^{-2H-2}\,k^{-2H-2}.
\end{aligned}
$$

Therefore, by Weierstrass's M test, the summation of the two infinite series is bounded above for all (λ,β) with $\lambda\in(0,2\pi)$:

$$
\sum_{k=1}^{\infty}\frac{\partial}{\partial\lambda}Q_{1,k}(\lambda)+\sum_{k=1}^{\infty}\frac{\partial}{\partial\lambda}Q_{2,k}(\lambda)<3\,(2\pi)^{-2H-2}\sum_{k=1}^{\infty}k^{-2H-2}<\infty. \tag{33}
$$

Since $\frac{\partial}{\partial\lambda}Q_0\,(\lambda)$ is continuous at all (λ,θ), following the standard argument, we can show the continuity of $\frac{\partial}{\partial\lambda}S(\lambda;\beta)$ at all (λ,θ) with $\lambda\in(0,2\pi)$.

- Similar results can be obtained for $\frac{\partial}{\partial H}S(\lambda;\beta)$, $\frac{\partial}{\partial\kappa}S(\lambda;\beta)$, $\frac{\partial^2}{\partial H^2}S(\lambda;\beta)$, $\frac{\partial^2}{\partial\kappa^2}S(\lambda;\beta)$, $\frac{\partial^2}{\partial H\partial\kappa}S(\lambda;\beta)$, $\frac{\partial^2}{\partial\lambda\partial H}S(\lambda;\beta)$, and $\frac{\partial^2}{\partial\lambda\partial\kappa}S(\lambda;\beta)$, which are omitted here for brevity.

We now verify the order property in **Assumption** 3. Since $f_y^{\Delta}\,(\lambda)=\sigma^2\frac{1}{2\pi}C\,(H)\,\Delta^{2H}S\,(\lambda;\beta)$, the orders of terms related to $f_y^{\Delta}\,(\lambda)$ are same as those of $S\,(\lambda;\beta)$.

- $S(\lambda;\beta)$: From (32), $\sum_{k=1}^{\infty}Q_{1,k}(\lambda)+\sum_{k=1}^{\infty}Q_{2,k}(\lambda)$ converges uniformly across (λ,β). As $\lambda\to 0$,

$$
Q_0\,(\lambda)=\frac{\lambda^{1-2H}}{(\Delta\kappa)^2+\lambda^2}\begin{cases}\sim c|\lambda|^{1-2H}\to\infty, & \text{if } H>1/2,\\ \to 0, & \text{if } H\leq 1/2.\end{cases} \tag{34}
$$

Therefore, $Q_0\,(\lambda)$ dominates the other two terms in $S\,(\lambda;\beta)$ when $H>1/2$. Whereas, when $H\leq 1/2$, all three terms in $S\,(\lambda;\beta)$ are bounded. It follows that

$$
S\,(\lambda;\beta)=\begin{cases} O\left(\lambda^{1-2H}\right), & \text{if } H>1/2,\\ O\,(1)\,, & \text{if } H\leq 1/2.\end{cases} \tag{35}
$$

and that for any $\delta>0$,

$$
\frac{S\,(\lambda;\beta)}{\lambda^{-\alpha(\beta)-\delta}}\to 0.
$$

- $S^{-1}(\lambda;\beta)$: from (35) we have

$$S(\lambda;\beta)/\lambda^{1-2H} = \frac{\lambda^{-(1-2H)}}{S(\lambda;\beta)^{-1}} = O(1), \quad \text{if } H > 1/2,$$
$$S(\lambda;\beta) = \frac{1}{S(\lambda;\beta)^{-1}} = O(1), \quad \text{if } H \le 1/2.$$

This implies that

$$S(\lambda;\beta)^{-1} = \begin{cases} O\left(\lambda^{-1+2H}\right), & \text{if } H > 1/2, \\ O(1), & \text{if } H \le 1/2. \end{cases}$$

It follows that for any $\delta > 0$,

$$\frac{S(\lambda;\beta)^{-1}}{\lambda^{\alpha(\beta)-\delta}} \to 0.$$

- $\frac{\partial}{\partial\lambda}S(\lambda;\beta)$: from (33), we know that $\sum_{k=1}^{\infty}\frac{\partial}{\partial\lambda}Q_{1,k}(\lambda)+\sum_{k=1}^{\infty}\frac{\partial}{\partial\lambda}Q_{2,k}(\lambda)$ is bounded. When $\lambda \to 0$,

$$\begin{aligned} \frac{\partial}{\partial\lambda}Q_0(\lambda) &= \lambda^{-2H}\frac{1-2H}{(\Delta\kappa)^2+\lambda^2} - \frac{2\lambda^{2-2H}}{[(\Delta\kappa)^2+\lambda^2]^2} \\ &= \lambda^{-2H}\frac{1-2H}{(\Delta\kappa)^2+\lambda^2} + o(1) \\ &= \begin{cases} O\left(\lambda^{-2H}\right) \to \infty, & \text{if } H < 1/2, \\ o(1), & \text{if } H = 1/2, \\ O\left(\lambda^{-2H}\right) \to -\infty, & \text{if } H > 1/2. \end{cases} \end{aligned}$$

Consequently, given the settings of $\alpha(\phi)$ for fOU, as $\lambda \to 0$, we have

$$\frac{\frac{\partial}{\partial\lambda}S(\lambda;\beta)}{\lambda^{-\alpha(\beta)-1-\delta}} = \begin{cases} \frac{\lambda^{-2H}}{\lambda^{-1-\delta}} = \lambda^{1-2H+\delta} \to 0, & \text{if } H < 1/2, \\ \frac{\lambda^{2-2H}}{\lambda^{-1-\delta}} = \lambda^{3-2H+\delta} \to 0, & \text{if } H = 1/2, \\ \frac{\lambda^{-2H}}{\lambda^{-(2H-1)-1-\delta}} = \lambda^{\delta} \to 0, & \text{if } H > 1/2. \end{cases}$$

- Similar results can be obtained for $\frac{\partial}{\partial H}S(\lambda;\beta)$, $\frac{\partial}{\partial\kappa}S(\lambda;\beta)$, $\frac{\partial^2}{\partial H^2}S(\lambda;\beta)$, $\frac{\partial^2}{\partial\kappa^2}S(\lambda;\beta)$, $\frac{\partial^2}{\partial H\partial\kappa}S(\lambda;\beta)$, $\frac{\partial^3}{\partial H^3}S(\lambda;\beta)$, $\frac{\partial^3}{\partial\kappa^3}S(\lambda;\beta)$, $\frac{\partial^3}{\partial H\partial\kappa^2}S(\lambda;\beta)$, $\frac{\partial^3}{\partial\kappa\partial H^2}S(\lambda;\beta)$, $\frac{\partial^2}{\partial\lambda\partial H}S(\lambda;\beta)$, $\frac{\partial^2}{\partial\lambda\partial\kappa}S(\lambda;\beta)$.

Therefore, all three assumptions (**Assumptions** 2-4) are satisfied by the discretely sampled fOU process. This means that for fOU processes where the true parameters $(\kappa_0, H_0, \sigma_0^2) \in Interior(\Theta)$ with Θ being a compact set in $\mathbb{R}^+ \times (0,1) \times \mathbb{R}^+$, the Whittle ML estimator is consistent an asymptotically normally distributed. ■

Proof of Theorem 4.2. To obtain the asymptotic theory for $\hat{\beta}_{AW}$ and $\hat{\sigma}_{AW}$, it is sufficient to show that

$$\begin{aligned} \sqrt{n}(\hat{\beta}_{AW}-\beta_0) &= \sqrt{n}\left(\hat{\beta}_W-\beta_0\right)+\sqrt{n}\left(\hat{\beta}_{AW}-\hat{\beta}_W\right) = \sqrt{n}\left(\hat{\beta}_W-\beta_0\right)+o_p(1), \\ \hat{\sigma}_{AW}^2-\hat{\sigma}_W^2 &= o_p(1). \end{aligned}$$

From Theorem 3.1, we know that

$$\sup_{\lambda\in(0,2\pi)} |\tilde{\eta}(\lambda;\beta)-\eta(\lambda;\beta)| = O(K^{-2H-1}). \tag{36}$$

By definition, $\eta(\lambda;\beta) = \frac{1}{2\pi}C(H)\Delta^{2H}S(\lambda;\beta)$ where

$$S(\lambda;\beta) = \sum_{k=1}^{\infty} Q_{1,k}(\lambda) + Q_0(\lambda) + \sum_{k=1}^{\infty} Q_{2,k}(\lambda).$$

From (32), $\sum_{k=1}^{\infty} Q_{1,k}(\lambda) + \sum_{k=1}^{\infty} Q_{2,k}(\lambda)$ converges uniformly in (λ,β) with $\lambda \in (0,2\pi)$. Moreover,

$$Q_0(\lambda) = \frac{\lambda^{1-2H}}{(\Delta\kappa)^2+\lambda^2}\begin{cases} \to\infty & \text{if } H>0.5 \text{ and } \lambda\to 0 \\ <\infty & \text{otherwise} \end{cases}.$$

Consequently, when $H > 0.5$ and $\lambda \to 0$,

$$\eta(\lambda;\beta) = \frac{1}{2\pi}C(H)\Delta^{2H}Q_0(\lambda)[1+o(1)] = \frac{1}{2\pi}C(H)\Delta^{2H}\frac{\lambda^{1-2H}}{(\Delta\kappa)^2+\lambda^2}[1+o(1)] \to \infty; \tag{37}$$

otherwise, $\eta(\lambda;\beta) < \infty$. It follows that

$$\sup_{\lambda\in(0,2\pi)} \frac{|\tilde{\eta}(\lambda;\beta)-\eta(\lambda;\beta)|}{\eta(\lambda;\beta)} \preceq O(K^{-2H-1}). \tag{38}$$

Therefore, as $K\to\infty$, for all λ_s,

$$\tilde{\eta}(\lambda_s;\beta) = \eta(\lambda_s;\beta)\left[1+\frac{\tilde{\eta}(\lambda;\beta)-\eta(\lambda;\beta)}{\eta(\lambda;\beta)}\right] \preceq \eta(\lambda_s;\beta)\left[1+O(K^{-2H-1})\right], \tag{39}$$

and

$$\frac{1}{\tilde{\eta}(\lambda_s;\beta)} - \frac{1}{\eta(\lambda_s;\beta)} = \frac{\eta(\lambda_s;\beta)-\tilde{\eta}(\lambda_s;\beta)}{\tilde{\eta}(\lambda_s;\beta)\eta(\lambda_s;\beta)} \preceq O(K^{-2H-1}). \tag{40}$$

Next, we examine the properties of $1/g(\beta)$. When $H \le 0.5$, $S(\lambda;\beta)$ is bounded above for all (λ,β). It follows that

$$\frac{1}{g(\beta)} = \exp\left(\frac{1}{2\pi}\int_0^{2\pi}\log\eta(\lambda;\beta)\,\mathrm{d}\lambda\right) = \frac{2\pi}{C(H)\Delta^{2H}}\exp\left(\frac{1}{2\pi}\int_0^{2\pi}\log S(\lambda;\beta)\,\mathrm{d}\lambda\right) < \infty.$$

When $H > 0.5$, $\eta(\lambda;\beta)$ is not defined at $\lambda = 0$ and $\eta(\lambda;\beta)\to\infty$ as $\lambda\to 0$. Otherwise, it is bounded. Assume that $\varepsilon\to 0$, and $\underline{\lambda}$ is the lowest frequency in $(0,2\pi)$ such that $Q_0(\lambda)$ is bounded. It follows that

$$\frac{1}{g(\beta)} = \lim_{\varepsilon\to 0}\frac{2\pi}{C(H)\Delta^{2H}}\exp\left(\frac{1}{2\pi}\int_{\varepsilon}^{2\pi}\log S(\lambda;\beta)\,\mathrm{d}\lambda\right)$$

$$
\begin{aligned}
&\leq \lim_{\varepsilon\to 0}\frac{2\pi}{C(H)\Delta^{2H}}\exp\left[\frac{1}{2\pi}\left(\int_{\varepsilon}^{\underline{\lambda}}\log\frac{\lambda^{1-2H}}{(\Delta\kappa)^2+\lambda^2}\,\mathrm{d}\lambda+\int_{\underline{\lambda}}^{2\pi}\log S(\lambda;\beta)\,\mathrm{d}\lambda\right)\right]\\
&=\lim_{\varepsilon\to 0}\frac{2\pi\exp\left\{\frac{1}{2\pi}\left[(1-2H)\lim_{\varepsilon\to 0}\int_{\varepsilon}^{\underline{\lambda}}\log\lambda\,\mathrm{d}\lambda-\lim_{\varepsilon\to 0}\int_{\varepsilon}^{\underline{\lambda}}\log\left((\Delta\kappa)^2+\lambda^2\right)\mathrm{d}\lambda+\int_{\underline{\lambda}}^{2\pi}\log S(\lambda;\beta)\,\mathrm{d}\lambda\right]\right\}}{C(H)\Delta^{2H}}\\
&<\infty,
\end{aligned}
$$

since $\lim_{\varepsilon\to 0}\int_{\varepsilon}^{\underline{\lambda}}\log\left((\Delta\kappa)^2+\lambda^2\right)\mathrm{d}\lambda<\infty$, $\int_{\underline{\lambda}}^{2\pi}\log S(\lambda;\beta)\,\mathrm{d}\lambda<\infty$, and

$$
\lim_{\varepsilon\to 0}\int_{\varepsilon}^{\underline{\lambda}}\log\lambda\,\mathrm{d}\lambda=(1-2H)\left(\underline{\lambda}\log\underline{\lambda}-\underline{\lambda}+\varepsilon\right)<\infty.
$$

That is, for all β,

$$
\frac{1}{g(\beta)}=O(1). \tag{41}
$$

By definition and using results in (38),

$$
\begin{aligned}
\frac{g(\beta)}{\tilde{g}(\beta)}&=\exp\left(-\frac{1}{2\pi}\int_0^{2\pi}\log\eta(\lambda;\beta)\,\mathrm{d}\lambda+\frac{1}{2\pi}\int_0^{2\pi}\log\tilde{\eta}(\lambda;\beta)\,\mathrm{d}\lambda\right)\\
&=\exp\left(\frac{1}{2\pi}\int_0^{2\pi}\log\frac{\tilde{\eta}(\lambda;\beta)}{\eta(\lambda;\beta)}\,\mathrm{d}\lambda\right)\\
&=\exp\left(\frac{1}{2\pi}\int_0^{2\pi}\log\left[1+\frac{\tilde{\eta}(\lambda;\beta)-\eta(\lambda;\beta)}{\eta(\lambda;\beta)}\right]\mathrm{d}\lambda\right)\\
&\preceq 1+O(K^{-2H-1}).
\end{aligned}
$$

Let Θ^S be a compact subsect in $\mathbb{R}^+\times(0,1)$. Using (39), (40) and (41), for any given β,

$$
\begin{aligned}
&\sup_{\beta\in\Theta^s}\left\{\frac{1}{m}\sum_{s=1}^{m}\frac{I(\lambda_s)}{\tilde{h}(\lambda_s;\beta)}-\frac{1}{m}\sum_{s=1}^{m}\frac{I(\lambda_s)}{h(\lambda_s;\beta)}\right\}\\
&=\sup_{\beta\in\Theta^s}\frac{1}{m}\left[\frac{1}{\tilde{g}(\beta)}\sum_{s=1}^{m}\frac{I(\lambda_s)}{\tilde{\eta}(\lambda_s;\beta)}-\frac{1}{g(\beta)}\sum_{s=1}^{m}\frac{I(\lambda_s)}{\eta(\lambda_s;\beta)}\right]\\
&=\sup_{\beta\in\Theta^s}\frac{1}{m}\frac{1}{g(\beta)}\sum_{s=1}^{m}I(\lambda_s)\left[\frac{g(\beta)}{\tilde{g}(\beta)}\frac{1}{\tilde{\eta}(\lambda_s;\beta)}-\frac{1}{\eta(\lambda_s;\beta)}\right]\\
&=\sup_{\beta\in\Theta^s}\frac{1}{m}\frac{1}{g(\beta)}\sum_{s=1}^{m}I(\lambda_s)\left[\left(\frac{1}{\tilde{\eta}(\lambda_s;\beta)}-\frac{1}{\eta(\lambda_s;\beta)}\right)+\left(\frac{g(\beta)}{\tilde{g}(\beta)}-1\right)\frac{1}{\tilde{\eta}(\lambda_s;\beta)}\right]\\
&\preceq\sup_{\beta\in\Theta^s}\frac{1}{m}\frac{1}{g(\beta)}\sum_{s=1}^{m}I(\lambda_s)\left[O(K^{-2H-1})+O(K^{-2H-1})\times O(1)\right]\\
&=\sup_{\beta\in\Theta^s}\frac{1}{m}\frac{1}{g(\beta)}\sum_{s=1}^{m}I(\lambda_s)O(K^{-2H-1})=O_p(K^{-2H-1}).
\end{aligned}
$$

The last equality is guaranteed by the stationarity and ergodicity of the fOU process (Hu and Nualart, 2010b) and the results in Lemma 8.2.2 of Giraitis et al. (2012).

This uniform convergence in the Whittle likelihood function implies that, for any n, the difference between the maximum Whittle likelihood function and the maximum approximate Whittle likelihood function can be made smaller than any positive number ε/n, as $K \to \infty$. Consequently, for any n, as $K \to \infty$, the difference between β_{AW} and β_W can be made smaller than any positive number ε_β/n, where ε_β is independent of n. Hence,

$$
\begin{aligned}
\sqrt{n}\left(\hat{\beta}_{AW} - \beta_0\right) &= \sqrt{n}\left(\hat{\beta}_W - \beta_0\right) + \sqrt{n}\left(\hat{\beta}_{AW} - \hat{\beta}_W\right) \\
&= \sqrt{n}\left(\hat{\beta}_W - \beta_0\right) + \sqrt{n}\varepsilon_\beta/n, \\
&= \sqrt{n}\left(\hat{\beta}_W - \beta_0\right) + o_p(1).
\end{aligned}
$$

Similarly, we can show that as $n \to \infty$ and $K \to \infty$, $\hat{\sigma}^2_{AW} = \hat{\sigma}^2_W\left[1 + o_p(1)\right]$. ■