Is Volatility Rough ?[*]

Masaaki Fukasawa[†1], Tetsuya Takabatake[‡2], and Rebecca Westphal[§3]

[1,2]Graduate School of Engineering Science, Osaka University[¶]
[3]Department of Management, Technology, and Economics, ETH Zürich[∥]

Abstract

Rough volatility models are continuous time stochastic volatility models where the volatility process is driven by a fractional Brownian motion with the Hurst parameter smaller than half, and have attracted much attention since a seminal paper titled "Volatility is rough" was posted on SSRN in 2014 showing that the log realized volatility time series of major stock indices have the same scaling property as such a rough fractional Brownian motion has. We however find by simulations that the impressive approach tends to suggest the same roughness irrespectively whether the volatility is actually rough or not; an overlooked estimation error of latent volatility often results in an illusive scaling property. Motivated by this preliminary finding, here we develop a statistical theory for a continuous time fractional stochastic volatility model to examine whether the Hurst parameter is indeed estimated smaller than half, that is, whether the volatility is really rough. We construct a quasi-likelihood estimator and apply it to realized volatility time series. Our quasi-likelihood is based on the error distribution of the realized volatility and a Whittle-type approximation to the auto-covariance of the log-volatility process. We prove the consistency of our estimator under high frequency asymptotics, and examine by simulations its finite sample performance. Our empirical study suggests that the volatility is indeed rough; actually it is even rougher than considered in the literature.

Keywords Rough volatility, Stochastic volatility, Fractional Brownian motion, Realized variance, Whittle estimator, High frequency data analysis

1 Introduction

Nowadays it is widely recognized that the volatility of an asset price is not a constant but a stochastic process. The property of the process is, however, not very clear because it is not a directly observable process. Even in a simple continuous framework where the volatility process σ defined through

$$\mathrm{d}S_u = \sigma_u S_u \mathrm{d}B_u \tag{1}$$

with an asset price process S and a Brownian motion B, one can only examine indirectly its properties via a statistic like the realized variance

$$\hat{\sigma}^2_{\delta,t} := \sum_{(t-1)\delta \le u \le t\delta} \left| \Delta \log \bar{S}_u \right|^2,$$

where $\bar{S}$ is a piecewise constant process which jumps at every sampling time of S to the observed value of S at the time. In a hypothetical situation where sampling frequency goes to infinity without any measurement error,

$$\hat{\sigma}^2_{\delta,t} \to \int_{(t-1)\delta}^{t\delta} \sigma^2_u \, du \tag{2}$$

in probability; one therefore expects $\hat{\sigma}^2_{\delta,t}$ to work as a proxy of the unobservable quantity. Since the high frequency asymptotics does not require any ergodicity or stationarity assumption, it particularly fits the analysis of recent financial market data, where the sampling frequency is really high. The two remarkable empirical properties of daily realized variance time series that were already documented in the earliest work by Andersen et al. [1] are that their unconditional distributions are approximately log Gaussian, and that their auto-covariances decay slowly. Various modifications of the realized variance taking into account market microstructure noise and asset price jumps have been proposed and associated limit theorems have been proven in the literature; see Aït-Sahalia and Jacod [2] for an overview.

In 2014, an interesting paper, Gatheral et al. [21] titled "Volatility is rough" was posted on SSRN. Since then, it has been so influential in the community of Mathematical Finance that a number of papers[1] have already appeared dealing with the so-called rough volatility models. In that paper, the authors looked at the historical volatility proxy data including those from the Oxford-Man realized library[2]. Let $\hat{\sigma}_t$ be such a volatility proxy as the realized volatility

$$\hat{\sigma}_t \equiv \hat{\sigma}_{\delta,t} := \sqrt{\hat{\sigma}^2_{\delta,t}}$$

for a day t computed from intraday asset price data, where δ corresponds to the length of one day. They did a linear regression to find an impressive fit

$$\log \frac{1}{n} \sum_{t=1}^{n} |\log \hat{\sigma}_{t+\Delta} - \log \hat{\sigma}_t|^q \approx \zeta_q \log \Delta + \eta_q \tag{3}$$

for various values of q; see Figure 1 (left). Then, for the regression coefficients ζ_q, they did another linear regression to find another impressive fit $\zeta_q \approx Hq$ with $H \approx 0.1$; see Figure 1 (right). Naively, this scaling property together with the above mentioned stylized fact that the realized variance is approximately log Gaussian suggests a simple dynamics

$$d \log \sigma^2_u = \eta dW^H_u, \tag{4}$$

where η is a constant and W^H is a fractional Brownian motion[3] with the Hurst parameter H. Note that the estimate $H \approx 0.1$ is not consistent to a widespread belief that the volatility is a process of long memory. Gatheral et al. [21] showed by some simulations that such a "short memory" process pretends to be of long

---

[1] Antoine Jacquier established and has maintained a website
https://sites.google.com/site/roughvol/home
as a reference point for the fast growing literature of rough volatility.

[2] The Oxford-Man Institute provides daily nonparametric volatility estimates at
https://realized.oxford-man.ox.ac.uk

[3] A fractional Brownian motion W^H is characterized as a continuous centered Gaussian process with $W^H_0 = 0$ a.s., stationary increments and $E[|W^H_{t+\Delta} - W^H_t|^q] = C_q \Delta^{Hq}$ for any $q, \Delta > 0$, where C_q is the absolute qth moment of the standard normal distribution. See Mishura [26] for further detail. The fractional Brownian motion in volatility dynamics does not imply an arbitrage opportunity because the asset price process (1) is a local martingale.
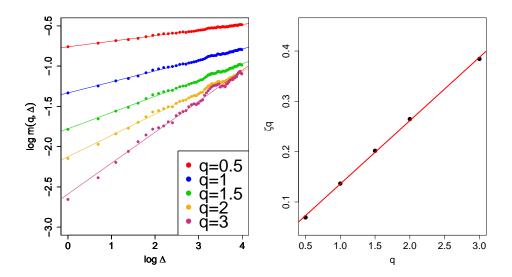
Figure 1: A reproduction of the linear regressions in Gatheral et al. [21] using SPX 5-minute realized volatility from the Oxford-Man Institute's Realized Library. Period: 03/01/2000 - 13/07/2018. The regression coefficient (Right): $H = 0.1258$.

memory. They demonstrated also a good prediction performance of this simple model. The analysis is extended by Bennedsen et al. [5] to a wider set of assets. The estimate $H \approx 0.1$ means that the volatility path is rougher than semimartingales, and is consistent to a power law for the term structure of implied volatility skew empirically observed in option markets; see [3, 16, 4, 17, 11, 20, 10]. A market microstructural foundation of a rough volatility model is given by El Euch et al. [9].

The statement $H \approx 0.1$ by Gatheral et al. [21] should be, however, understood not as a statistical estimate but as the proposal of a model which is consistent to a number of empirical evidences. In fact, as noted in that paper itself, what they "show here is that we cannot find any evidence against the RFSV [4] model". Our numerical experiments show that when using 5-minute realized volatility, the linear regression methods in Gatheral et al. [21] or in Bennedsen et al. [5] often give almost perfect fit with $H \approx 0.1$ irrespectively to the true value of H used to simulate paths; see Figure 2 just for one example, and see Westphal [33] for more extensive simulation results. Figure 2 indicates also that this striking phenomenon is due to the use of a volatility proxy; the approximation error of $\hat{\sigma}_t$ to σ_t results in an illusive scaling property. These observations from simple numerical experiments bring us a question whether the volatility is really rough, which the present paper aims at providing the first step to answer.

This is a question about the smoothness of a hidden process. Therefore any nonparametric spot volatility estimation method or filtering approach in the literature is not helpful here; such an estimator is not meant to preserve the regularity of the hidden path. Further, notice that for continuous time models like (1), most of theoretical studies in the high frequency data analysis so far have assumed that the volatility process σ is an Itô semimartingale. This is an indispensable assumption because the analysis is typically based on a piecewise constant approximation of σ, and the path regularity of σ determines the convergence rate of the approximation error; we refer again to Aït-Sahalia and Jacod [2]. There is a work by Rosenbaum [30] about fractional volatility models including (4); this however assumes $H \geq 1/2$ a priori and so, is not helpful here to study whether $H < 1/2$ (rough) or not. Other results from high frequency statistics for the model (1) that do not require σ to be an Itô semimartingale include the most primitive convergence (2), associated central

[4]Rough Fractional Stochastic Volatility. It is a special case of our model (5) below.
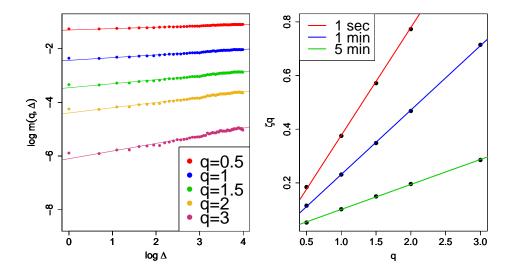
3

Figure 2: (Left) The linear regression (3) using 5-minute realized volatility of a simulated price path from the model (1) with volatility dynamics $d \log \sigma_u^2 = 10(-3.2 - \log \sigma_u^2)du + 0.8dW_u$, where W is a standard Brownian motion independent of B. (Right) For the same simulated path, using realized volatility with different sampling frequencies, the regression coefficients ζ_q of (3) are plotted and regressed on q. The regression coefficients are $H = 0.4016$ for 1 second, $H = 0.2398$ for 1 minute and $H = 0.0930 \approx 0.1$ for 5 minutes.

limit theorems by Jacod and Protter [22] and Fukasawa [14, 15], and some limit theorems for the so-called two-scales, or multi-scales realized volatility that takes the market microstructure noise into account; see Aït-Sahalia and Jacod [2].

This paper proposes a novel estimator of the Hurst and diffusion parameters under a fractional volatility model extending (4). Taking the difference between σ_t and its proxy $\hat{\sigma}_t$ into account, we derive an estimation function combining the three ideas: (1) a normal approximation of the log-realized volatility estimation error based on the above mentioned central limit theorem, (2) a local Gaussian approximation of the log-realized variance time series, and (3) a Whittle-type estimation for high frequency self-similar Gaussian models developed in Fukasawa and Takabatake [18]. The asymptotic results in this previous work are, however, not directly applicable here because the observed sequence is not Gaussian but only "locally Gaussian". The local Gaussian approximation error causes several technical difficulties in proving the consistency of our estimator. The consistency result in this paper is the first to show that a Whittle-type estimation function is effective to a non-Gaussian model under high frequency asymptotics.

Our empirical study for major stock indices indicates that H is even smaller than 0.1; so our tentative answer to the question is affirmative. It is tentative because constructing statistical tests still remains for future research. We remark that a model with $H = 0$ formally corresponds to a Gaussian multiplicative chaos [24], or a multifractal process [27, 29]. A framework including those is also a topic for future research.

The paper is organized as follows. We propose a model, construct an estimator and state a consistency theorem in Section 2, examine the finite sample performance of our estimator by simulations in Section 3, and then apply it to the Oxford-Man realized library data to get an estimate of H in Section 4. The proofs are deferred to Appendix.

4

2 Model, Quasi-likelihood Estimator, and its Consistency

2.1 Model

Denote by $(\Omega, \mathcal{F}, P, \mathbb{F})$, $\mathbb{F} = \{\mathcal{F}_u\}_{u\in[0,\infty)}$, a filtered probability space satisfying the usual condition on which an asset price process S and its volatility process σ are defined. Extending the simplest rough volatility model (4) and a fractional volatility model of Comte and Renault [8], we assume the volatility process to satisfy

$$d \log \sigma_u^2 = \kappa_u du + \eta dW_u^H, \tag{5}$$

where κ is an unknown $\mathbb{F}$-adapted càdlàg (or càglàd) process and W^H is a fractional Brownian motion which is also $\mathbb{F}$-adapted. The parameter to be estimated is $(H, \eta) \in \Theta$, where $\Theta := \Theta_H \times \Theta_\eta$ is a compact set of the form $\Theta_H := [H_-, H_+] \subset (0, 1]$ and $\Theta_\eta := [\eta_-, \eta_+] \subset (0, \infty)$. The true value is denoted by $\vartheta_0 = (H_0, \eta_0)$ and assumed to be an interior point of Θ. The process σ^2 is not directly observed and so needs to be estimated from discrete observations of the asset price S. As a proxy of the unobservable σ_t^2, we adopt the realized variance with m equidistant sampling

$$\hat{\sigma}_t^2 \equiv \hat{\sigma}_t^2(m, \delta) := \sum_{j=0}^{m-1} \left|\log S_{(t-1+(j+1)/m)\delta} - \log S_{(t-1+j/m)\delta}\right|^2, \quad m \in \mathbb{N}, \quad \delta > 0,$$

and model directly the law of the proxy error for the log realized variance

$$\epsilon_t := \log \hat{\sigma}_t^2 - \log \int_{(t-1)\delta}^{t\delta} \sigma_u^2 \, du, \quad t = 1, 2, \ldots, n+1, \tag{6}$$

as an i.i.d. sequence independent of W^H and normally distributed with mean 0 and variance $2/m$, where m is the size of intraday price data used to compute $\hat{\sigma}_t^2$. This specification of the law of $\{\epsilon_t\}$ is motivated by the following limit theorem.

Theorem 2.1. *Consider a positive sequence $\{\delta_n\}_{n\in\mathbb{N}}$ and a sequence of natural numbers $\{m_n\}_{n\in\mathbb{N}}$ satisfying $\delta_n \to 0$ and $m_n \to \infty$ as $n \to \infty$, and $\hat{\sigma}_t^2 \equiv \hat{\sigma}_t^2(m_n, \delta_n)$. Assume that a log-asset price process $\log S = A + M$ given by*

$$dM_u = \sigma_u \, dB_u, \quad dA_u = \psi_u \, du, \tag{7}$$

where $\psi = \{\psi_s\}_{s\in[0,\infty)}$ is a $\mathbb{F}$-adapted locally bounded left-continuous process, B is a standard $\mathbb{F}$-Brownian motion and the volatility process σ^2 is given in (5). Then we have

$$\left\{ \sqrt{m_n} \left(\log \hat{\sigma}_t^2 - \log \int_{(t-1)\delta_n}^{t\delta_n} \sigma_u^2 \, du\right) \right\}_{t\in\mathbb{N}} \stackrel{n\to\infty}{\to} \{ \sqrt{2}\xi_t\}_{t\in\mathbb{N}} \quad \textit{in law,}$$

where $\{\xi_t\}_{t\in\mathbb{N}}$ is an i.i.d. standard Gaussian sequence independent of $\mathcal{F}$.

The proof of Theorem 2.1 is given in Appendix H. Here we give some remarks in order.

Remark 2.2. We model the volatility dynamics (5) in the business time scale, which means that the time variable u in the model evolves only when a market of the asset is open. Therefore, the volatility is freezing when markets are closed. Including volatility jumps remains for future research.

Remark 2.3. In Theorem 2.1 and in the sequel, we consider the double high frequency limits $(\delta, 1/m) \to (0, 0)$. For example, $\delta = 0.04 = 1/250$ corresponding to the 1 day length in a year consisting of 250 business days.

For the 5-minute realized variance of market data with 6 opening hours, $m = 6 \times 60/5 = 72$.

Remark 2.4. Daily volatility proxy data including the 5-minute realized variance are readily available thanks to the Oxford-Man Institute, while high frequency price data (tick data) are not easily accessible. This motivates us to include a proxy as a model element. Among many volatility proxies, we adopt the realized variance by the following 4 reasons: i) As mentioned in Introduction, the high frequency limit theorems for the realized variance are valid without assuming the volatility σ to be an Itô semimartingale while those for others are not in general. ii) The asymptotic theory of two-scales and multi-scales realized volatilities assumes the market microstructure noise to be independent of the price and the volatility processes. While this is a popular assumption in the literature, the authors do not consider it enough realistic. iii) The realized variance with modest frequency like 5 minutes, for which the market microstructure noises are negligible but still high frequency limit theorems are valid, is easy to compute from modest frequency price data that are nowadays easily obtained online for free. iv) As shown in Theorem 2.1 above, the realized variance with equidistant sampling admits a particularly simple limit law. Note that the limit law is different for a different proxy and even so for the realized variance with a different sampling scheme; see [14, 15]. It is remarkable that the limit law in Theorem 2.1 does not depend on σ, which enables us to quantify the size of the approximation error to σ without knowing the exact value of σ.

Remark 2.5. In view of Theorem 2.1, for our model (6), more plausible would be a weaker assumption that the law of $\{\epsilon_t\}_{t \in \mathbb{N}}$ is not exactly but only asymptotically i.i.d. Gaussian with mean 0 and variance $2/m$. We believe that the same quasi-likelihood estimator given below enjoys the same consistency property also under this weaker assumption plus a suitable uniform integrability condition; we however refrain from increasing the complexity of this already technical and lengthy paper.

2.2 Construction of Adapted Whittle Estimator

Here, for a sequence of integers m_n and a positive sequence δ_n, we define a quasi-likelihood estimator of the unknown parameter $\vartheta = (H, \eta)$ based on the log-realized variance increments

$$Y_t^n := \log \hat{\sigma}_{t+1}^2(m_n, \delta_n) - \log \hat{\sigma}_t^2(m_n, \delta_n), \quad t = 1, 2, \ldots, n.$$

Firstly, we define an estimator of a reparametrized parameter (H, ν), where $\nu := \eta \delta_n^H \in \Theta_\nu^n := [\eta_- \delta_n^{H_+}, \eta_+ \delta_n^{H_-}]$, by

$$(\widehat{H}_n, \widehat{\nu}_n) := \underset{(H,\nu) \in \Theta_H \times \Theta_\nu^n}{\arg \min} \; U_n(H, \nu),$$

$$U_n(H, \nu) := \frac{1}{4\pi} \int_{-\pi}^{\pi} \left(\log g_{H,\nu}^n(\lambda) + \frac{I_n(\lambda, \mathbf{Y}_n)}{g_{H,\nu}^n(\lambda)} \right) d\lambda, \tag{8}$$

where $\mathbf{Y}_n := (Y_1^n, \cdots, Y_n^n)$, and $I_n(\cdot, \mathbf{y})$ and $g_{H,\nu}^n$ are a periodogram of $\mathbf{y} \in \mathbb{C}^n$ and an approximate spectral density of $\mathbf{Y}_n$ with respect to the reparametrized parameter (H, ν) respectively given by

$$I_n(\lambda, \mathbf{y}) := \frac{1}{2\pi n} \left| \sum_{t=1}^{n} y_t \exp\left(\sqrt{-1} t \lambda \right) \right|^2, \quad \mathbf{y} \equiv (y_1, \cdots, y_n) \in \mathbb{C}^n, \tag{9}$$

$$g_{H,\nu}^n(\lambda) := \nu^2 f_H(\lambda) + \frac{2}{m_n} \ell(\lambda), \quad \lambda \in [-\pi, \pi],$$

where f_H and ℓ are given in Appendix D. Note that, for each $n \in \mathbb{N}$, the estimator $(\widehat{H}_n, \widehat{\nu}_n)$ always exists because $\Theta_H \times \Theta_\nu^n$ is compact. Then we define an estimator of the parameter η by substituting $\widehat{H}_n$ and $\widehat{\nu}_n$ into the relation $\nu = \eta \delta_n^H$, i.e. an estimator of the original unknown parameter ϑ is defined by

$$\widehat{\vartheta}_n := \left(\widehat{H}_n, \widehat{\eta}_n \right) \quad \text{with} \quad \widehat{\eta}_n := \delta_n^{-\widehat{H}_n} \widehat{\nu}_n, \quad n \in \mathbb{N}. \tag{10}$$

We call the estimator $\widehat{\vartheta}_n$ as the adapted Whittle estimator through this paper.

Now the idea for (8) is summarized in the following remark.

Remark 2.6. Our idea to derive the approximate likelihood function (8) is based on a local approximation of $\mathbf{Y}_n$ by a certain Gaussian vector and the Whittle likelihood of a sequence of the approximate Gaussian vectors. Indeed, the Taylor and the Euler-Maruyama approximations of $\mathbf{Y}_n$ yield

$$Y_t^n \approx \frac{1}{\delta_n} \int_{t\delta_n}^{(t+1)\delta_n} \eta(W_u^H - W_{u-\delta_n}^H) \, \mathrm{d}u + (\epsilon_{t+1}^n - \epsilon_t^n) =: G_t^n, \quad t = 1, \cdots, n, \tag{11}$$

as $\delta_n \to 0$, where $\{\epsilon_t^n\}_{t \in \mathbb{Z}}$ is an i.i.d. sequence independent of W^H and normally distributed with mean 0 and variance $2/m_n$. See Appendix B for a precise statement of the above approximation. Furthermore, we can show that a covariance function of the approximate Gaussian vector $\mathbf{G}_n := (G_1^n, \cdots, G_n^n)$ is characterized by a spectral density $f_{H,\eta}^n$ given by

$$f_{H,\eta}^n(\lambda) := \eta^2 \delta_n^{2H} f_H(\lambda) + \frac{2}{m_n} \ell(\lambda), \quad \lambda \in [-\pi, \pi].$$

See Appendix D for more detail. Finally, we adopt the Whittle likelihood of the Gaussian vector $\mathbf{G}_n$, which was investigated in Fukasawa and Takabatake [18] under high frequency observations without the noise $\{\epsilon_t^n\}_{t \in \mathbb{N}}$, as an approximate likelihood of $\mathbf{Y}_n$.

Remark 2.7 (*Why we need to reparametrize ?*)**.** Under high frequency observations, due to a self-similarity property of fractional Gaussian noises, the effects of η and H fuse in the limit and the asymptotic Fisher information matrix becomes singular. As a result, it is necessary to reparametrize the parameter η in order to obtain a limit theorem of estimator. See Brouste and Fukasawa [7] and Fukasawa and Takabatake [18] for more details.

2.3 Main Theorem

We state our main theorem in this paper.

Theorem 2.8. *Assume the true value ϑ_0 is an interior point of Θ and the following three conditions $(H.1) - (H.3)$ hold:*

(H.1) $\lim_{n \to \infty} \delta_n = 0$ *and* $\lim_{n \to \infty} m_n = \infty$.

(H.2) $0 < \underline{\lim}_{n \to \infty} T_n \leq \overline{\lim}_{n \to \infty} T_n < \infty$, *where* $T_n := n\delta_n$.

(H.3) $\underline{\lim}_{n \to \infty} \inf_{H \in \Theta_H} m_n \delta_n^{2H} = \underline{\lim}_{n \to \infty} m_n \delta_n^{2H_+} > 0$.

Then a sequence of estimators $\{\widehat{\vartheta}_n\}_{n \in \mathbb{N}}$ is weakly consistent, i.e. $\widehat{\vartheta}_n \to \vartheta_0$ in probability as $n \to \infty$.

The proof of Theorem 2.8 is deferred to Appendix. Here we make comments on technical difficulties for the proof in the following remark.

Remark 2.9. One of the difficulties is that the parameter space $\Theta_H \times \Theta_\nu^n$ where the estimation function $U_n(H, \nu)$ is minimized depends on the asymptotic parameter $n \in \mathbb{N}$ and $\lim_{n\to\infty} \Theta_\nu^n = \emptyset$. As a result, $U_n(H, \nu)$ fails to satisfy the identifiability condition of the parameter (H, ν) in the limit as $n \to \infty$. In order to circumvent this difficulty, we appropriately rescale the estimator $(\widehat{H}_n, \widehat{\nu}_n)$ and its estimation function $U_n(H, \nu)$, and attempt to find a function which can identify a rescaled parameter in the limit. Actually, we can find a function $U_{n,0}(H, \widetilde{\nu})$, where $\widetilde{\nu} := \nu \delta_n^{-H_0} \in \Theta_{\widetilde{\nu}}^n := [\eta_- \delta_n^{H_+ - H_0}, \eta_+ \delta_n^{H_- - H_0}]$, which satisfies

$$(\widehat{H}_n, \delta_n^{-H_0}\widehat{\nu}_n) = \underset{(H,\widetilde{\nu})\in\Theta_H\times\Theta_{\widetilde{\nu}}^n}{\arg\min}\ U_{n,0}(H, \widetilde{\nu}), \tag{12}$$

where

$$U_{n,0}(H, \widetilde{\nu}) := \frac{1}{4\pi} \int_{-\pi}^{\pi} \left(\log h_{H,\widetilde{\nu}}^n(\lambda) + \frac{I_n\left(\lambda, \widetilde{\mathbf{Y}}_n\right)}{h_{H,\widetilde{\nu}}^n(\lambda)} \right) d\lambda, \tag{13}$$

$$h_{H,\widetilde{\nu}}^n(\lambda) := \widetilde{\nu}^2 f_H(\lambda) + \frac{2}{m_n \delta_n^{2H_0}} \ell(\lambda) \tag{14}$$

with $\widetilde{\mathbf{Y}}_n := \delta_n^{-H_0}\mathbf{Y}_n$. Indeed, $U_n(H, \nu)$ and $U_{n,0}(H, \widetilde{\nu})$ are connected by the relation $U_n(H, \nu) = H_0 \log \delta_n + U_{n,0}(H, \nu/\delta_n^{H_0})$ so that the estimator $(\widehat{H}_n, \widehat{\nu}_n)$ also minimizes

$$(\widehat{H}_n, \widehat{\nu}_n) = \underset{(H,\nu)\in\Theta_H\times\Theta_\nu^n}{\arg\min}\ U_{n,0}(H, \nu/\delta_n^{H_0}).$$

As a result, (12) follows from the one-to-one correspondence between ν and $\widetilde{\nu}$. Then we can show that $U_{n,0}(H, \widetilde{\nu})$ and $(\widehat{H}_n, \delta_n^{-H_0}\widehat{\nu}_n)$ converge to a certain function which can identify the rescaled parameter $(H, \widetilde{\nu})$ and to the true value $(H_0, \widetilde{\nu}_0) = (H_0, \eta_0)$, where denote $\nu_0 := \eta_0\delta_n^{H_0}$ and $\widetilde{\nu}_0 := \nu_0\delta_n^{-H_0}$, respectively when, at least, the proxy error rapidly vanishes in the sense of (H.3). Furthermore, we can also show that the estimator $\widehat{\vartheta}_n$ converges to (H_0, η_0) by using the convergence $(\widehat{H}_n, \delta_n^{-H_0}\widehat{\nu}_n)$ to (H_0, η_0). Note that $U_{n,0}(H, \widetilde{\nu})$ is not a true estimation function because the true value H_0 is used in its definition. It plays, however, the similar role to the usual estimation function due to (12). The final remark is that a sequence of rescaled parameter spaces $\{\Theta_{\widetilde{\nu}}^n\}_{n\in\mathbb{N}}$ converges to the unbounded set $(0, \infty)$ so that several additional cares in the proof are necessary.

3 Numerical Study

In this section, we examine the finite sample performance of the adapted Whittle estimator proposed in Section 2.2 when the log-volatility dynamics is given by a fractional Ornstein-Uhlenbeck process with mean-reverting property. We explain how to simulate a sample path of an asset price process following the fractional volatility model in Section 3.1 and how to implement the adapted Whittle estimation in Section 3.2. We summarize several numerical results in Section 3.3.

3.1 Simulation Method for Asset Price Process

In our numerical studies, we simulate an asset price process whose log-volatility process is given by the fractional Ornstein-Uhlenbeck process, i.e.

$$d \log S_u = \sigma_u\, dB_u, \quad d \log \sigma_u^2 = \alpha(c - \log \sigma_u^2)\, du + \eta_0\, dW_u^{H_0}, \tag{15}$$

8

by using the Euler-Maruyama approximation, where B is a Brownian motion independent of W^{H_0}. Here we generate the fractional Brownian motion W^{H_0} by using the R-function "SimulateFGN" given in the R-package "FGN". We consider the case of $\delta = 1/250$ and $T := n\delta = 10$. For the size of the price data m used to compute the realized volatility, we consider three cases: $m = 80$ and 80×5 of which the values are corresponding to those of 5-minute and 1-minute realized volatilities respectively. Moreover, our model parameters are given by $H_0 = 0.01, 0.05, 0.1, 0.3, 0.5, 0.7, \eta_0 = 1, 2, 3, \alpha = 0.001, c = \log \sigma_0^2 = -3.2$ and $S_0 = 100$. We generate 100 paths to have 100 samples of the estimator.

3.2 Implementation of Adapted Whittle Estimator

Denote by $m \equiv m_n$ and $g_{H,\nu} \equiv g_{H,\nu}^n$ for notational simplicity. In order to implement the adapted Whittle estimator, we evaluate the estimation function $U_n(H, \nu)$ by

$$U_n(H, \nu) \approx \frac{1}{2\pi} \int_\psi^\pi \left(\log g_{H,\nu}(\lambda) + \frac{I_n(\lambda, \mathbf{Y}_n)}{g_{H,\nu}(\lambda)} \right) d\lambda + A_{H,\nu}^1(\psi) + A_{H,\nu}^2(\psi) \tag{16}$$

for sufficiently small $\psi > 0$, where the above integral is calculated using the R-function "integrate" and additional correction terms $A_{H,\nu}^1(\psi)$ and $A_{H,\nu}^2(\psi)$ are respectively given by

$$A_{H,\nu}^1(\psi) := \frac{1}{2\pi} \left(\psi \log(\nu^2 C_H) + \psi(\log \psi - 1)(1 - 2H) + \frac{\psi^{2+2H}}{\nu^2 C_H m\pi(2 + 2H)} \right),$$

$$A_{H,\nu}^2(\psi) := \frac{1}{2\pi} \left(a_{H,\nu}(0, \psi)\widehat{\gamma}_n(0) + 2 \sum_{\tau=1}^{n-1} a_{H,\nu}(\tau, \psi)\widehat{\gamma}_n(\tau) \right), \quad \psi \in (0, \pi],$$

with $\widehat{\gamma}_n(\tau) := \frac{1}{n} \sum_{t=1}^{n-|\tau|} Y_t^n Y_{t+|\tau|}^n$ and $a_{H,\nu}(\tau, \psi) \equiv a_{H,\nu}(\tau, \psi, J)$ given by

$$a_{H,\nu}(\tau, \psi, J) := \frac{1}{2\pi} \sum_{j=0}^J \frac{(-1)^j \tau^{2j}}{(2j)!} \frac{1}{\nu^2 C_H} \left(\frac{\psi^{2j+2H}}{2j + 2H} - \frac{\psi^{1+2j+4H}}{\nu^2 C_H m\pi(1 + 2j + 4H)} \right)$$

for a sufficiently large $J \in \mathbb{N}$ and each $\tau \in \{0, 1, \cdots, n - 1\}$. The derivation (16) is given in Appendix I. Note that the auto-covariance function $\widehat{\gamma}$ can be effectively computed using the fast Fourier transform algorithm. Moreover, we adopt the Paxson approximation of spectral densities for the spectral density $g_{H,\nu}$ used in (16), i.e. $g_{H,\nu}$ is approximated by

$$g_{H,\nu}(\lambda) \approx \nu^2 C_H \{2(1 - \cos \lambda)\}^2 \left\{ |\lambda|^{-3-2H} + \sum_{k=1}^K d_H^1(k, \lambda) + \frac{1}{2} \left(d_H^2(K, \lambda) + d_H^2(K + 1, \lambda) \right) \right\} + \frac{2}{m} \ell(\lambda)$$

with a sufficiently large $K \in \mathbb{N}$, where

$$d_H^1(x, \lambda) := (2\pi x + \lambda)^{-3-2H} + (2\pi x - \lambda)^{-3-2H},$$

$$d_H^2(x, \lambda) := \frac{1}{2\pi(2 + 2H)} \left\{ (2\pi x + \lambda)^{-2-2H} + (2\pi x - \lambda)^{-2-2H} \right\}$$

for $x \in (1/2, \infty)$ and $\lambda \in [-\pi, \pi]$; see Fukasawa and Takabatake [19] for more detail. We fix $\psi = 10^{-5}, K = 500$ and $J = 20$ in our numerical studies.

Finally, we briefly explain how to numerically evaluate the minimizer $(\widehat{H}_n, \widehat{\nu}_n)$ of the estimation function $U_n(H, \nu)$. In our numerical studies, we use the R-function "optim" in order to obtain the minimizer and

select the option "L-BFGS-B" as the optimization method of $U_n(H, \nu)$. Then we consider the parameter space $\Theta = \Theta_H \times \Theta_\eta = [0.001, 0.99] \times [0.1, 10]$ and take the true value (H_0, ν_0), where $\nu_0 = \eta_0(1/250)^{H_0}$, as the initial value of the optimization for $U_n(H, \nu)$.

3.3 The numerical results

In Table 1 and Table 2, we give the mean and variance of $\widehat{H}_n$ and $\widehat{\eta}_n$ respectively. The tables show that when the Hurst parameter is greater than 0.05, both of the Hurst and diffusion parameters are estimated reasonably well even with 5-minute realized volatility. In the case of $H_0 = 0.01$, positive and negative estimation biases are observed for the estimator $\widehat{H}_n$ and $\widehat{\eta}_n$ respectively. There would be, however, no problem in examining whether the volatility is rough ($H_0 < 0.5$ or not) because there are few estimation biases in the case of $H_0 \geq 0.05$ and the size of them in the case of $H_0 < 0.05$ would not be too large. Thus, we conclude that the adapted Whittle estimator gives a reliable answer to our question with data analysis using 5-minute realized volatility.

	$\eta_0 = 1$		$\eta_0 = 2$		$\eta_0 = 3$	
	Mean	Variance	Mean	Variance	Mean	Variance
$H_0 = 0.01$						
1 min	0.03189	0.0006475	0.02437	0.0004847	0.02348	0.0004543
5 min	0.04543	0.0011467	0.02659	0.0005737	0.02179	0.0004091
$H_0 = 0.05$						
1 min	0.06662	0.0003567	0.06300	0.0002407	0.05853	0.0003239
5 min	0.06606	0.0004830	0.05850	0.0002776	0.05405	0.0003375
$H_0 = 0.1$						
1 min	0.10717	0.0001890	0.10267	0.0002609	0.09905	0.0002079
5 min	0.10527	0.0003103	0.09709	0.0002675	0.09303	0.0002458
$H_0 = 0.3$						
1 min	0.30185	0.0002681	0.30102	0.0002316	0.29975	0.0002678
5 min	0.30029	0.0005437	0.29672	0.0003572	0.29557	0.0003542
$H_0 = 0.5$						
1 min	0.50131	0.0007236	0.50107	0.0003558	0.49946	0.0002952
5 min	0.49196	0.0016275	0.49921	0.0006041	0.49863	0.0004874
$H_0 = 0.7$						
1 min	0.70793	0.0017105	0.70928	0.0008751	0.70514	0.0005364
5 min	0.70212	0.0033792	0.71401	0.0014436	0.70388	0.0014050

Table 1: The mean and variance of the adapted Whittle estimator of the Hurst parameter.

	$\eta_0 = 1$		$\eta_0 = 2$		$\eta_0 = 3$	
	Mean	Variance	Mean	Variance	Mean	Variance
$H_0 = 0.01$						
1 min	0.8014	0.0438057	1.741	0.140862	2.687	0.296818
5 min	0.7293	0.0459881	1.728	0.147558	2.757	0.250566
$H_0 = 0.05$						
1 min	0.9895	0.0021232	2.004	0.004104	3.114	0.038000
5 min	1.0111	0.0049482	2.067	0.010018	3.223	0.057503
$H_0 = 0.1$						
1 min	1.0217	0.0006927	2.047	0.002700	3.091	0.004235
5 min	1.0341	0.0007719	2.063	0.002196	3.124	0.004053
$H_0 = 0.3$						
1 min	1.0101	0.0047328	2.012	0.016563	3.026	0.046931
5 min	1.0117	0.0064975	1.993	0.020801	2.996	0.052943
$H_0 = 0.5$						
1 min	1.0175	0.0120100	2.016	0.032845	3.008	0.068443
5 min	0.9987	0.0181695	2.009	0.040863	3.009	0.085446
$H_0 = 0.7$						
1 min	1.0584	0.0338098	2.114	0.090472	3.109	0.123645
5 min	1.0582	0.0504322	2.160	0.136346	3.120	0.231304

Table 2: The mean and variance of the adapted Whittle estimator of the diffusion parameter.

4 Application to Daily Realized Volatility Data of Stock Indices

In this section, we apply the adapted Whittle estimator to the 5-minute daily realized volatility data for several major stock indices provided by the Oxford-Man realized library. We give the estimated values in Section 4.1 and give an additional discussion in Section 4.2.

4.1 Estimation Results

First of all, we make several remarks on the optimization of the estimation function. In our data analysis, we use the same implementation and optimization methods of the estimation function $U_n(H, \nu)$ and the same parameter space $\Theta = \Theta_H \times \Theta_\eta$ mentioned in Section 3.2. Then we calculate the optimal value in the candidates of the estimated values each of which is obtained from the optimization method starting at each initial value (H_{ini}, ν_{ini}) with $H_{ini} \in \{0.01, 0.05, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ and $\nu_{ini} \in \{0.5, 1.5, 2.5, 3.5\}$.

Next, we briefly explain how to compute the value of m which is the size of price data used to compute the 5-minute daily realized volatility for each stock index. In our data analysis below, we consider the following 5 stock indices: S&P 500, FTSE 100, Nikkei 225, DAX, Russell 3000. Then we can easily calculate the value of m for each stock index since we know the opening hours of the markets are from 9:30 to 16:00 for S&P 500, from 8:00 to 16:30 for FTSE 100, from 9:00 to 11:30 and from 12:30 to 15:00 for Nikkei 225, from 9:00 to 17:40 for DAX, from 9:30 to 16:00 for Russell 3000, see Remark 2.3 for an example of computation of

m. In the first row of Table 3, we summarize the values of *m* for the stock indices. We give the estimated values $(\widehat{H}_n, \widehat{\eta}_n)$ in Table 3. For all indices, the Hurst parameter is estimated between 0.02 and 0.06. Our data analysis suggests $H < 0.5$, that is, the volatility is indeed rough; it is even rougher than claimed in Gatheral et al. [21] and Bennedsen et al. [5]. It is noteworthy that the estimate $H < 0.1$ is consistent to the calibrated parameters from the option market data in Bayer et al. [4].

	SPX 500	FTSE 1000	Nikkei 225	DAX	Russell 3000
m	78	102	60	105	78
$\widehat{H}$	0.04272	0.02255	0.05928	0.03551	0.03926
$\widehat{\eta}$	2.53112	2.89098	2.01702	2.21585	2.39789

Table 3: Estimated value of the adapted Whittle estimator $(\widehat{H}_n, \widehat{\eta}_n)$ of major stock indices for the period: 02/01/2008-29/12/2017.

4.2 Additional Discussion

In this subsection, we check whether the estimated values given in Section 4.1 are adequate from another aspect. More specifically, we apply the linear regression method of Gatheral et al. [21] to simulated 5-minute realized volatility data for which the Hurst and diffusion parameters are chosen to be the estimated values.

In Figure 3, we compare the linear regression results of SPX and simulated data for the same period of the estimation results given in Table 3. Taking into account the estimation bias of the adapted Whittle estimator mentioned in Section 3.3, we used slightly smaller value of the Hurst parameter than its estimated value given in Section 4.1. We obtained similar figures and linear regression coefficients of ζ_q against q from the SPX and simulated 5-minute realized volatilities. In particular, we confirm that the linear regression method of Gatheral et al. [21] does not give a proper estimate of H and our estimated value of (H, η) does not contradict the analysis in Gatheral et al. [21].

5 Conclusion

We have questioned whether the volatility is really rough, that is, whether the Hurst index of the fractional Brownian motion driving the volatility process is smaller than 0.5 or not. We have proposed an estimator for the Hurst and diffusion parameters under a fractional stochastic volatility model (5). We have proved its consistency under high frequency asymptotics. We have also confirmed by numerical simulations its reasonably good performance with finite samples. The estimated Hurst parameters for various stock indices and periods are all smaller than 0.06, indicating that the volatility is rough. This is however a tentative answer to our question; in particular constructing statistical tests remains for future research.
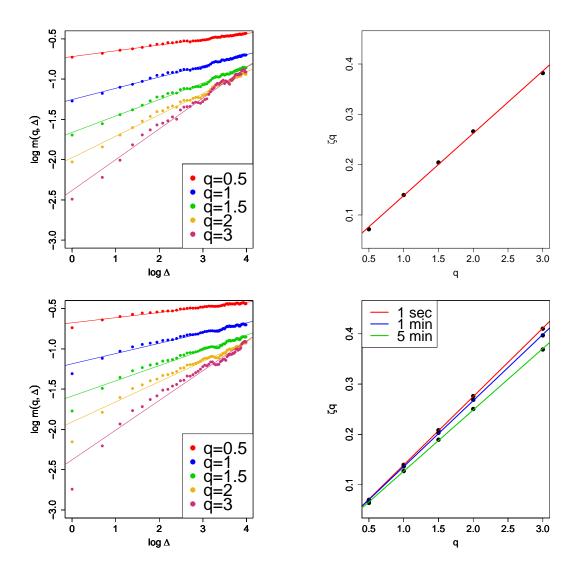
Figure 3: The upper figure is a reproduction of the linear regressions in Gatheral et al. [21] using SPX 5-minute realized volatility from the Oxford-Man Institute's Realized Library. Period: 02/01/2008 - 29/12/2017. The regression coefficient is $H = 0.1215$. The lower left figure is the linear regression (3) using 5-minute realized volatility of a simulated price path from the model (15) whose parameters are $H = 0.03$, $\eta = 2.5$, $\alpha = 0.005$, $c = -3.2$. For the same simulated path, using realized volatility with different sampling frequencies, the regression coefficients ζ_q of (3) are plotted and regressed on q in the lower right figure. The regression coefficients are $H = 0.1359$ for 1 second, $H = 0.1313$ for 1 minute and $H = 0.1219$ for 5 minutes.

References

[1] Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P. (2001): The Distribution of Realized Exchange Rate Volatility, *Journal of the American Statistical Association*, 96, 42-55.

[2] Aït-Sahalia, Y. and Jacod, J. (2014): *High-Frequency Financial Econometrics*, Princeton University Press.

[3] Alòs, E., León, J.A., Vives, J. (2007): On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. *Finance Stoch.* 11, 571-589.

[4] Bayer, C., Friz, P. and Gatheral, J. (2016): Pricing under rough volatility, *Quant. Finance.* 16(6), 887-904.

[5] Bennedsen, M., Lunde, A. and Pakkanen, M. S. (2017): Decoupling the Short- and Long-Term Behavior of Stochastic Volatility, *arXiv.*

[6] Billingsley, P. (1968): *Convergence of probability measures*, John Wiley.

[7] Brouste, A. and Fukasawa, M. (2018): Local asymptotic normality property for fractional Gaussian noise under high-frequency observations, *Ann. Statist.* 46, 2045-2061.

[8] Comte, F. and Renault, E. (1998): Long Memory in Continuous-Time Stochastic Volatility Models, *Math. Finance* 8(4), 291-323.

[9] El Euch, O., Fukasawa, M. and Rosenbaum, M. (2018): The microstructual foundations of leverage effect and rough volatility, *Finance Stoch.* 22, 241-280.

[10] El Euch,O., Fukasawa, M., Gatheral, J. and Rosenbaum, M. (2019): Short-term at-the-money asymptotics under stochastic volatility models, *SIAM J. Finan. Math.* 10, 491-511.

[11] Forde, M. and Zhang, H.(2017): Asymptotics for rough stochastic volatility models, *SIAM J. Financial Math.* 8, 114-145.

[12] Fox, R. and Taqqu, M. S. (1986): Large-Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series, *Ann. Statist.* 14(2), 517-532.

[13] Fox, R. and Taqqu, M. S. (1987): Central Limit Theorems for Quadratic Forms in Random Variables having Long-Range Dependence, *Probab. Theory Relat. Fields* 74(2), 213-240.

[14] Fukasawa, M. (2010): Realized volatility with stochastic sampling, *Stochastic Process. Appl.* 120, 829-852.

[15] Fukasawa, M. (2011): Discretization error of stochastic integrals, *Ann. Appl. Probab.* 21, 1436-1465.

[16] Fukasawa, M. (2011): Asymptotic analysis for stochastic volatility: martingale expansion, *Finance Stoch.* 15, 635-654.

[17] Fukasawa, M. (2017): Short-time at-the-money skew and rough fractional volatility, *Quant. Finance.* 17(2), 189-198.

[18] Fukasawa, M. and Takabatake, T. (2019): Asymptotically efficient estimators for self-similar stationary Gaussian noises under high frequency observations, *Bernoulli*, forthcoming.

[19] Fukasawa, M. and Takabatake, T. (2019): Supplement to "Asymptotically efficient estimators for self-similar stationary Gaussian noises under high frequency observations", *Bernoulli*, forthcoming.

[20] Garnier, J. and Solna, K. (2017): Correction to Black-Scholes formula due to fractional stochastic volatility, *SIAM J. Financial Math.* 8, 560-588.

[21] Gatheral, J., Jaisson, T. and Rosenbaum, M. (2018): Volatility is rough, *Quant. Finance.* 18(6), 933-949.

[22] Jacod, J. and Protter, P. (1998): Asymptotic error distributions for the Euler methods for stochastic differential equations, *Ann. Probab.* 26(1), 267-307.

[23] Jacod, J. and Shiryaev, A. N. (2003): *Limit Theorems for Stochastic Processes*, 2nd ed., Springer, Berlin.

[24] Kahane, J.P. (1985): Sur le chaos multiplicatif, *Ann. Sci. Math. Québec*, 9, no.2, 105-150.

[25] Karatzas, I. and Shreve, S. E. (1998): *Brownian Motion and Stochastic Calculus*, second edition, Springer.

[26] Mishura, Y. (2008): *Stochastic Calculus for Fractional Brownian Motion and Related Processes*, Springer-Verlag Berlin Heidelberg.

[27] Muzy, J.F., Delour, J. and Bacry, E. (2000): Modelling fluctuations of financial time series: from cascade process to stochastic volatility model, *Eur. Phys. J. B*, 17:537.

[28] Nourdin and Peccati (2012): *Normal approximations with Malliavin calculus: From Stein's method to universality*, Cambridge university press, New York.

[29] Pochart, B. and Bouchaud, J.P. (2002): The skewed multifractal random walk with applications to option smiles, *Quantitative Finance*, vol. 2, issue 4, 303-314.

[30] Rosenbaum, M. (2008): Estimation of the volatility persistence in a discretely observed diffusion model, *Stochastic Process. Appl.* 118, 1434-1462.

[31] Tudor, C. A. (2013): *Analysis of Variations for Self-similar Processes*, Springer.

[32] Velasco, C. and Robinson, P. M. (2000): Whittle Pseudo-Maximum Likelihood Estimation for Nonstationary Time Series, *J. Amer. Statist. Assoc.* 95(452), 1229-1243.

[33] Westphal, R. (2017): Empirical Analysis of High-Frequency Financial Data under the Rough Fractional Stochastic Volatility Model, Master thesis, ETH Zurich.

A   Notation

In this section, we summarize notation used throughout the appendix in this paper.

A.1   Notation of Bilinear Form

Denote by $L^1[-\pi, \pi]$ the set of the Lebesgue integrable functions on $[-\pi, \pi]$. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $k \in L^1[-\pi, \pi]$ be an even function. We define a symmetric bilinear form of $\mathbf{x}$ and $\mathbf{y}$ with a certain symmetric $n \times n$-matrix $\Sigma_n(k)$ by

$$B_n(\mathbf{x}, \mathbf{y}, k) := \frac{1}{2\pi n} \mathbf{x}^{\mathrm{T}} \Sigma_n(k) \mathbf{y},$$

where $\mathbf{x}^{\mathrm{T}}$ denotes the transpose of vector $\mathbf{x}$ and $\Sigma_n(k)$ denotes a symmetric matrix whose (i, j)-element is given by the $(i - j)$th Fourier coefficient of k, denoted by $\widehat{k}(i - j)$, for each $i, j \in \{1, \cdots, n\}$, i.e.

$$\widehat{k}(\tau) := \int_{-\pi}^{\pi} e^{\sqrt{-1}\tau\lambda} k(\lambda) \, \mathrm{d}\lambda, \quad \tau \in \mathbb{Z}.$$

In particular, we denote by $Q_n(\mathbf{x}, k) := B_n(\mathbf{x}, \mathbf{x}, k)$. Note that

$$Q_n(\mathbf{x}, k) = \int_{-\pi}^{\pi} I_n(\lambda, \mathbf{x}) k(\lambda) \, \mathrm{d}\lambda$$

holds, where the periodogram $I_n(\lambda, \mathbf{x})$ is defined in (9).

A.2 Notation of Stochastic Sequences

Set $\Lambda_n := \{1, 2, \cdots, n\}$. Denote by $(\Omega, \mathcal{F}, P)$ a probability space on which the sequence of observations $\{\mathbf{Y}_n\}_{n \in \mathbb{N}}$ is defined and by $\| \cdot \|_p$, $p \in [1, \infty]$, the L^p-norm on the probability space. Furthermore, we denote $\Delta X_t := X_{t+1} - X_t$ for a discrete-time stochastic process $X = \{X_t\}_{t \in \mathbb{N}}$ and $t \in \mathbb{N}$, and

$$\mathbf{Y}_n^\dagger := (Y_1^{n,\dagger}, Y_2^{n,\dagger}, \cdots, Y_n^{n,\dagger}) \ \text{ with } \ Y_t^{n,\dagger} \equiv Y_t^{n,\dagger,\vartheta} := \Delta \left(\log \int_{(t-1)\delta_n}^{t\delta_n} \sigma_u^2 \, du \right),$$

$$\mathbf{G}_n^\dagger := (G_1^{n,\dagger}, G_2^{n,\dagger}, \cdots, G_n^{n,\dagger}) \ \text{ with } \ G_t^{n,\dagger} \equiv G_t^{n,\dagger,\vartheta} := \Delta \left(\frac{1}{\delta_n} \int_{(t-1)\delta_n}^{t\delta_n} \eta W_u^H \, du \right),$$

$$\mathbf{V}_n := (V_1^n, V_2^n, \cdots, V_n^n) \ \text{ with } \ V_t^n \equiv V_t^{n,\vartheta} := \Delta \left(\frac{1}{\delta_n} \int_{(t-1)\delta_n}^{t\delta_n} \log \sigma_u^2 \, du \right),$$

Recall that

$$\mathbf{Y}_n := \left(Y_1^n, Y_2^n, \cdots, Y_n^n \right) \ \text{ with } \ Y_t^n := \Delta \left(\log \int_{(t-1)\delta_n}^{t\delta_n} \sigma_u^2 \, du \right) + \Delta \epsilon_t^n,$$

$$\mathbf{G}_n := \left(G_1^n, G_2^n, \cdots, G_n^n \right) \ \text{ with } \ G_t^n := G_t^{n,\dagger} + \Delta \epsilon_t^n,$$

where $\{\epsilon_t^n\}_{t \in \mathbb{Z}}$ is an i.i.d. sequence independent of W^H and normally distributed with mean 0 and variance $2/m_n$. For each $K \in \mathbb{N}$ and $\mathbf{p} \equiv (p_1, \cdots, p_K) \in \mathbb{N}^K$, denote by $\mathbf{p}! := \prod_{k=1}^K p_k!$, $|\mathbf{p}| := p_1 + \cdots + p_K$ and

$$\mathbf{Z}_n^{\mathbf{p}} := \left(Z_1^{n,\mathbf{p}}, Z_2^{n,\mathbf{p}}, \cdots, Z_{n+1}^{n,\mathbf{p}} \right), \ \ \mathbf{W}_n^{\mathbf{p}} := \left(W_1^{n,\mathbf{p}}, W_2^{n,\mathbf{p}}, \cdots, W_{n+1}^{n,\mathbf{p}} \right),$$

where

$$Z_t^{n,p_k} := \frac{1}{\delta_n} \int_{(t-1)\delta_n}^{t\delta_n} \left(\log \sigma_u^2 - \log \sigma_{(t-1)\delta_n}^2 \right)^{p_k} du, \ \ Z_t^{n,\mathbf{p}} := \prod_{k=1}^K Z_t^{n,p_k},$$

$$W_t^{n,p_k} := \frac{1}{\delta_n} \int_{(t-1)\delta_n}^{t\delta_n} \eta^{p_k} \left(W_u^H - W_{(t-1)\delta_n}^H \right)^{p_k} du, \ \ W_t^{n,\mathbf{p}} := \prod_{k=1}^K W_t^{n,p_k}.$$

Note that for each $t \in \mathbb{Z}$, P-a.s. $\omega \in \Omega$,

$$W_t^{n,\mathbf{p}}(\omega) = \prod_{k=1}^K \int_0^1 \eta^{p_k} \left(W_{(u+t-1)\delta_n}^H(\omega) - W_{(t-1)\delta_n}^H(\omega) \right)^{p_k} du$$

$$= \eta^{|\mathbf{p}|} \int_{[0,1]^K} \prod_{k=1}^K \left(W_{(u_k+t-1)\delta_n}^H(\omega) - W_{(t-1)\delta_n}^H(\omega) \right)^{p_k} d\mathbf{u}, \ \ \mathbf{u} \equiv (u_1, \cdots, u_K), \tag{17}$$

follows from the change of variables and Fubini's theorem. Moreover, the Hölder continuity of the fractional Brownian motion and $\log \sigma^2$ yield that for any $\psi > 0$,

$$\max_{t \in \Lambda_{n+1}} \left\| W_t^{n,\mathbf{p}} \right\|_\infty = o \left(\delta_n^{|\mathbf{p}|H - \psi} \right), \ \ \max_{t \in \Lambda_{n+1}} \left\| Z_t^{n,\mathbf{p}} \right\|_\infty = o \left(\delta_n^{|\mathbf{p}|H - \psi} \right) \ \text{ as } \ n \to \infty. \tag{18}$$

B Approximation of Data

The following proposition gives a precise statement of the approximation (11) in Remark 2.6, which follows from a Taylor expansion of $\mathbf{Y}_n$ around the Gaussian vector $\mathbf{G}_n$ under high frequency observations, i.e. $\delta_n \to 0$.

Proposition B.1. *For any $\psi \in (0, H)$ and $J \in \mathbb{N}$, there exists a positive random variable $M \equiv M(\psi, J, T, \vartheta)$, which is independent of the asymptotic parameter $n \in \mathbb{N}$, such that*

$$\max_{t \in \Lambda_n} \left| Y_t^n - G_t^n - \sum_{j=2}^{J} \sum_{k=1}^{j} \frac{(-1)^{k-1}}{k} \sum_{\mathbf{p} \in \mathbb{N}^k, |\mathbf{p}|=j} \frac{1}{\mathbf{p}!} \Delta W_t^{n,\mathbf{p}} \right| \leq M \cdot \delta_n^{\min\{1, (J+1)H - \psi\}} \tag{19}$$

holds P-a.s. for sufficiently small δ_n.

Note that the lhs of the inequality in Proposition B.1 is dominated as follows:

$$\max_{t \in \Lambda_n} \left| Y_t^n - G_t^n - \sum_{j=2}^{J} \sum_{k=1}^{j} \frac{(-1)^{k-1}}{k} \sum_{\mathbf{p} \in \mathbb{N}^k, |\mathbf{p}|=j} \frac{1}{\mathbf{p}!} \Delta W_t^{n,\mathbf{p}} \right| \tag{20}$$

$$\leq \max_{t \in \Lambda_n} \left| Y_t^{n,\dagger} - V_t^n - \sum_{j=2}^{J} \sum_{k=1}^{j} \frac{(-1)^{k-1}}{k} \sum_{\mathbf{p} \in \mathbb{N}^k, |\mathbf{p}|=j} \frac{1}{\mathbf{p}!} \Delta Z_t^{n,\mathbf{p}} \right|$$

$$+ \max_{t \in \Lambda_n} \left| V_t^n - G_t^{n,\dagger} \right| + \max_{t \in \Lambda_n} \left| \sum_{j=2}^{J} \sum_{k=1}^{j} \frac{(-1)^{k-1}}{k} \sum_{\mathbf{p} \in \mathbb{N}^k, |\mathbf{p}|=j} \frac{1}{\mathbf{p}!} \Delta \left(Z_t^{n,\mathbf{p}} - W_t^{n,\mathbf{p}} \right) \right|.$$

In the rest of this section, we evaluate the asymptotic order of the three terms in the rhs of (20) when $\delta_n \to 0$. At first, we treat the first term in (20) in the following lemma.

Lemma B.2. *For any $\psi \in (0, H)$ and $J \in \mathbb{N}$, there exists a positive random variable $M \equiv M(\psi, J, T, \vartheta)$, which is independent of the asymptotic parameter $n \in \mathbb{N}$, such that*

$$\max_{t \in \Lambda_n} \left| Y_t^{n,\dagger} - V_t^n - \sum_{j=2}^{J} \sum_{k=1}^{j} \frac{(-1)^{k-1}}{k} \sum_{\mathbf{p} \in \mathbb{N}^k, |\mathbf{p}|=j} \frac{1}{\mathbf{p}!} \Delta Z_t^{n,\mathbf{p}} \right| \leq M \cdot \delta_n^{(J+1)H - \psi}$$

holds P-a.s. for sufficiently small δ_n.

Proof. At first, Taylor's theorem yields that any infinitely differentiable function f on an ϵ-open ball $B_\epsilon(a)$ at the point of $a \in \mathbb{R}$ is expanded by

$$f(x) = f(a) + \sum_{j=1}^{J} \frac{f^{(j)}(a)}{j!} x^j + (x-a)^{J+1} \int_0^1 \frac{(1-z)^J}{J!} f^{(J+1)} (a + z(x-a)) \, dz$$

for each $x \in B_\epsilon(a)$ and $J \in \mathbb{N}$. Moreover, if the function f and its derivatives of any order are also continuous on $\overline{B_\epsilon(a)}$, then it holds that

$$\sup_{x \in \overline{B_\epsilon(a)}} \left| \int_0^1 \frac{(1-z)^J}{J!} f^{(J+1)} (a + z(x-a)) \, dz \right| < \infty. \tag{21}$$

17

Therefore, using the Hölder continuity of $\log \sigma^2$, we can derive the following Taylor's expansion:

$$
\begin{aligned}
\log\left[\frac{1}{\delta_n}\int_{(t-1)\delta_n}^{t\delta_n}\sigma_u^2\,du\right] &= \log\sigma_{(t-1)\delta_n}^2 + \log\left[\frac{1}{\delta_n}\int_{(t-1)\delta_n}^{t\delta_n}e^{\log\sigma_u^2-\log\sigma_{(t-1)\delta_n}^2}\,du\right]\\
&= \log\sigma_{(t-1)\delta_n}^2 + \log\left[1 + \sum_{p=1}^{J}\frac{1}{p!}Z_t^{n,p} + o\left(\delta_n^{(J+1)H-\psi}\right)\right]\\
&= \log\sigma_{(t-1)\delta_n}^2 + \frac{1}{\delta_n}\int_{(t-1)\delta_n}^{t\delta_n}\left(\log\sigma_u^2 - \log\sigma_{(t-1)\delta_n}^2\right)\,du + \sum_{p=2}^{J}\frac{1}{p!}Z_t^{n,p}\\
&\quad + \sum_{j=2}^{J}\frac{(-1)^{j-1}}{j}\left\{\sum_{p=1}^{J}\frac{1}{p!}Z_t^{n,p} + o\left(\delta_n^{(J+1)H-\psi}\right)\right\}^j + o\left(\delta_n^{(J+1)H-\psi}\right)\\
&= \frac{1}{\delta_n}\int_{(t-1)\delta_n}^{t\delta_n}\log\sigma_u^2\,du + \sum_{j=2}^{J}\sum_{k=1}^{j}\frac{(-1)^{k-1}}{k}\sum_{\mathbf{p}\in\mathbb{N}^k,|\mathbf{p}|=j}\frac{1}{\mathbf{p}!}Z_t^{n,\mathbf{p}} + o\left(\delta_n^{(J+1)H-\psi}\right). \quad (22)
\end{aligned}
$$

Note that the Hölder continuity property also implies that all reminder terms in the above equality are independent of $t \in \Lambda_n$ and $\omega \in \Omega$ if δ_n is sufficiently small, see also (21). Therefore, the conclusion follows from taking a difference of both sides of (22). $\quad\square$

The second term is also negligible because the following inequality holds.

Lemma B.3. *The following inequality holds:*

$$
\max_{t\in\Lambda_n}\left|V_t^n - G_t^{n,\dagger}\right| \le \left(\sup_{u\in[0,T]}|\kappa_u|\right)\cdot\delta_n.
$$

Finally, we show the negligibility of the third term. In order to achieve this purpose, it suffices to prove that the error between $\mathbf{Z}_n^{\mathbf{p}}$ and $\mathbf{W}_n^{\mathbf{p}}$ is negligible for each $\mathbf{p} \in \mathbb{N}^K$ by using the triangle inequality of $\|\cdot\|_\infty$. Therefore, we show the following result.

Lemma B.4. *For each $\mathbf{p} \equiv (p_1, p_2, \cdots, p_K) \in \mathbb{N}^K$ with $|\mathbf{p}| \ge 2$, the following relation holds for any $\psi \in (0, H)$,*

$$
\max_{t\in\Lambda_{n+1}}\left|Z_t^{n,\mathbf{P}} - W_t^{n,\mathbf{P}}\right| \le \left(2^{|\mathbf{p}|}-1\right)\left(A_{H-\psi,\eta}\vee 1\right)^{(|\mathbf{p}|-1)}\left(\sup_{u\in[0,T]}|\kappa_u|\vee 1\right)^{|\mathbf{p}|}\cdot\delta_n^{1+(|\mathbf{p}|-1)(H-\psi)},
$$

where $A_{H-\psi,\eta} \equiv A_{H-\psi,\eta}(T, \vartheta)$ given by

$$
A_{H-\psi,\eta} := \sup_{s,u\in[0,T]}\frac{\eta|W_u^H - W_s^H|}{|u-s|^{H-\psi}} < \infty.
$$

Proof. At first, consider the case where $K = 1$, i.e. $p \in \mathbb{N}$ with $p \ge 2$. Note that the binomial theorem yields that the integrand of $Z_t^{n,\mathbf{P}}$ is given by

$$
\left(\log\sigma_u^2 - \log\sigma_{(t-1)\delta_n}^2\right)^p = \sum_{j=0}^{p}\frac{\Gamma(p+1)}{\Gamma(p-j+1)\Gamma(j+1)}\eta^{p-j}\left(W_u^H - W_{(t-1)\delta_n}^H\right)^{p-j}\left(\int_{(t-1)\delta_n}^{u}\kappa_s\,ds\right)^j.
$$

Then $Z_t^{n,p}$ is represented by

$$Z_t^{n,p} = \frac{1}{\delta_n} \int_{(t-1)\delta_n}^{t\delta_n} \eta^p \left(W_u^H - W_{(t-1)\delta_n}^H\right)^p \, du + \sum_{j=1}^{p} \frac{\Gamma(p+1)}{\Gamma(p-j+1)\Gamma(j+1)} R_{t,n,j},$$

where

$$R_{t,n,j}^p := \frac{1}{\delta_n} \int_{(t-1)\delta_n}^{t\delta_n} \eta^{p-j} \left(W_u^H - W_{(t-1)\delta_n}^H\right)^{p-j} \left(\int_{(t-1)\delta_n}^{u} \kappa_s \, ds\right)^j \, du.$$

Therefore, the conclusion when $K = 1$ follows from the Hölder continuity of the fractional Brownian motion W^H. Next, we consider the case where $K \geq 2$. Then the multinomial theorem yield that

$$\max_{t \in \Lambda_{n+1}} \left|Z_t^{n,\mathbf{p}} - W_t^{n,\mathbf{p}}\right| = \max_{t \in \Lambda_{n+1}} \left|\prod_{k=1}^{K} \left\{W_t^{n,p_k} + \left(Z_t^{n,p_k} - W_t^{n,p_k}\right)\right\} - W_t^{n,\mathbf{p}}\right|$$

$$\leq \sum_{j_1, j_2, \cdots, j_K} \max_{t \in \Lambda_{n+1}} \left|\prod_{k=1}^{K} \left(W_t^{n,p_k}\right)^{j_i} \left(Z_t^{n,p_k} - W_t^{n,p_k}\right)^{1-j_i}\right|,$$

where the last sum is taken over all $j_1, \cdots, j_K \in \{0,1\}$ satisfying that there exists $i \in \{1, \cdots, K\}$ such that $j_i = 0$. As a result, the conclusion when $K \geq 2$ follows from (18) and the conclusion when $K = 1$. $\qquad\square$

C Asymptotic Decay of Covariance Function for Stationary Process Associated with Some Functionals of Fractional Brownian Motion

In this section, we will show an asymptotic decay of covariance function for the stationary process $\mathbf{W}_n^{\mathbf{p}}$ appeared in the reminder terms of the Taylor approximation given in Proposition B.1. This result plays a key role in order to prove that the reminder terms $\mathbf{W}_n^{\mathbf{p}}$ are asymptotically negligible in the case where the consistency of the adapted Whittle estimator holds. We will state the key result in Section C.1, several preliminary results used in its proof are summarized in Section C.2 and its proof is given in Section C.3.

C.1 Notation and Statement of Key Result

At first, we prepare notation in order to state a general result for Proposition C.1. Denote by $C_{\mathbb{R}}$ a set of real-valued continuous functions on $\mathbb{R}$ and by $\mathcal{B}(C_{\mathbb{R}})$ a Borel σ-algebra on $C_{\mathbb{R}}$ generated by a topology associated with the compact convergence. Let μ_H be the distribution of the two-sided standard fractional Brownian motion with the Hurst parameter $H \in (0, 1]$ on $(C_{\mathbb{R}}, \mathcal{B}(C_{\mathbb{R}}))$, and a continuous shift operator $\theta = \{\theta_u\}_{u \in \mathbb{R}}$ be defined by $\theta_u x_. := x_{.+u} - x_u$ for $(u, x) \in \mathbb{R} \times C_{\mathbb{R}}$. Note that μ_H is θ-invariant, i.e. $\mu_H \circ \theta_u^{-1} = \mu_H$ for each $u \in \mathbb{R}$ since the fractional Brownian motion enjoys the stationary increments property. Moreover, $U = \{U_u\}_{u \in \mathbb{R}}$ denotes the canonical process on $(C_{\mathbb{R}}, \mathcal{B}(C_{\mathbb{R}}))$, i.e. $U_u(x) := x_u$ for each $(u, x) \in \mathbb{R} \times C_{\mathbb{R}}$. Furthermore, for each $\mathbf{p} = (p_1, \cdots, p_K) \in \mathbb{N}^K$, $K \in \mathbb{R}$, and compact set $A_{\mathbf{p}} \subset \mathbb{R}^K$, we define a functional $F^{\mathbf{p}}$ by

$$F^{\mathbf{p}}(x) := \int_{A_{\mathbf{p}}} \prod_{k=1}^{K} x_{u_k}^{p_k} \, du_1 \cdots du_K = \int_{A_{\mathbf{p}}} \prod_{k=1}^{K} \{U_{u_k}(x)\}^{p_k} \, du_1 \cdots du_K,$$

for $x = \{x_u\}_{u \in \mathbb{R}} \in C_{\mathbb{R}}$, and set a stochastic process $G_u^{\mathbf{p}} := F^{\mathbf{p}}(\theta_u)$ for $u \in \mathbb{R}$.

Next, let us recall the following definition, e.g. see Tudor [31], p.172.

Definition C.1. *A filter of length $J \in \mathbb{N}$ and order $r \in \mathbb{N}$ is a $(J+1)$-dimensional vector $\mathbf{a} := \{a_0, a_1, \cdots, a_J\}$ such that for any $k \in \mathbb{N} \cup \{0\}$ with $k < r$,*

$$\sum_{j=0}^{J} a_j j^k = 0, \tag{23}$$

where we use $0^0 := 1$ for convenience, and

$$\sum_{j=0}^{J} a_j j^r \neq 0. \tag{24}$$

Moreover, we also call $\mathbf{a} = \{a_0, a_1, \cdots, a_J\}$ as a filter of length J and order 0 if it satisfies (24) for $r = 0$.

Remark C.2. *For any filter $\mathbf{a} = \{a_0, a_1, \cdots, a_J\}$ of order $r \in \mathbb{N}$, the property (23) yield that for any $k \in \mathbb{N} \cup \{0\}$ with $k < 2r$,*

$$\sum_{i,j=0}^{J} a_i a_j (j-i)^k = 0. \tag{25}$$

For a filter $\mathbf{a} = \{a_0, a_1, \cdots, a_J\}$ and a stochastic process $X = \{X_u\}_{u \in \mathbb{R}}$, we define

$$\Delta_{\mathbf{a}} X_u := \sum_{j=0}^{J} a_j X_{u-j}, \quad u \in \mathbb{R}. \tag{26}$$

For example, if we set $\mathbf{a} = (a_0, a_1)$ with $a_0 = -1$, $a_1 = 1$, then $\mathbf{a}$ is a filter of length 1 and order 1, and $\Delta_{\mathbf{a}} X_\cdot = X_\cdot - X_{\cdot - 1}$.

Finally, we will state a main result in this section.

Proposition C.3. *Let $\mathbf{a}$ be a filter of length $J \in \mathbb{N}$ and order $r \in \mathbb{N} \cup \{0\}$. Then for any $\mathbf{p} \in \mathbb{N}^K$ with $K \in \mathbb{N}$, the stochastic process $\{\Delta_{\mathbf{a}} G_u^{\mathbf{p}}\}_{u \in \mathbb{R}}$ is stationary and for any $\mathbf{p} \in \mathbb{N}^K$, $\mathbf{q} \in \mathbb{N}^L$ with $K, L \in \mathbb{N}$ and $u \in \mathbb{R}$,*

$$\mathrm{Cov}_{\mu_H} \left[\Delta_{\mathbf{a}} G_u^{\mathbf{p}}, \Delta_{\mathbf{a}} G_{u+\tau}^{\mathbf{q}} \right] = O\left(|\tau|^{2H-2-2r} \right) \quad \text{as } |\tau| \to \infty. \tag{27}$$

As a corollary of Proposition C.3, we can obtain the following result from the self-similarity property of the fractional Brownian motion.

Proposition C.4. *For any $\mathbf{p} \in \mathbb{N}^K$, $\mathbf{q} \in \mathbb{N}^L$ with $K, L \in \mathbb{N}$, the stochastic process $\{W_t^{n, \mathbf{p}}\}_{t \in \mathbb{Z}}$ is stationary for each $n \in \mathbb{N}$ and the following relation holds for any $t \in \mathbb{R}$:*

$$\sup_{n \in \mathbb{N}} \left| \delta_n^{-(|\mathbf{p}|+|\mathbf{q}|)H} \mathrm{Cov} \left[\Delta W_t^{n, \mathbf{p}}, \Delta W_{t+\tau}^{n, \mathbf{q}} \right] \right| = O\left(|\tau|^{2H-4} \right) \quad \text{as } |\tau| \to \infty.$$

C.2   Preliminary Results

We summarize several preliminary results used in the proof of Proposition C.3 in this subsection. The first result is proven in the similar way to that in Billingsley [6], p.230-231.

Proposition C.5. *Let $A \in \mathcal{B}(\mathbb{R}^K)$ with $K \in \mathbb{N}$ and f be a measurable function on $\mathbb{R}^K$. For $x = \{x_u\}_{u \in \mathbb{R}} \in C_{\mathbb{R}}$, we define a functional F by*

$$F(x) := \int_A f(x_{u_1}, \cdots, x_{u_K}) \, du_1 \cdots du_K. \tag{28}$$

Then the functional F is $\mathcal{B}(C_{\mathbb{R}})$-measurable if $(u_1, \cdots, u_K) \mapsto f(x_{u_1}, \cdots, x_{u_K})$ is integrable on A. Furthermore, if A is compact and f is continuous, then $x \mapsto F(x)$ is also continuous.

Since the shift operator θ is continuous and μ_H is θ-invariant, we can obtain the following result using Proposition C.5.

Corollary C.6. *Let us consider a functional F of the form (28) with a continuous function f and a compact set $A \in \mathcal{B}(\mathbb{R}^K)$. Then a stochastic process $G = \{G_u\}_{u \in \mathbb{R}}$ defined by $G_u(x) := F(\theta_u x)$ for $(u, x) \in \mathbb{R} \times C_{\mathbb{R}}$ is continuous and strong stationary on the probability space $(C_{\mathbb{R}}, \mathcal{B}(C_{\mathbb{R}}), \mu_H)$.*

The following result is a consequence of the well-know Wick formula which expresses the higher moments of centered multivariate Gaussian vectors in terms of its second moments, e.g. see Nourdin and Peccati [28]. Given a finite set b the number of which is even, we denote by $\mathcal{P}(b)$ the class of all partitions of b such that each block of a partition π contains exactly two elements, and recall $\Lambda_M := \{1, 2, \cdots, M\}$.

Lemma C.7. *For any $K_0, L_0 \in \mathbb{N}$ and $(K_0 + L_0)$-dimensional centered Gaussian vector $(X_1, \cdots, X_{K_0+L_0})$,*

$$\mathrm{Cov}\left[\prod_{k=1}^{K_0} X_k, \prod_{\ell=1}^{L_0} X_{K_0+\ell}\right] = \begin{cases} \sum_{\pi = \{\{k_1,\ell_1\},\cdots,\{k_{M_0},\ell_{M_0}\}\} \in \mathcal{P}_0(\Lambda_{2M_0})} \mathrm{Cov}[X_{k_1}, X_{\ell_1}] \cdots \mathrm{Cov}[X_{k_{M_0}}, X_{\ell_{M_0}}] & \text{if } K_0 + L_0 \text{ is even,} \\ 0 & \text{if } K_0 + L_0 \text{ is odd,} \end{cases}$$

where $M_0 := (K_0+L_0)/2$ and $\mathcal{P}_0(\Lambda_{2M_0})$ denotes the subset of $\mathcal{P}(\Lambda_{2M_0})$ whose elements are partitions $\pi = \{\{k_1, \ell_1\}, \cdots, \{k_{M_0}, \ell_{M_0}\}\} \in \mathcal{P}(\Lambda_{2M_0})$ such that there exists $m \in \Lambda_{M_0}$ satisfying $k_m \le K_0 < \ell_m$.

Proof. Let us consider only the case that both K_0 and L_0 are even since the other cases are trivial from the Wick formula. Since K_0 and L_0 are even, the Wick formula yields that

$$\begin{aligned} E\left[\prod_{k=1}^{K_0+L_0} X_k\right] &= \sum_{\{\{k_1,\ell_1\},\cdots,\{k_{M_0},\ell_{M_0}\}\} \in \mathcal{P}(\Lambda_{2M_0})} \mathrm{Cov}[X_{k_1}, X_{\ell_1}] \cdots \mathrm{Cov}[X_{k_{M_0}}, X_{\ell_{M_0}}] \\ &= \left(\sum_{\{\{k_1,\ell_1\},\cdots,\{k_{K_0/2},\ell_{K_0/2}\}\} \in \mathcal{P}(\Lambda_{K_0})} \mathrm{Cov}[X_{k_1}, X_{\ell_1}] \cdots \mathrm{Cov}[X_{k_{K_0/2}}, X_{\ell_{K_0/2}}]\right) \\ &\quad \times \left(\sum_{\{\{k_1,\ell_1\},\cdots,\{k_{L_0/2},\ell_{L_0/2}\}\} \in \mathcal{P}(\Lambda_{L_0})} \mathrm{Cov}[X_{K_0+k_1}, X_{K_0+\ell_1}] \cdots \mathrm{Cov}[X_{K_0+k_{L_0/2}}, X_{K_0+\ell_{L_0/2}}]\right) \\ &\quad + \sum_{\{\{k_1,\ell_1\},\cdots,\{k_{M_0},\ell_{M_0}\}\} \in \mathcal{P}_0(\Lambda_{2M_0})} \mathrm{Cov}[X_{k_1}, X_{\ell_1}] \cdots \mathrm{Cov}[X_{k_{M_0}}, X_{\ell_{M_0}}] \\ &= E\left[\prod_{k=1}^{K_0} X_k\right] E\left[\prod_{\ell=1}^{L_0} X_{K_0+\ell}\right] + \sum_{\{\{k_1,\ell_1\},\cdots,\{k_{M_0},\ell_{M_0}\}\} \in \mathcal{P}_0(\Lambda_{2M_0})} \mathrm{Cov}[X_{k_1}, X_{\ell_1}] \cdots \mathrm{Cov}[X_{k_{M_0}}, X_{\ell_{M_0}}]. \end{aligned}$$

Therefore, the conclusion follows. $\square$

C.3 Proof of Proposition C.3

Before proving Proposition C.3, we will show the following two lemmas. Denote by $\gamma_{s,u}(\tau) := \mathrm{Cov}_{\mu_H}[U_s(\theta_0), U_u(\theta_\tau)]$ for $s, u, \tau \in \mathbb{R}$.

Lemma C.8. *For each $s, u \in \mathbb{R}$, $\tau \mapsto \gamma_{s,u}(\tau)$ is infinitely differentiable a.e. and, for any $k \in \mathbb{N} \cup \{0\}$ and compact set $A \subset \mathbb{R}$, its kth derivative satisfies*

$$\sup_{s,u \in A} \left|\frac{\partial^k \gamma_{s,u}}{\partial \tau^k}(\tau)\right| = O(|\tau|^{2H-2-k}) \text{ as } |\tau| \to \infty.$$

Proof. Fix $s, u \in \mathbb{R}$ and a compact set $A \subset \mathbb{R}$. Since μ_H is a distribution of the two-sided standard fractional Brownian motion with the Hurst parameter H, we have

$$\gamma_{s,u}(\tau) = -\frac{1}{2}\left(|\tau + u - s|^{2H} - |\tau + u|^{2H} - |\tau - s|^{2H} + |\tau|^{2H}\right), \quad \tau \in \mathbb{R}.$$

As a result, the first assertion is obvious and for any $k \in \mathbb{N}$, we obtain

$$\frac{\partial^k \gamma_{s,u}}{\partial \tau^k}(\tau) = -\frac{(\operatorname{sgn}(\tau))^k}{2} \prod_{\ell=0}^{k-1}(2H - \ell)\left(|\tau + u - s|^{2H-k} - |\tau + u|^{2H-k} - |\tau - s|^{2H-k} + |\tau|^{2H-k}\right) \tag{29}$$

if $|\tau|$ is sufficiently large, where $\operatorname{sgn}(\cdot)$ denotes the sign function defined by

$$\operatorname{sgn}(\tau) = \begin{cases} 1 & \tau \geq 0, \\ -1 & \tau < 0. \end{cases}$$

Then the second assertion follows from (29) because Taylor's theorem yields that for any $L \in \mathbb{N}$,

$$|\tau + u - s|^{2H-k} - |\tau + u|^{2H-k} - |\tau - s|^{2H-k} + |\tau|^{2H-k}$$

$$= |\tau|^{2H-k}\left\{\left(1 + \frac{u-s}{\tau}\right)^{2H-k} - \left(1 + \frac{u}{\tau}\right)^{2H-k} - \left(1 + \frac{-s}{\tau}\right)^{2H-k} + 1\right\}$$

$$= |\tau|^{2H-k} \sum_{\ell_0=1}^{L} \frac{1}{\ell_0!}\left\{\prod_{\ell=0}^{\ell_0-1}(2H - k - \ell)\right\}\left\{(u-s)^{\ell_0} - (-s)^{\ell_0} + u^{\ell_0}\right\}\tau^{-\ell_0} + o(|\tau|^{2H-k-L})$$

as $|\tau| \to \infty$ uniformly in $s, u \in A$ and $(u-s)^{\ell_0} - (-s)^{\ell_0} + u^{\ell_0} = 0$ for $\ell_0 = 1$. $\qquad\square$

Lemma C.9. *Let* $\mathbf{a} = (a_0, a_1, \cdots, a_J)$ *be a filter of length* $J \in \mathbb{N}$ *and order* $r \in \mathbb{N} \cup \{0\}$. *For any compact set* $A \subset \mathbb{R}$ *and* $\mathbf{p} = (p_1, \cdots, p_K) \in \mathbb{N}^K$, $\mathbf{q} = (q_1, \cdots, q_L) \in \mathbb{N}^L$ *with* $K, L \in \mathbb{N}$,

$$\sup_{s_1, u_1, \cdots, s_v, u_v \in A}\left|\sum_{i,j=0}^{J} a_i a_j \operatorname{Cov}_{\mu_H}\left[\prod_{k=1}^{K}\{U_{s_k}(\theta_i)\}^{p_k}, \prod_{\ell=1}^{L}\{U_{u_\ell}(\theta_{j+\tau})\}^{q_\ell}\right]\right| = O\left(|\tau|^{2H-2-2r}\right) \text{ as } |\tau| \to \infty.$$

Proof. By using Lemma C.7 in the case that $K_0 := |\mathbf{p}|$, $L_0 := |\mathbf{q}|$ and $(K_0 + L_0)$-dimensional centered Gaussian vector $\mathbf{X} \equiv (X_1, \cdots, X_{K_0+L_0})$ given by

$$\mathbf{X} := (\underbrace{U_{s_1}(\theta_i), \cdots, U_{s_1}(\theta_i)}_{p_1 \text{ times}}, \cdots, \underbrace{U_{s_K}(\theta_i), \cdots, U_{s_K}(\theta_i)}_{p_K \text{ times}}, \underbrace{U_{u_1}(\theta_{j+\tau}), \cdots, U_{u_1}(\theta_{j+\tau})}_{q_1 \text{ times}}, \cdots, \underbrace{U_{u_L}(\theta_{j+\tau}), \cdots, U_{u_L}(\theta_{j+\tau})}_{q_L \text{ times}}),$$

it suffices to prove that for any compact set $A \subset \mathbb{R}$ and $v \in \mathbb{N}$,

$$\sup_{s_1, u_1, \cdots, s_v, u_v \in A}\left|\sum_{i,j=0}^{J} a_i a_j \prod_{w=1}^{v} \operatorname{Cov}_{\mu_H}\left[U_{s_w}(\theta_i), U_{u_w}(\theta_{j+\tau})\right]\right| = O\left(|\tau|^{2H-2-2r}\right) \text{ as } |\tau| \to \infty \tag{30}$$

22

since the stationary increments property of the fractional Brownian motion implies

$$\text{Cov}_{\mu_H}\left[U_{s_1}, U_{s_2}\right] = \text{Cov}_{\mu_H}\left[U_{s_1}(\theta_i), U_{s_2}(\theta_i)\right],$$
$$\text{Cov}_{\mu_H}\left[U_{u_1}, U_{u_2}\right] = \text{Cov}_{\mu_H}\left[U_{u_1}(\theta_{j+\tau}), U_{u_2}(\theta_{j+\tau})\right]$$

for any $s_1, s_2, u_1, u_2 \in \mathbb{R}$.

Fix a compact set $A \subset \mathbb{R}$ and recall $\gamma_{s,u}(\tau) := \text{Cov}_{\mu_H}[U_s(\theta_0), U_u(\theta_\tau)]$. Since Taylor's theorem and Lemma C.8 yield that for any $K \in \mathbb{N}$,

$$\sup_{\substack{s,u\in A \\ i,j=0,\cdots,J}} \left| \gamma_{s,u}(\tau + (j-i)) - \sum_{k=0}^{K} \frac{(j-i)^k}{k!} \frac{\partial^k \gamma_{s,u}}{\partial \tau^k}(\tau) \right| = o\left(|\tau|^{2H-2-K}\right) \text{ as } |\tau| \to \infty, \tag{31}$$

(30) in the case of $v = 1$ follows from (25) if we take $K \in \mathbb{N}$ satisfying $K \geq 2r$. Moreover, the Taylor approximation (31), the multinomial theorem and Lemma C.8 yield that

$$\sup_{\substack{s_1,u_1,\cdots,s_v,u_v\in A \\ i,j=0,\cdots,J}} \left| \prod_{w=1}^{v} \gamma_{s_w,u_w}(\tau + (j-i)) - \sum_{k_1,\cdots,k_v=0}^{K} \frac{(j-i)^{k_1+\cdots+k_v}}{k_1!\cdots k_v!} \prod_{w=1}^{v} \frac{\partial^{k_w} \gamma_{s_w,u_w}}{\partial \tau^{k_w}}(\tau) \right| = o\left(|\tau|^{2H-2-K}\right) \tag{32}$$

as $|\tau| \to \infty$, and (25) and Lemma C.8 yield that

$$\sup_{s_1,u_1,\cdots,s_v,u_v\in A} \left| \sum_{i,j=0}^{J} a_i a_j \frac{(j-i)^{k_1+\cdots+k_v}}{k_1!\cdots k_v!} \prod_{w=1}^{v} \frac{\partial^{k_w} \gamma_{s_w,u_w}}{\partial \tau^{k_w}}(\tau) \right| \tag{33}$$

$$\begin{cases} = 0 & \text{if } \sum_{w=1}^{v} k_w < 2r, \\ = O\left(|\tau|^{\sum_{w=1}^{v}(2H-2-k_w)}\right) \text{ as } |\tau| \to \infty & \text{if } \sum_{w=1}^{v} k_w \geq 2r. \end{cases}$$

Then (30) in the case of $v \geq 2$ follows from (32) and (33) if we take $K \in \mathbb{N}$ satisfying $K \geq 2r$. Therefore, we finish the proof. □

Proof of Proposition C.3. Since $G^{\mathbf{p}}$ is stationary from Corollary C.6, the bilinearity of covariance functions and Fubini's theorem yield that

$$\text{Cov}_{\mu_H}\left[\Delta_{\mathbf{a}} G_u^{\mathbf{p}}, \Delta_{\mathbf{a}} G_{u+\tau}^{\mathbf{q}}\right] = \int_{A_{\mathbf{p}}\times A_{\mathbf{q}}} \sum_{i,j=0}^{J} a_i a_j \text{Cov}_{\mu_H}\left[\prod_{k=1}^{K}\{U_{s_k}(\theta_i)\}^{p_k}, \prod_{\ell=1}^{L}\{U_{u_\ell}(\theta_{j+\tau})\}^{q_\ell}\right] ds_1\cdots ds_K du_1\cdots du_L.$$

Therefore, the conclusion follows from the above equality and Lemma C.9. □

D Approximating Spectral Density of Data

D.1 Spectral Density of Stationary Gaussian Sequence $\{G_t^{n,\dagger}\}_{t\in\mathbb{Z}}$

Recall that a spectral density of the stationary Gaussian sequence $\{G_t^{n,\dagger}\}_{t\in\mathbb{Z}}$, which is obtained by [18], is characterized by

$$\text{Cov}\left[G_t^{n,\dagger}, G_{t+\tau}^{n,\dagger}\right] = \frac{\eta^2\delta_n^{2H}}{2}(|\tau+2|^{2H+2} - 4|\tau+1|^{2H+2}$$
$$+ 6|\tau|^{2H+2} - 4|\tau-1|^{2H+2} + |\tau-2|^{2H+2})$$
$$= \int_{-\pi}^{\pi} e^{\sqrt{-1}\tau\lambda}\eta^2\delta_n^{2H}f_H(\lambda)\,\mathrm{d}\lambda,$$

where

$$f_H(\lambda) := C_H\{2(1-\cos\lambda)\}^2 \sum_{j\in\mathbb{Z}} \frac{1}{|\lambda + 2\pi j|^{3+2H}}$$

with $C_H := (2\pi)^{-1}\Gamma(2H+1)\sin(\pi H)$. The following Lemma shows that the stationary Gaussian sequence $\{G_t^{n,\dagger}\}_{t\in\mathbb{Z}}$ satisfies Assumption 1 in [18], see Section 4.2 in [18].

Lemma D.1. *The spectral density $f(\lambda, H)$ satisfies the following relations.*

(1) *For any $H \in \Theta$, $\lambda \mapsto f(\lambda, H)$, $\lambda \in [-\pi, \pi]/\{0\}$, is a non-negative integrable even function with 2π-periodicity. Moreover, it satisfies that*
$$f \in C^{3,1}\left(\Theta \times [-\pi,\pi]/\{0\}\right).$$

(2) *If (H_1, η_1) and (H_2, η_2) are distinct elements of $\Theta \times \Sigma$, a set $\{\lambda \in [-\pi, \pi] : \eta_1 f(\lambda, H_1) \neq \eta_2 f(\lambda, H_2)\}$ has a positive Lebesgue measure.*

(3) *Let $\alpha(H) := 2H - 1$ with $H \in (0,1)$. There exist constants $c_1, c_2 > 0$ and for any $\iota > 0$, there exists a constant $c_{3,\iota}$, which only depends on ι, such that the following conditions hold for every $(H, \lambda) \in \Theta \times [-\pi, \pi]\backslash\{0\}$.*

    (a) $c_1|\lambda|^{-\alpha(H)} \leq f(\lambda, H) \leq c_2|\lambda|^{-\alpha(H)}$.

    (b) *For any $j \in \{1, 2, 3\}$,*

$$\left|\frac{\partial^j}{\partial H^j}f(\lambda, H)\right| \leq c_{3,\iota}|\lambda|^{-\alpha(H)-\iota}, \quad \left|\frac{\partial^{j+1}}{\partial\lambda\partial H^j}f(\lambda, H)\right| \leq c_{3,\iota}|\lambda|^{-\alpha(H)-1-\iota}.$$

D.2 Spectral Density of Stationary Gaussian Sequence $\{G_t^n\}_{t\in\mathbb{Z}}$

We derive a spectral density of the stationary sequence $\{G_t^n\}_{t\in\mathbb{Z}}$ in this subsection. Since $\{\epsilon_t^n\}_{t\in\mathbb{Z}}$ is an i.i.d. sequence, $\{\Delta\epsilon_t^n\}_{t\in\mathbb{Z}}$ is a MA(1) process and its auto-covariance function is given by

$$\gamma_n(\tau) := \text{Cov}\left[\Delta\epsilon_t^n, \Delta\epsilon_{t+\tau}^n\right] = \begin{cases} 4/m_n & (\tau = 0) \\ -2/m_n & (|\tau| = 1) \\ 0 & (\text{otherwise}) \end{cases}.$$

Then its spectral density ℓ_n is given by the Fourier series

$$\ell_n(\lambda) := \frac{1}{2\pi}\sum_{\tau\in\mathbb{Z}}\gamma_n(\tau)e^{\sqrt{-1}\tau\lambda} = \frac{2}{m_n}\ell(\lambda), \quad \text{where } \ell(\lambda) := \frac{1}{\pi}(1-\cos\lambda), \quad \lambda \in [-\pi, \pi].$$

Since $\{G_t^{n,\dagger}\}_{t\in\mathbb{Z}}$ and $\{\epsilon_t^n\}_{t\in\mathbb{Z}}$ are independent, the covariance function of $\{G_t^n\}_{t\in\mathbb{Z}}$ is characterized by

$$\mathrm{Cov}\left[G_t^n, G_{t+\tau}^n\right] = \mathrm{Cov}\left[G_t^{n,\dagger}, G_{t+\tau}^{n,\dagger}\right] + \mathrm{Cov}\left[\Delta\epsilon_t^n, \Delta\epsilon_{t+\tau}\right] = \int_{-\pi}^{\pi} e^{\sqrt{-1}\tau\lambda} f_{\vartheta}^n(\lambda)\, \mathrm{d}\lambda,$$

where the spectral density f_{ϑ}^n is given by

$$f_{\vartheta}^n(\lambda) \equiv f(\lambda, H, \eta, n) := \eta^2 \delta_n^{2H} f_H(\lambda) + \frac{2}{m_n}\ell(\lambda), \quad \lambda \in [-\pi, \pi], \quad \vartheta = (H, \eta).$$

E Extension of Some Results in Fox and Taqqu [12, 13]

We will show several extended lemmas and theorem developed in Fox and Taqqu [12, 13] in the case where functions appeared in their results depend on the asymptotic parameter $n \in \mathbb{N}$. They can be easily proven in the similar way to the corresponding results in Fox and Taqqu [12, 13]; we will however give their concise proofs in Section E.1 and Section E.2 for convenience. The following two results are extensions of Lemma 4 and Lemma 5 in [12] which show an asymptotic decay of the Fourier coefficient.

Lemma E.1 (cf.Lemma 4 and Lemma 5 in [12]). *Let $\beta \in (-1, 0) \cup (0, 1)$ and $n \in \mathbb{N}$. Suppose a sequence of 2π-periodic functions $k^n : \mathbb{R} \to [-\infty, \infty]$, $n \in \mathbb{N}$, satisfies the following conditions:*

(1) *If $\beta \in (0, 1)$, k^n is continuously differentiable on $[-\pi, \pi]\setminus\{0\}$ for each $n \in \mathbb{N}$ and*

$$\sup_{n\in\mathbb{N}, \lambda\in[-\pi,\pi]\setminus\{0\}} |\lambda|^{\beta} |k^n(\lambda)| < \infty, \quad \sup_{n\in\mathbb{N}, \lambda\in[-\pi,\pi]\setminus\{0\}} |\lambda|^{\beta+1} \left|\frac{\partial k^n}{\partial\lambda}(\lambda)\right| < \infty.$$

(2) *If $\beta \in (-1, 0)$, k^n is integrable and twice continuously differentiable on $[-\pi, \pi]\setminus\{0\}$ for each $n \in \mathbb{N}$ and*

$$\sup_{n\in\mathbb{N}, \lambda\in[-\pi,\pi]\setminus\{0\}} |\lambda|^{\beta+1} \left|\frac{\partial k^n}{\partial\lambda}(\lambda)\right| < \infty, \quad \sup_{n\in\mathbb{N}, \lambda\in[-\pi,\pi]\setminus\{0\}} |\lambda|^{\beta+2} \left|\frac{\partial^2 k^n}{\partial\lambda^2}(\lambda)\right| < \infty.$$

Then the sequence of the Fourier coefficients $\widehat{k^n}(\tau)$, $\tau \in \mathbb{Z}$, satisfies

$$\sup_{n\in\mathbb{N}} \left|\widehat{k^n}(\tau)\right| = O\left(|\tau|^{\beta-1}\right) \quad as \quad |\tau| \to \infty.$$

Lemma E.2. *Suppose a sequence of 2π-periodic functions $k^n : \mathbb{R} \to [-\infty, \infty]$, $n \in \mathbb{N}$, is continuously differentiable on $[-\pi, \pi]\setminus\{0\}$ for each $n \in \mathbb{N}$ and*

$$\sup_{n\in\mathbb{N}, \lambda\in[-\pi,\pi]\setminus\{0\}} |k^n(\lambda)| < \infty, \quad \sup_{n\in\mathbb{N}, \lambda\in[-\pi,\pi]\setminus\{0\}} |\lambda| \left|\frac{\partial k^n}{\partial\lambda}(\lambda)\right| < \infty.$$

Then the sequence of the Fourier coefficients $\widehat{k^n}(\tau)$, $\tau \in \mathbb{Z}$, satisfies

$$\sup_{n\in\mathbb{N}} \left|\widehat{k^n}(\tau)\right| = O\left(|\tau|^{-1} \log|\tau|\right) \quad as \quad |\tau| \to \infty.$$

The following result is an extension of Theorem 1 in [13] in the case where functions appeared in Theorem 1 in [13] depend on the asymptotic parameter $n \in \mathbb{N}$; they however have the same asymptotic behavior at the origin as that assumed in Theorem 1 in [13] uniformly to the asymptotic parameter $n \in \mathbb{N}$ and they uniformly converge to some functions almost everywhere as $n \to \infty$.

Theorem E.3 (cf. Theorem 1 in [13]). *Let $\alpha_1, \alpha_2 < 1$ and $p \in \mathbb{N}$. Suppose sequences of even functions $k_1^n, k_2^n :$ $[-\pi, \pi] \to [-\infty, \infty]$ satisfy the following two conditions:*

(1) *The following relations hold:*

$$\sup_{n \in \mathbb{N}, \lambda \in [-\pi,\pi] \setminus \{0\}} |\lambda|^{\alpha_1} \left|k_1^n(\lambda)\right| < \infty, \qquad \sup_{n \in \mathbb{N}, \lambda \in [-\pi,\pi] \setminus \{0\}} |\lambda|^{\alpha_2} \left|k_2^n(\lambda)\right| < \infty.$$

(2) *There exist functions $k_1, k_2 : [-\pi, \pi] \to [-\infty, \infty]$ such that*

$$\lim_{n \to \infty} \operatorname*{ess\,sup}_{\lambda \in [-\pi,\pi]} |k_1^n(\lambda) - k_1(\lambda)| = 0, \quad \lim_{n \to \infty} \operatorname*{ess\,sup}_{\lambda \in [-\pi,\pi]} |k_2^n(\lambda) - k_2(\lambda)| = 0.$$

   Moreover, the discontinuities of k_1 and k_2 have the Lebesgue measure 0.

Under the above conditions, we have

(*a*) *If $p(\alpha_1 + \alpha_2) < 1$,*
$$\lim_{n \to \infty} \frac{1}{n} \operatorname{Tr} \left[\left(\Sigma_n(k_1^n) \Sigma_n(k_2^n) \right)^p \right] = (2\pi)^{2p-1} \int_{-\pi}^{\pi} [k_1(\lambda) k_2(\lambda)]^p \, d\lambda.$$

(*b*) *If $p(\alpha_1 + \alpha_2) \geq 1$, then for any $\psi > 0$,*
$$\operatorname{Tr} \left[\left(\Sigma_n(k_1^n) \Sigma_n(k_2^n) \right)^p \right] = o \left(n^{p(\alpha_1 + \alpha_2) + \psi} \right) \quad as \ n \to \infty.$$

E.1 Proof of Lemma E.1 and Lemma E.2

Proof of Lemma E.1 in Case (1). Consider the case of $\beta \in (0, 1)$. Let $\tau \in \mathbb{Z} \setminus \{0\}$. Since k^n is 2π-periodic, we have

$$\widehat{k^n}(\tau) = \int_{-\pi+\pi/|\tau|}^{\pi+\pi/|\tau|} e^{\sqrt{-1}\tau\lambda} k^n(\lambda) \, d\lambda$$

$$= -\int_{-\pi+\pi/|\tau|}^{\pi+\pi/|\tau|} e^{\sqrt{-1}\tau(\lambda-\pi/|\tau|)} k^n(\lambda) \, d\lambda = -\int_{-\pi}^{\pi} e^{\sqrt{-1}\tau\lambda} k^n\left(\lambda + \frac{\pi}{|\tau|}\right) d\lambda.$$

As a result, we obtain

$$2 \left|\widehat{k^n}(\tau)\right| = \left| \int_{-\pi}^{\pi} e^{\sqrt{-1}\tau\lambda} \left[k^n(\lambda) - k^n\left(\lambda + \frac{\pi}{|\tau|}\right) \right] d\lambda \right|$$

$$\leq \int_{-\pi}^{\pi} \left| k^n(\lambda) - k^n\left(\lambda + \frac{\pi}{|\tau|}\right) \right| d\lambda = \int_{-\pi}^{-2\pi/|\tau|} + \int_{-2\pi/|\tau|}^{\pi/|\tau|} + \int_{\pi/|\tau|}^{\pi}. \tag{34}$$

The assumption implies that

$$c_1 := \sup_{n \in \mathbb{N}, \lambda \in [-\pi,\pi] \setminus \{0\}} \left\{ |\lambda|^\beta |k^n(\lambda)| + |\lambda|^{\beta+1} \left| \frac{\partial k^n}{\partial \lambda}(\lambda) \right| \right\} < \infty.$$

By the mean value theorem,

$$
\int_{-\pi}^{-2\pi/|\tau|} \left| k^n(\lambda) - k^n\left(\lambda + \frac{\pi}{|\tau|}\right) \right| d\lambda \le c_1 \frac{\pi}{|\tau|} \int_{-\pi}^{-2\pi/|\tau|} \left| \lambda + \frac{\pi}{|\tau|} \right|^{-\beta-1} d\lambda
$$

$$
= c_1 \frac{\pi}{|\tau|} \int_{-\pi+\pi/|\tau|}^{-\pi/|\tau|} |\lambda|^{-\beta-1} d\lambda
$$

$$
= c_1 \pi |\tau|^{\beta-1} \int_{\pi}^{(|\tau|-1)\pi} \lambda^{-\beta-1} d\lambda = O\left(|\tau|^{\beta-1}\right)
$$

as $|\tau| \to \infty$. Note that $\beta > 0$ is necessary to obtain the last asymptotic behavior. A similar argument shows that

$$
\sup_{n\in\mathbb{N}} \int_{\pi/|\tau|}^{\pi} \left| k^n(\lambda) - k^n\left(\lambda + \frac{\pi}{|\tau|}\right) \right| d\lambda = O\left(|\tau|^{\beta-1}\right) \text{ as } |\tau| \to \infty.
$$

We also have

$$
\int_{-2\pi/|\tau|}^{\pi/|\tau|} \left| k^n(\lambda) - k^n\left(\lambda + \frac{\pi}{|\tau|}\right) \right| d\lambda \le \int_{-2\pi/|\tau|}^{\pi/|\tau|} |k^n(\lambda)| \, d\lambda + \int_{-2\pi/|\tau|}^{\pi/|\tau|} \left| k^n\left(\lambda + \frac{\pi}{|\tau|}\right) \right| d\lambda
$$

$$
\le c_1 \int_{-2\pi/|\tau|}^{\pi/|\tau|} |\lambda|^{-\beta} \, d\lambda + c_1 \int_{-2\pi/|\tau|}^{\pi/|\tau|} \left| \lambda + \frac{\pi}{|\tau|} \right|^{-\beta} d\lambda
$$

$$
= 2c_1 \int_{-2\pi/|\tau|}^{\pi/|\tau|} |\lambda|^{-\beta} \, d\lambda = O\left(|\tau|^{\beta-1}\right) \text{ as } |\tau| \to \infty.
$$

This completes the proof in the case of $\beta \in (0,1)$. □

Proof of Lemma E.1 in Case (2). Consider the case of $\beta \in (-1,0)$. Let $\tau \in \mathbb{Z}\backslash\{0\}$. Since the continuity of k^n on $[-\pi, \pi]\backslash\{0\}$ implies $k^n(\pi) = k^n(-\pi)$, the integration by parts formula yields

$$
\widehat{k^n}(\tau) = -\frac{1}{\sqrt{-1}\tau} \int_{-\pi}^{\pi} e^{\sqrt{-1}\tau\lambda} \frac{\partial k^n}{\partial \lambda}(\lambda) \, d\lambda.
$$

Moreover, since the derivative $\frac{\partial k^n}{\partial \lambda}$ is also 2π-periodic from the assumption, the argument in the case (1) can be applied so that we obtain

$$
\sup_{n\in\mathbb{N}} \left| \widehat{k^n}(\tau) \right| = \frac{1}{|\tau|} O\left(|\tau|^{(\beta-1)-1}\right) = O\left(|\tau|^{\beta-1}\right) \text{ as } |\tau| \to \infty.
$$

This completes the proof in the case of $\beta \in (-1,0)$. □

Proof of Lemma E.2. The same argument in Lemma E.1 shows the inequality (34). The assumption implies that

$$
c_2 := \sup_{n\in\mathbb{N}, \lambda\in[-\pi,\pi]\backslash\{0\}} \left\{ |k^n(\lambda)| + |\lambda| \left| \frac{\partial k^n}{\partial \lambda}(\lambda) \right| \right\} < \infty.
$$

By the mean value theorem, the similar argument in Lemma E.1 yields

$$
\int_{-\pi}^{-2\pi/|\tau|} \left| k^n(\lambda) - k^n\left(\lambda + \frac{\pi}{|\tau|}\right) \right| d\lambda \le c_2 \frac{\pi}{|\tau|} \int_{\pi}^{(|\tau|-1)\pi} \lambda^{-1} \, d\lambda
$$

$$
= c_2 \frac{\pi}{|\tau|} \{\log((|\tau|-1)\pi) - \log\pi\} = O\left(|\tau|^{-1}\log|\tau|\right)
$$

27

as $|\tau| \to \infty$. A similar argument shows that

$$\sup_{n \in \mathbb{N}} \int_{\pi/|\tau|}^{\pi} \left| k^n(\lambda) - k^n\left(\lambda + \frac{\pi}{|\tau|}\right) \right| \, d\lambda = O\left(|\tau|^{-1} \log |\tau|\right) \text{ as } |\tau| \to \infty.$$

Since $k^n(\lambda)$ is bounded a.e. from the assumption, the same argument in Lemma E.1 yields

$$\sup_{n \in \mathbb{N}} \int_{-2\pi/|\tau|}^{\pi/|\tau|} \left| k^n(\lambda) - k^n\left(\lambda + \frac{\pi}{|\tau|}\right) \right| \, d\lambda = O\left(|\tau|^{-1}\right) \text{ as } |\tau| \to \infty.$$

This completes the proof of Lemma E.2. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

E.2 Proof of Theorem E.3

E.2.1 Outline of Proof of Theorem E.3

Fix $p \in \mathbb{N}$ and note that

$$\text{Tr}\left[\left(\Sigma_n(k_1^n)\Sigma_n(k_2^n)\right)^p\right]$$

$$= \sum_{j_1=0}^{n-1} \cdots \sum_{j_{2p}=0}^{n-1} \widehat{k_1^n}(j_1 - j_2)\widehat{k_2^n}(j_2 - j_3) \cdots \widehat{k_1^n}(j_{2p-1} - j_{2p})\widehat{k_2^n}(j_{2p} - j_1)$$

$$= \sum_{j_1=0}^{n-1} \cdots \sum_{j_{2p}=0}^{n-1} \left(\int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} e^{\sqrt{-1}(j_1-j_2)y_1} e^{\sqrt{-1}(j_2-j_3)y_2} \cdots e^{\sqrt{-1}(j_{2p}-j_1)y_{2p}} k_1^n(y_1)k_2^n(y_2) \cdots k_1^n(y_{2p-1})k_2^n(y_{2p}) \, dy_1 \cdots dy_{2p}\right)$$

$$= \int_{U_\pi} P_n(\mathbf{y})Q_n(\mathbf{y}) \, d\mathbf{y},$$

where $U_t := [-t, t]^{2p}$ for $t \in (0, \pi]$ and

$$P_n(\mathbf{y}) := h_n^*(y_1 - y_{2p})h_n^*(y_2 - y_1) \cdots h_n^*(y_{2p} - y_{2p-1}), \quad h_n^*(y) := \sum_{j=0}^{n-1} e^{\sqrt{-1}jy},$$

$$Q_n(\mathbf{y}) := k_1^n(y_1)k_2^n(y_2) \cdots k_1^n(y_{2p-1})k_2^n(y_{2p}).$$

Following the arguments of Fox and Taqqu [13], we divide U_π into three disjoint sets E_t, F_t, G given by

$$E_t := U_\pi \backslash \{U_t \cup W\}, \quad F_t := U_t \backslash W, \quad G := U_\pi \cap W,$$

where $t \in (0, \pi]$ and

$$W_j := \left\{ \mathbf{y} = (y_1, \cdots, y_{2p}) \in \mathbb{R}^{2p} : |y_j| \leq \frac{|y_{j+1}|}{2} \right\}, \quad j = 1, \cdots, 2p,$$

$$W := W_1 \cup W_2 \cup \cdots W_{2p}.$$

Note that we use the notation $y_{2p+1} \equiv y_1$ for simplicity.

In order to prove the first result of Theorem E.3, it suffices to prove that $p(\alpha_1 + \alpha_2) < 1$ implies the

following three results:

$$\lim_{n\to\infty} \frac{1}{n} \int_{E_t} P_n(\mathbf{y})Q_n(\mathbf{y}) \, d\mathbf{y} = (2\pi)^{2p-1} \int_{t\le|z|\le\pi} [f(z)g(z)]^p \, dz, \quad \forall t \in (0,1], \tag{35}$$

$$\lim_{t\to 0} \limsup_{n\to\infty} \frac{1}{n} \int_{F_t} P_n(\mathbf{y})Q_n(\mathbf{y}) \, d\mathbf{y} = 0, \tag{36}$$

$$\lim_{n\to\infty} \frac{1}{n} \int_{G} P_n(\mathbf{y})Q_n(\mathbf{y}) \, d\mathbf{y} = 0. \tag{37}$$

Remark E.4. In order to prove (36), we will show that $p(\alpha_1 + \alpha_2) < 1$ implies

$$\lim_{t\to 0} \limsup_{n\to\infty} \frac{1}{n} \int_{U_t} P_n(\mathbf{y})Q_n(\mathbf{y}) \, d\mathbf{y} = 0. \tag{38}$$

Remark E.5. Since $G = \bigcup_{j=1}^{2p}[U_\pi \cap W_j]$, the relation (37) will hold if we prove that $p(\alpha_1 + \alpha_2) < 1$ implies

$$\lim_{n\to\infty} \frac{1}{n} \int_{U_\pi \cap W_j} |P_n(\mathbf{y})Q_n(\mathbf{y})| \, d\mathbf{y} = 0, \quad j = 1, \cdots, 2p. \tag{39}$$

From the definition of P_n and Q_n, it is clear that

$$\int_{U_\pi \cap W_1} |P_n(\mathbf{y})Q_n(\mathbf{y})| \, d\mathbf{y} = \int_{U_\pi \cap W_3} |P_n(\mathbf{y})Q_n(\mathbf{y})| \, d\mathbf{y} = \cdots = \int_{U_\pi \cap W_{2p-1}} |P_n(\mathbf{y})Q_n(\mathbf{y})| \, d\mathbf{y}$$

and

$$\int_{U_\pi \cap W_2} |P_n(\mathbf{y})Q_n(\mathbf{y})| \, d\mathbf{y} = \int_{U_\pi \cap W_4} |P_n(\mathbf{y})Q_n(\mathbf{y})| \, d\mathbf{y} = \cdots = \int_{U_\pi \cap W_{2p}} |P_n(\mathbf{y})Q_n(\mathbf{y})| \, d\mathbf{y}.$$

Because of the symmetry between α_1 and α_2 in the hypothesis of theorem, it is clear that we prove that $p(\alpha_1 + \alpha_2) < 1$ implies

$$\lim_{n\to\infty} \frac{1}{n} \int_{U_\pi \cap W_1} |P_n(\mathbf{y})Q_n(\mathbf{y})| \, d\mathbf{y} = 0, \tag{40}$$

then we will have also established

$$\lim_{n\to\infty} \frac{1}{n} \int_{U_\pi \cap W_2} |P_n(\mathbf{y})Q_n(\mathbf{y})| \, d\mathbf{y} = 0.$$

Thus (39) will follow from (40).

In conclusion, the first result of Theorem E.3 will be proven if we show that $p(\alpha_1 + \alpha_2) < 1$ implies (35), (38) and (40). Moreover, the second result of Theorem E.3 will be proven if we show that $p(\alpha_1 + \alpha_2) \ge 1$ implies

$$\forall \psi > 0, \quad \int_{U_\pi} |P_n(\mathbf{y})Q_n(\mathbf{y})| \, d\mathbf{y} = O(n^{p(\alpha_1+\alpha_2)+\psi}) \text{ as } n \to \infty. \tag{41}$$

These results will be proven in Section E.2.3. In the next subsection, we summarize several preliminaries used in the proof of Theorem E.3 following with Fox and Taqqu [13].

E.2.2 Preliminaries

To state the lemma, introduce the diagonal

$$D := \{\mathbf{y} = (y_1, \cdots, y_{2p}) \in U_\pi : y_1 = y_2 = \cdots = y_{2p}\}.$$

Let μ be the measure on U_π which is concentrated on D and satisfies $\mu(\{\mathbf{y} : a \le y_1 = y_2 = \cdots = y_{2p} \le b\}) = b - a$ for all $-\pi \le a \le b \le \pi$. Thus μ is Lebesgue measure on D, normalized so that $\mu(D) = 2\pi$.

Lemma E.6 (cf. Lemma 7.1. in [13]). *Define a measure μ_n on U_π by*

$$\mu_n(A) := \frac{1}{(2\pi)^{2p-1}n} \int_A P_n(\mathbf{y}) \, d\mathbf{y} \tag{42}$$

for each measurable set $A \subset U_\pi$. Then μ_n converges weakly to μ as $n \to \infty$.

For each $n \in \mathbb{N}$, define the function

$$h_n(z) := \begin{cases} \min\left(\frac{1}{|z+2\pi|}, n\right) & \text{if } -2\pi \le z < -\pi, \\ \min\left(\frac{1}{|z|}, n\right) & \text{if } -\pi \le z < \pi, \\ \min\left(\frac{1}{|z-2\pi|}, n\right) & \text{if } \pi \le z \le 2\pi. \end{cases}$$

and the function $f_n : \mathbb{R}^{2p} \to \mathbb{R}$ by

$$\begin{aligned} f_n(\mathbf{y}) :=& h_n\left(y_1 - y_{2p}\right) h_n\left(y_2 - y_1\right) h_n\left(y_3 - y_2\right) \cdots h_n\left(y_{2p} - y_{2p-1}\right) \\ &\times |y_1|^{-\alpha_1} |y_2|^{-\alpha_2} |y_3|^{-\alpha_1} \cdots |y_{2p}|^{-\alpha_2}, \end{aligned}$$

where $\alpha_1, \alpha_2 < 1$.

Lemma E.7. *There exists a constant $c > 0$ such that for each $n \in \mathbb{N}$ and $\mathbf{y} \in U_\pi$,*

$$\left|P_n(\mathbf{y})Q_n(\mathbf{y})\right| \le c f_n(\mathbf{y}).$$

Proof. As shown in [13], p.237, we have

$$|P_n(\mathbf{y})| \le 4^{2p} h_n\left(y_1 - y_{2p}\right) h_n\left(y_2 - y_1\right) h_n\left(y_3 - y_2\right) \cdots h_n\left(y_{2p} - y_{2p-1}\right)$$

for each $n \in \mathbb{N}$ and $\mathbf{y} \in U_\pi$. Therefore, the conclusion follows from the assumption. □

Proposition E.8 (cf. Proposition 6.1. in [13]). *Let $\alpha_1, \alpha_2 < 1$ and $W_1 = \{\mathbf{y} \in \mathbb{R}^{2p} : |y_1| \le \frac{|y_2|}{2}\}$.*

a) *If $\alpha_1 + \alpha_2 \le 0$, then for any $\psi > 0$,*

$$\int_{U_\pi \cap W_1} f_n(\mathbf{y}) \, d\mathbf{y} = O(n^\psi) \text{ as } n \to \infty.$$

b) *If $\alpha_1 + \alpha_2 > 0$, then for any $\psi > 0$,*

$$\int_{U_\pi \cap W_1} f_n(\mathbf{y}) \, d\mathbf{y} = O(n^{p(\alpha_1+\alpha_2)+\psi}) \text{ as } n \to \infty.$$

Proposition E.9 (cf. Proposition 6.2. in [13]). *Let $\alpha_1, \alpha_2 < 1$.*

a) If $p(\alpha_1 + \alpha_2) < 1$, then

$$\lim_{t \to 0} \limsup_{n \to \infty} \frac{1}{n} \int_{U_t} f_n(\mathbf{y}) \, d\mathbf{y} = 0.$$

b) If $p(\alpha_1 + \alpha_2) \geq 1$, then for any $\psi > 0$,

$$\int_{U_\pi} f_n(\mathbf{y}) \, d\mathbf{y} = O(n^{p(\alpha_1 + \alpha_2) + \psi}) \quad \text{as } n \to \infty.$$

E.2.3 Proof of Theorem E.3

As mentioned in Fox and Taqqu [13], p.237-238, the results (38), (40) and (41) immediately follow from Proposition 6.1., Proposition 6.2. in [13] in addition to Lemma E.7. In the rest of this section, we will prove (35). Note that

$$\frac{1}{n} \int_{E_t} P_n(\mathbf{y}) Q_n(\mathbf{y}) \, d\mathbf{y} = (2\pi)^{2p-1} \int_{E_t} Q_n(\mathbf{y}) \, \mu_n(d\mathbf{y}),$$

where μ_n is given in (42), and set

$$Q(\mathbf{y}) := k_1(y_1) k_2(y_2) \cdots k_1(y_{2p-1}) k_2(y_{2p}), \quad \mathbf{y} = (y_1, \cdots, y_{2p}) \in E_t.$$

Since the assumptions imply

$$\lim_{n \to \infty} \operatorname*{ess\,sup}_{\lambda \in [-\pi, \pi]} |Q^n(\lambda) - Q(\lambda)| = 0$$

and the limit function Q is continuous a.e. and bounded on E_t for each $t \in (0, \pi]$, see Fox and Taqqu [13], p.237, for more detail, Lemma 7.1. in Fox and Taqqu [13] yields

$$
\begin{aligned}
\frac{1}{n} \int_{E_t} P_n(\mathbf{y}) Q_n(\mathbf{y}) \, d\mathbf{y} &= (2\pi)^{2p-1} \int_{E_t} Q_n(\mathbf{y}) \, \mu_n(d\mathbf{y}) \\
&= (2\pi)^{2p-1} \int_{E_t} (Q_n(\mathbf{y}) - Q(\mathbf{y})) \, \mu_n(d\mathbf{y}) + (2\pi)^{2p-1} \int_{E_t} Q(\mathbf{y}) \, \mu_n(d\mathbf{y}) \\
&\stackrel{n \to \infty}{\to} (2\pi)^{2p-1} \int_{E_t} Q(\mathbf{y}) \, \mu(d\mathbf{y}) = (2\pi)^{2p-1} \int_{[-\pi,\pi] \setminus [-t,t]} [f(z)g(z)]^p \, dz.
\end{aligned}
$$

Therefore, the conclusion follows.

F Limit Theorems of Quadratic Forms

In this section, we derive several limit theorems of the quadratic form of random sequence which are used in the proof of Proposition G.1 and Proposition G.2 under the following assumptions.

Assumption F.1. Recall $\Theta := \Theta_H \times \Theta_\eta$ is a compact set of the form $\Theta_H := [H_-, H_+] \subset (0, 1]$ and $\Theta_\eta := [\eta_-, \eta_+] \subset (0, \infty)$. Let us consider a function $k : [-\pi, \pi] \times \Theta \times \mathbb{N} \to [-\infty, \infty]$, denoted by $k_\vartheta^n(\lambda) \equiv k(\lambda, \vartheta, n)$, be even and integrable on $[-\pi, \pi]$ for each $\vartheta \in \Theta$ and $n \in \mathbb{N}$ and assume there exist monotonically increasing continuous functions $\beta_0, \beta_1 : \Theta_H \to (-1, 1)$ such that the function k satisfies the conditions (C.1)-(C.3) below on a restricted parameter space $\Theta_0(\xi) := \Theta_{H,0}(\xi) \times \mathcal{K}$, where $\mathcal{K}$ be a compact interval of $(0, \infty)$ and

$$\Theta_{H,0}(\xi) := \{H \in \Theta_H : -\beta_0(H) - \alpha(H_0) \geq -1 + \xi, -\beta_1(H) - \alpha(H_0) \geq -1 + \xi\}, \quad \xi \in (0, 1).$$

Here H_0 denotes the true value of $H \in \Theta_H$, the function $\alpha : \Theta_H \to (-1, 1)$ is given in Lemma D.1 and we only

31

consider sufficiently small $\xi \in (0,1)$ such that $\mathring{\Theta}_{H,0}(\xi) \neq \emptyset$, where $\mathring{\Theta}_{H,0}(\xi)$ is the set of all interior points of $\Theta_{H,0}(\xi)$.

(C.1) For each $\vartheta \in \Theta_0(\xi)$, there exists a function k_ϑ such that

$$\lim_{n\to\infty} \operatorname*{ess\,sup}_{\lambda\in[-\pi,\pi]} |k_\vartheta^n(\lambda) - k_\vartheta(\lambda)| = 0,$$

and the discontinuities of k_ϑ has the Lebesgue measure 0 for each $\vartheta \in \Theta_0(\xi)$.

(C.2) For each $\vartheta \in \Theta_0(\xi)$, the following relations hold:

$$\sup_{n\in\mathbb{N},\lambda\in[-\pi,\pi]\setminus\{0\}} |\lambda|^{\beta_0(H)} \left|k_\vartheta^n(\lambda)\right| < \infty.$$

(C.3) For each $\lambda \in [-\pi,\pi]\setminus\{0\}$, $k_\vartheta^n(\lambda)$ is differentiable with respect to $\vartheta \in \Theta_0(\xi)$ and its partial derivatives satisfy

$$\sup_{\substack{n\in\mathbb{N},\lambda\in[-\pi,\pi]\setminus\{0\},\\ \vartheta=(\vartheta_1,\vartheta_2)\in\Theta_0(\xi)}} |\lambda|^{\beta_1(\vartheta_1)} \left|\frac{\partial k_\vartheta^n}{\partial \vartheta_j}(\lambda)\right| < \infty, \quad j = 1,2.$$

F.1 Basic Properties of Bilinear and Quadratic Forms

At first, we summarize several basic properties of the bilinear form B_n and the quadratic form Q_n as functionals on $L^1[-\pi,\pi]$ without proofs.

Lemma F.2. *Let $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$. The functionals $B_n(\mathbf{x}, \mathbf{y}, \cdot)$ and $Q_n(\mathbf{x}, \cdot)$ on $L^1[-\pi,\pi]$ satisfy the following properties.*

(1) *For each $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, the functional $B_n(\mathbf{x}, \mathbf{y}, \cdot)$ is linear on $L^1[-\pi,\pi]$.*

(2) *For each $\mathbf{x} \in \mathbb{C}^n$, the functional $Q_n(\mathbf{x}, \cdot)$ is non-decreasing on $L^1[-\pi,\pi]$, i.e. for each $k_1, k_2 \in L^1[-\pi,\pi]$,*

$$Q_n(\mathbf{x}, k_1) \leq Q_n(\mathbf{x}, k_2) \text{ if } k_1 \leq k_2,$$

where $k_1 \leq k_2$ means $k_1(\lambda) \leq k_2(\lambda)$ for a.e. $\lambda \in [-\pi,\pi]$.

(3) *For each $\mathbf{x} \in \mathbb{C}^n$, $Q_n(\mathbf{x}, k) \geq 0$ if $k \in L^1[-\pi,\pi]$ satisfies $k \geq 0$.*

(4) *For each $\mathbf{x} \in \mathbb{C}^n$ with $\mathbf{x} \neq 0$, $Q_n(\mathbf{x}, k) > 0$ if $k \in L^1[-\pi,\pi]$ satisfies $k \geq 0$ and the set $\{\lambda \in [-\pi,\pi] : k(\lambda) > 0\}$ has a positive Lebesgue measure.*

Next lemma is useful to evaluate asymptotic behaviors of bilinear forms.

Lemma F.3. *Suppose a sequence of functions k_ϑ^n, $n \in \mathbb{N}$, satisfies the condition (C.2) in Assumption F.1. Then there exists an even and 2π-periodic function $k_\vartheta^\dagger$, which is independent of the asymptotic parameter $n \in \mathbb{N}$, such that*

$$\sup_{n\in\mathbb{N}} \left|k_\vartheta^n(\lambda)\right| \leq \left|k_\vartheta^\dagger(\lambda)\right| \quad and \quad \sup_{\vartheta\in\Theta_0,\lambda\in[-\pi,\pi]\setminus\{0\}} \left\{|\lambda|^{\beta_0(H)} \left|k_\vartheta^\dagger(\lambda)\right|\right\} < \infty.$$

Moreover, the following two inequalities hold for each $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ and $\vartheta \in \Theta_0$:

$$\left|Q_n\left(\mathbf{x}, k_\vartheta^n\right)\right| \leq Q_n\left(\mathbf{x}, \left|k_\vartheta^n\right|\right) \leq Q_n(\mathbf{x}, k_\vartheta^\dagger), \tag{43}$$

$$\left|B_n\left(\mathbf{x}, \mathbf{y}, k_\vartheta^n\right)\right| \leq 2\sqrt{Q_n\left(\mathbf{x}, k_\vartheta^\dagger\right)} \sqrt{Q_n\left(\mathbf{y}, k_\vartheta^\dagger\right)}. \tag{44}$$

32

Proof. Define a function $k_\vartheta^\dagger$ by

$$k_\vartheta^\dagger(\lambda) := c\{2(1 - \cos \lambda)\} \sum_{j \in \mathbb{Z}} |\lambda + 2\pi j|^{-\beta_0(H)-2}$$

$$\text{with } c := \sup_{\vartheta \in \Theta_0, \lambda \in [-\pi,\pi] \setminus \{0\}, n \in \mathbb{N}} \left\{ \frac{|\lambda|^2}{2(1 - \cos \lambda)} \cdot |\lambda|^{\beta_0(H)} \left| k_\vartheta^n(\lambda) \right| \right\}.$$

Then it is obvious that the function $k_\vartheta^\dagger$ satisfies all conditions mentioned at the beginning. Moreover, the first inequality immediately follows from Lemma F.2 (2). In the rest of this proof, we will prove the second inequality. Decompose k_ϑ^n into the following two non-negative functions:

$$k_\vartheta^n(\lambda) = k_{\vartheta,+}^n(\lambda) - k_{\vartheta,-}^n(\lambda), \text{ where } k_{\vartheta,+}^n(\lambda) := \max(k_\vartheta^n(\lambda), 0), \ k_{\vartheta,-}^n(\lambda) := \max(-k_\vartheta^n(\lambda), 0).$$

Note that both of $k_{\vartheta,+}^n$ and $k_{\vartheta,-}^n$ are even functions and satisfy the condition (C.2) from the assumptions of k_ϑ^n. At first, consider the case where both of $k_{\vartheta,+}^n$ and $k_{\vartheta,-}^n$ are positive almost everywhere. Since Lemma F.2 (4) yields the matrix $\Sigma_n(k_\vartheta^n)$ is positive definite, Lemma F.2 (1), Schwartz's inequality of bilinear forms and (43) yield that for each $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$,

$$\left| B_n\left(\mathbf{x}, \mathbf{y}, k_\vartheta^n\right) \right| \leq \left| B_n\left(\mathbf{x}, \mathbf{y}, k_{\vartheta,+}^n\right) \right| + \left| B_n\left(\mathbf{x}, \mathbf{y}, k_{\vartheta,-}^n\right) \right|$$

$$\leq \sum_{i \in \{+,-\}} \sqrt{Q_n\left(\mathbf{x}, k_{\vartheta,i}^n\right)} \sqrt{Q_n\left(\mathbf{y}, k_{\vartheta,i}^n\right)} \leq 2 \sqrt{Q_n\left(\mathbf{x}, k_\vartheta^\dagger\right)} \sqrt{Q_n\left(\mathbf{y}, k_\vartheta^\dagger\right)}.$$

Note that the above inequalities also follows even if $k_{\vartheta,+}^n \equiv 0$ or $k_{\vartheta,-}^n \equiv 0$. Therefore, the conclusion follows. $\square$

The following result immediately follows from Lemma F.2 and Lemma F.3.

Corollary F.4. *Let $J \in \mathbb{N}$ and suppose a sequence of functions k_ϑ^n, $n \in \mathbb{N}$, satisfies the condition (C.2) in Assumption F.1. For any n-dimensional vector of the form $\mathbf{y} := \sum_{j=0}^J a_j \mathbf{w}_j$ with $\mathbf{w}_j \in \mathbb{C}^n$ and $a_j \in \mathbb{C}$ for $j \in \{0, 1, 2, \cdots, J\}$, the following inequality holds:*

$$\left| Q_n\left(\mathbf{y}, k_\vartheta^n\right) - Q_n\left(a_0 \mathbf{w}_0, k_\vartheta^n\right) \right| \leq \sum_{i=0}^J \sum_{j=1}^J |a_i| |a_j| \left| B_n\left(\mathbf{w}_i, \mathbf{w}_j, k_\vartheta^n\right) \right|$$

$$\leq 2 \sum_{i=0}^J \sum_{j=1}^J |a_i| |a_j| \sqrt{Q_n\left(\mathbf{w}_i, k_\vartheta^\dagger\right)} \sqrt{Q_n\left(\mathbf{w}_j, k_\vartheta^\dagger\right)},$$

where $k_\vartheta^\dagger$ is given in Lemma F.3.

F.2 Pointwise Convergence of Gaussian Quadratic Form

Denote by $\widetilde{\mathbf{G}}_n := \delta_n^{-H_0} \mathbf{G}_n$. In the next lemma, we show a pointwise convergence of the quadratic form of the stationary Gaussian sequence $\widetilde{\mathbf{G}}_n$, $n \in \mathbb{N}$.

Lemma F.5. *Suppose a sequence of functions k_ϑ^n, $n \in \mathbb{N}$, satisfies the conditions (C.1) and (C.2) in Assumption F.1. Under the conditions (H.1) and (H.3), the following convergence holds for each $\vartheta \in \Theta_0(\xi)$:*

$$\lim_{n \to \infty} \left\| Q_n\left(\widetilde{\mathbf{G}}_n, k_\vartheta^n\right) - Q_{\vartheta_0}\left(k_\vartheta\right) \right\|_2 = 0,$$

where k_ϑ is the limit function given in (C.1) and

$$Q_{\vartheta_0}(k_\vartheta) := \int_{-\pi}^{\pi} \eta_0^2 f_{H_0}(\lambda) k_\vartheta(\lambda) \, d\lambda. \tag{45}$$

Proof. At first, we obtain

$$\left\| Q_n\left(\widetilde{\mathbf{G}}_n, k_\vartheta^n\right) - Q_{\vartheta_0}(k_\vartheta) \right\|_2^2 = \mathrm{Var}\left[Q_n\left(\widetilde{\mathbf{G}}_n, k_\vartheta^n\right) \right] + \left\{ \mathrm{E}\left[Q_n\left(\widetilde{\mathbf{G}}_n, k_\vartheta^n\right) \right] - Q_{\vartheta_0}(k_\vartheta) \right\}^2$$

$$= \frac{2}{(2\pi n)^2} \mathrm{Tr}\left[\left(\Sigma_n\left(h_{\vartheta_0}^n\right) \Sigma_n\left(k_\vartheta^n\right) \right)^2 \right] + \left(\frac{1}{2\pi n} \mathrm{Tr}\left[\Sigma_n\left(h_{\vartheta_0}^n\right) \Sigma_n\left(k_\vartheta^n\right) \right] - Q_{\vartheta_0}(k_\vartheta) \right)^2,$$

where $h_{\widetilde{\vartheta}}^n \equiv h_{H,\widetilde{\nu}}^n$ is given in (14). Note that $\widetilde{\vartheta}_0 := (H_0, \delta_n^{-H_0} \nu_0) = \vartheta_0$. Since $\vartheta = (H, \nu) \in \Theta_0(\xi)$ implies $\beta_0(H) + \alpha(H_0) < 1$ and under the conditions (H.1) and (H.3), we have

$$\sup_{\lambda \in [-\pi,\pi]} |h_{\vartheta_0}^n(\lambda) - \eta_0^2 f_{H_0}(\lambda)| = \frac{1}{m_n \delta_n^{H_0}} \sup_{\lambda \in [-\pi,\pi]} |\ell(\lambda)| \overset{n \to \infty}{\to} 0,$$

the conclusion follows from the conditions (C.1), (C.2) and Theorem E.3. $\qquad\square$

The following result is easily proven in the similar way to the proof of Lemma F.5.

Corollary F.6. *Suppose a sequence of functions k_ϑ^n, $n \in \mathbb{N}$, satisfies the condition (C.2) in Assumption F.1. Under the conditions (H.1) and (H.3), the following convergence holds for each $\vartheta = (H, \nu) \in \Theta_0(\xi)$ satisfying $\alpha(H_0) + \beta_0(H) < 1/2$:*

$$Q_n\left(\widetilde{\mathbf{G}}_n, k_\vartheta^n\right) - E\left[Q_n\left(\widetilde{\mathbf{G}}_n, k_\vartheta^n\right) \right] = O_P\left(1/\sqrt{n}\right) \text{ as } n \to \infty.$$

F.3 Pointwise Convergence of Quadratic Form of Observation $\mathbf{Y}_n$

Denote by $\widetilde{\mathbf{Y}}_n := \delta_n^{-H_0} \mathbf{Y}_n$. Our goal in this subsection is to prove that the quadratic form of the rescaled observation $\widetilde{\mathbf{Y}}_n$ and that of the Gaussian vector $\widetilde{\mathbf{G}}_n$ are asymptotically equivalent as $\delta_n \to 0$. Namely, we show the following result.

Proposition F.7. *Suppose a sequence of functions k_ϑ^n, $n \in \mathbb{N}$, satisfies the condition (C.2) in Assumption F.1. Under the conditions (H.1)−(H.3), there exists a constant $\psi > 0$ such that the following convergence holds for each $\vartheta \in \Theta_0(\xi)$:*

$$Q_n\left(\mathbf{Y}_n, k_\vartheta^n\right) = Q_n\left(\mathbf{G}_n, k_\vartheta^n\right) + o_P\left(\delta_n^{2H_0+\psi}\right) \text{ as } n \to \infty.$$

Proof. From Proposition B.1, Corollary F.4 and Lemma F.5, it suffices to prove the following two results for the non-negative function $k_\vartheta^\dagger$ given in Lemma F.3 and each $\vartheta \in \Theta_0(\xi) \times (0, \infty)$:

(R.1) For any $K \in \mathbb{N}$ and $\mathbf{p} \equiv (p_1, \cdots, p_K) \in \mathbb{N}^K$, the following relation holds:

$$Q_n\left(\Delta \mathbf{W}_n^{\mathbf{p}}, k_\vartheta^\dagger\right) = O_P\left(\delta_n^{2|\mathbf{p}|H_0}\right) \text{ as } n \to \infty.$$

(R.2) Assume that there exists a positive random variable A, which is independent of the asymptotic parameter $n \in \mathbb{N}$, such that a random vector $\mathbf{R}_n := (R_1^n, R_2^n, \cdots, R_n^n)$ satisfies

$$\sup_{t \in \Lambda_n} |R_t^n| \leq A \cdot \delta_n. \tag{46}$$

Then there exists a constant $\psi > 0$ such that the following relation holds:

$$Q_n\left(\mathbf{R}_n, k_\vartheta^\dagger\right) = o\left(\delta_n^{2H_0+\psi}\right) \text{ as } n \to \infty.$$

In the rest of this proof section, we prove (R.1) and (R.2). $\qquad\square$

Proof of (R.1). Fix $\vartheta \in \Theta_0(\xi) \times (0, \infty)$. At first, Chebyshev's inequality and Lemma F.2 (3) yield that the following inequality holds for any $M > 0$:

$$\begin{aligned}
P\left[\left|Q_n\left(\Delta\mathbf{W}_n^\mathbf{p}, k_\vartheta^\dagger\right)\right| > M\right] &\leq \frac{1}{M}E\left[Q_n\left(\Delta\mathbf{W}_n^\mathbf{p}, k_\vartheta^\dagger\right)\right] \\
&= \frac{1}{2\pi M} \cdot \frac{1}{n}\sum_{s,t=1}^n \widehat{k_\vartheta^\dagger}(s-t)\mathrm{Cov}\left[\Delta W_s^{n,\mathbf{p}}, \Delta W_t^{n,\mathbf{p}}\right] \\
&= \frac{1}{2\pi M}\sum_{|\tau|<n}\left(1 - \frac{|\tau|}{n}\right)\widehat{k_\vartheta^\dagger}(\tau)\mathrm{Cov}\left[\Delta W_1^{n,\mathbf{p}}, \Delta W_{1+|\tau|}^{n,\mathbf{p}}\right],
\end{aligned} \tag{47}$$

where the stationarity property of $\{W_t^{n,\mathbf{p}}\}_{t\in\mathbb{Z}}$ is used in the last equality, see Proposition C.4. Since the function $k_\vartheta^\dagger$ satisfies the all assumptions in Lemma E.1 and Lemma E.2 with respect to $\beta \equiv \beta_0(H)$, we obtain

$$\widehat{k_\vartheta^\dagger}(\tau) = O\left(|\tau|^{\beta_0(H)-1}\right) \text{ as } |\tau| \to \infty. \tag{48}$$

As a result, (48) and Proposition C.4 yield that there exists a constant $c > 0$ such that the last quantity of (47) is dominated by

$$\begin{aligned}
\frac{1}{2\pi M}\sum_{|\tau|<n}\left|\widehat{k_\vartheta^\dagger}(\tau)\right|\left|\mathrm{Cov}\left[\Delta W_1^p, \Delta W_{1+|\tau|}^p\right]\right| &\leq \frac{c\delta_n^{2|\mathbf{p}|H_0}}{2\pi M}\sum_{|\tau|<n}|\tau|^{\beta_0(H)-1+(2H_0-4)} \\
&\leq \frac{c\delta_n^{2|\mathbf{p}|H_0}}{2\pi M}\sum_{\tau\in\mathbb{Z}}|\tau|^{\beta_0(H)+\alpha(H_0)-4}.
\end{aligned} \tag{49}$$

Note that the series in (49) converges because $H \in \Theta_{H,0}(\xi)$ implies $\beta_0(H) + \alpha(H_0) - 4 < -1$. Since the last quantity of (49) is independent of the asymptotic parameter $n \in \mathbb{N}$, the conclusion follows as $M \to \infty$. $\qquad\square$

Proof of (R.2). Fix $\xi \in (0, 1)$. At first, (46) and (48) yield that there exists a constant $c > 0$ such that

$$\begin{aligned}
Q_n(\mathbf{R}_n, k_\vartheta^\dagger) &= \frac{1}{2\pi n}\sum_{s,t=1}^n \widehat{k_\vartheta^\dagger}(s-t)R_s^n R_t^n \\
&\leq \frac{A^2}{2\pi} \cdot \frac{\delta_n^2}{n}\sum_{s,t=1}^n\left|\widehat{k_\vartheta^\dagger}(s-t)\right| \\
&= \frac{A^2}{2\pi}\cdot\delta_n^2\sum_{|\tau|<n}\left(1-\frac{|\tau|}{n}\right)\left|\widehat{k_\vartheta^\dagger}(\tau)\right| \leq \frac{cA^2}{2\pi}\cdot\delta_n^2\sum_{|\tau|<n}|\tau|^{\beta_0(H)-1}.
\end{aligned} \tag{50}$$

Moreover, for any $\psi \in (0, \xi)$, the last quantity of (50) is dominated by

$$\delta_n^2\sum_{|\tau|<n}|\tau|^{\beta_0(H)-1} \leq \delta_n^2\sum_{|\tau|<n}\frac{n^{1-\alpha(H_0)-\psi}}{|\tau|^{1-\alpha(H_0)-\psi}}|\tau|^{\beta_0(H)-1} \leq \delta_n^{2H_0+\psi}T_n^{2-2H_0-\psi}\sum_{\tau\in\mathbb{Z}}|\tau|^{\alpha(H_0)+\beta_0(H)-2+\psi}. \tag{51}$$

Note that the series in (51) converges because $\psi \in (0, \xi)$ and $H \in \Theta_{H,0}(\xi)$ imply $\alpha(H_0) + \beta_0(H) - 2 + \psi \leq -1 + \psi - \xi < -1$. Then the conclusion follows from (50), (51) and the assumptions (H.1) and (H.2). □

We can obtain the following result from Lemma F.5 and Proposition F.7.

Corollary F.8. *Suppose a sequence of functions k_ϑ^n, $n \in \mathbb{N}$, satisfies the conditions (C.1) and (C.2) in Assumption F.1. Under the conditions (H.1) − (H.3), the following convergence holds for each $\vartheta \in \Theta_0(\xi)$:*

$$Q_n\left(\widetilde{\mathbf{Y}}_n, k_\vartheta^n\right) = Q_{\vartheta_0}(k_\vartheta) + o_P(1) \quad as \quad n \to \infty.$$

F.4 Uniform Convergence of Quadratic Form of Observations $\mathbf{Y}_n$

In this subsection, we prove a uniform convergence of the quadratic form of $\widetilde{\mathbf{Y}}_n$ which is an extension of Corollary F.8 given in the previous subsection.

Proposition F.9. *Suppose a sequence of functions k_ϑ^n, $n \in \mathbb{N}$, satisfies the conditions (C.1)-(C.3) in Assumption F.1. Under the conditions (H.1) − (H.3), the following uniform convergence holds:*

$$\sup_{\vartheta \in \Theta_0(\xi)} \left| Q_n\left(\widetilde{\mathbf{Y}}_n, k_\vartheta^n\right) - Q_{\vartheta_0}(k_\vartheta) \right| = o_P(1) \quad as \quad n \to \infty.$$

Proof. Fix $\xi \in (0, 1)$. At first, the compactness of $\Theta_0(\xi)$ yields that for each $r > 0$, there exists $j(r) \in \mathbb{N}$ and a finite open covering $\{B_r(\vartheta_i)\}_{i \in \Lambda_{j(r)}}$ given by

$$B_r(\vartheta_i) := \{\vartheta \in \Theta_0(\xi) : \|\vartheta - \vartheta_i\|_{\mathbb{R}^2} < r\} \quad for \quad \vartheta_i = (H_i, \eta_i) \in \Theta_0(\xi), i \in \Lambda_{j(r)}.$$

Then we obtain the following inequality:

$$\begin{aligned}
\sup_{\vartheta \in \Theta_0(\xi)} \left| Q_n\left(\widetilde{\mathbf{Y}}_n, k_\vartheta^n\right) - Q_{\vartheta_0}(k_\vartheta) \right| &\leq \sup_{i \in \Lambda_{j(r)}, \vartheta \in B_r(\vartheta_i)} \left| Q_n\left(\widetilde{\mathbf{Y}}_n, k_\vartheta^n\right) - Q_{\vartheta_0}(k_\vartheta) \right| \\
&\leq \max_{i \in \Lambda_{j(r)}} \left| Q_n(\widetilde{\mathbf{Y}}_n, k_{\vartheta_i}^n) - Q_{\vartheta_0}(k_{\vartheta_i}) \right| \\
&\quad + \sup_{\substack{\|\vartheta_1 - \vartheta_2\|_{\mathbb{R}^2} < r \\ \vartheta_1, \vartheta_2 \in \Theta_0(\xi) \times \mathcal{K}}} \left| Q_{\vartheta_0}(k_{\vartheta_1}) - Q_{\vartheta_0}(k_{\vartheta_2}) \right| \\
&\quad + \sup_{i \in \Lambda_{j(r)}, \vartheta \in B_r(\vartheta_i)} \left| Q_n\left(\widetilde{\mathbf{Y}}_n, k_\vartheta^n\right) - Q_n\left(\widetilde{\mathbf{Y}}_n, k_{\vartheta_i}^n\right) \right|.
\end{aligned} \tag{52}$$

Here Corollary F.8 yields that for each $r > 0$, the first term of the last quantity of (52) converges to 0 as $n \to \infty$. Moreover, the second term of it also converges to 0 as $r \downarrow 0$ because $\vartheta \mapsto Q_{\vartheta_0}(k_\vartheta)$ is uniformly continuous on $\Theta_0(\xi)$ under the condition (C.3). As a result, it suffices to show that the third term of it is negligible for sufficiently small $r > 0$ and large $n \in \mathbb{N}$.

Without loss of generality, we assume $r \in (0, \xi/2)$ and

$$\sup_{H_1^\dagger, H_2^\dagger \in \Theta_{H,0}(\xi), |H_1^\dagger - H_2^\dagger| < r} \|\beta_1(H_1^\dagger) - \beta_1(H_2^\dagger)\| < \xi/2$$

since β_1 is uniformly continuous on $\Theta_{H,0}(\xi)$. Here the condition (C.3) implies

$$c_1 := \sup_{\substack{n \in \mathbb{N}, \lambda \in [-\pi, \pi] \setminus \{0\}, \\ \vartheta = (H, \eta) \in \Theta_0(\xi)}} |\lambda|^{\beta_1(H)} \left\| \nabla k_\vartheta^n(\lambda) \right\|_{\mathbb{R}^2} < \infty.$$

Then the mean value theorem and Schwartz's inequality yield that for any $\vartheta_i^{\dagger,1}, \vartheta_i^{\dagger,2} \in B_r(\vartheta_i)$, $i \in \Lambda_{j(r)}$ and $\lambda \in [-\pi, \pi] \setminus \{0\}$,

$$\left| k_{\vartheta_i^{\dagger,1}}^n(\lambda) - k_{\vartheta_i^{\dagger,2}}^n(\lambda) \right| \leq \left\| \nabla k_{\vartheta_i^{\dagger}}^n(\lambda) \right\|_{\mathbb{R}^2} \left\| \vartheta_i^{\dagger,1} - \vartheta_i^{\dagger,2} \right\|_{\mathbb{R}^2} \leq rc_1 |\lambda|^{-\beta_1(H_i^{\dagger})} \leq rc_2 |\lambda|^{-\beta_1(H_i)-\xi/2}, \tag{53}$$

where $c_2 := c_1 \pi^{\xi}$ and $\vartheta_i^{\dagger} \equiv (H_i^{\dagger}, \eta_i^{\dagger}) \in B_r(\vartheta_i)$ is determined by the relation $\vartheta_i^{\dagger} = \vartheta_i^{\dagger,1} + t(\vartheta_i^{\dagger,1} - \vartheta_i^{\dagger,2})$ with $t \equiv t(\vartheta_i^{\dagger,1}, \vartheta_i^{\dagger,2}) \in (0,1)$. Since $\vartheta_i^{\dagger} \equiv (H_i^{\dagger}, \eta_i^{\dagger}) \in \Theta_0(\xi)$ implies $-\beta_1(H_i^{\dagger}) - \alpha(H_0) - \xi/2 > -1$, Lemma F.2 and (53) yield that the third term of the last quantity of (52) is dominated by

$$\max_{i \in \Lambda_{j(r)}} Q_n \left(\widetilde{\mathbf{Y}}_n, \sup_{\vartheta \in B_r(\vartheta_i)} |k_\vartheta^n - k_{\vartheta_i}^n| \right) \leq r \frac{c_2}{2\pi} \max_{i \in \Lambda_{j(r)}} \int_{-\pi}^{\pi} I_n \left(\lambda, \widetilde{\mathbf{Y}}_n \right) |\lambda|^{-\beta_1(H_i)-\xi/2} \, d\lambda$$

$$\leq r \frac{c_2}{2\pi} \left(\max_{i \in \Lambda_{j(r)}} Q_{H_0,\xi}(H_i) + \max_{i \in \Lambda_{j(r)}} R_{n,\xi}(H_i) \right), \tag{54}$$

where

$$Q_{H_0,\xi}(H_i) := \int_{-\pi}^{\pi} \eta_0^2 f_{H_0}(\lambda) |\lambda|^{-\beta_1(H_i)-\xi/2} \, d\lambda,$$

$$R_{n,\xi}(H_i) := \left| \int_{-\pi}^{\pi} I_n \left(\lambda, \widetilde{\mathbf{Y}}_n \right) |\lambda|^{-\beta_1(H_i)-\xi/2} \, d\lambda - Q_{H_0,\xi}(H_i) \right|.$$

Moreover, Lemma D.1 and $\vartheta_i^{\dagger} \equiv (H_i^{\dagger}, \eta_i^{\dagger}) \in \Theta_0(\xi)$, $i \in \Lambda_{j(r)}$, yield that there exists a constant $c_3 \equiv c_3(\xi) > 0$, which is independent of $r \in \mathbb{N}$, such that the first term of the last quantity of (54) is dominated by

$$r \frac{c_2}{2\pi} \cdot \max_{i \in \Lambda_{j(r)}} Q_{H_0,\xi}(H_i) \leq r \frac{c_3}{2\pi} \int_{-\pi}^{\pi} |\lambda|^{-1+\xi/2} \, d\lambda = rc_3 \pi^{\xi/2-1}.$$

As a result, the first term of the last quantity of (54) converges to 0 as $r \downarrow 0$ irrespectively to the asymptotic parameter $n \in \mathbb{N}$. Moreover, Corollary F.8 yields that for each $r \in (0, \xi/2)$, the second term of the last quantity of (54) converges to 0 as $n \to \infty$. Therefore, the conclusion follows. □

G Proof of Theorem 2.8

Main purpose in this section is to give a proof of Theorem 2.8. We prepare notation used in its proof in Section G.1 and several limit theorems of estimation and its score functions are summarized in Section G.2. A key proposition and Theorem 2.8 are proven in Section G.3.

G.1 Notation of Parameter Space and Estimation Function

Recall $\Theta := \Theta_H \times \Theta_\eta$ is a compact set of the form $\Theta_H := [H_-, H_+] \subset (0,1]$, $\Theta_\eta := [\eta_-, \eta_+] \subset (0,\infty)$ and $\Theta_{\underline{\vartheta}}^n := \Theta_H \times \Theta_{\underline{\nu}}^n$ with $\Theta_{\underline{\nu}}^n := [\eta_- \delta_n^{H_+ - H_0}, \eta_+ \delta_n^{H_- - H_0}]$. Following the argument in Velasco and Robinson [32], we divide the parameter space Θ_H into the following two subsets:

$$\Theta_{H,1}(\xi) := \{ H \in \Theta_H : \alpha(H) - \alpha(H_0) \geq -1 + \xi \},$$
$$\Theta_{H,2}(\xi) := \{ H \in \Theta_H : \alpha(H) - \alpha(H_0) < -1 + \xi \},$$

37

where $\alpha(H) := 2H - 1$ and $\xi \in (0,1)$. Moreover, we also divide the rescaled parameter space $\Theta_{\widetilde{\vartheta}}^n$ into the following four subsets:

$$\Theta_j^n(\xi, L) := \Theta_{H,j}(\xi) \times \Theta_{\widetilde{\nu},0}^n(L), \quad \Theta_k^n(L) := \Theta_H \times \Theta_{\widetilde{\nu},k}^n(L)$$

for $L \in \mathbb{R}_+$ and $j, k \in \{1, 2\}$, where

$$\Theta_{\widetilde{\nu},0}^n(L) := \Theta_{\widetilde{\nu}}^n \cap [1/L, L], \quad \Theta_{\widetilde{\nu},1}^n(L) := \Theta_{\widetilde{\nu}}^n \cap (0, 1/L), \quad \Theta_{\widetilde{\nu},2}^n(L) := \Theta_{\widetilde{\nu}}^n \cap (L, \infty).$$

Denote by $U_{n,0}(\widetilde{\vartheta}) \equiv U_{n,0}(H, \widetilde{\nu})$, defined in Section 2.2, and

$$U_{\vartheta_0}(\widetilde{\vartheta}) := \frac{1}{4\pi} \int_{-\pi}^{\pi} \left\{ \log\left(\widetilde{\nu}^2 f_H(\lambda)\right) + \frac{\eta_0^2 f_{H_0}(\lambda)}{\widetilde{\nu}^2 f_H(\lambda)} \right\} d\lambda.$$

Note that $\vartheta_0 = (H_0, \eta_0) = (H_0, \widetilde{\nu}_0)$ and for any $\xi \in (0,1)$ and $L \in \mathbb{R}_+$, we can show that $U_{\vartheta_0}(\widetilde{\vartheta})$ satisfies the identifiability condition with respect to the parameter $\widetilde{\vartheta}$ on $\Theta_1^n(\xi, L)$, i.e. for any $\iota > 0$,

$$\inf_{\widetilde{\vartheta} \in \Theta_1^n(\xi, L), \|\widetilde{\vartheta} - \vartheta_0\| \geq \iota} U_{\vartheta_0}(\widetilde{\vartheta}) > U_{\vartheta_0}(\vartheta_0), \tag{55}$$

by using Lemma D.1 (2) and the elementary inequality $\log x \leq x - 1$ for any $x > 0$ that is actually an equality only when $x = 1$.

G.2 Convergence of Estimation and Its Score Functions

Proposition G.1. *Let $\mathcal{K}$ be a compact interval of $(0, \infty)$ and ϑ_0 be an interior point of Θ. Under the conditions $(H.1) - (H.3)$, the following uniform convergences on $\Theta_{H,1}(\xi) \times \mathcal{K}$ hold for any $\xi \in (0,1)$:*

$$\sup_{\widetilde{\vartheta} \in \Theta_{H,1}(\xi) \times \mathcal{K}} \left| U_{n,0}(\widetilde{\vartheta}) - U_{\vartheta_0}(\widetilde{\vartheta}) \right| = o_P(1) \text{ as } n \to \infty,$$

$$\sup_{\widetilde{\vartheta} \in \Theta_{H,1}(\xi) \times \mathcal{K}} \left| \nabla^2 U_{n,0}(\widetilde{\vartheta}) - \nabla^2 U_{\vartheta_0}(\widetilde{\vartheta}) \right| = o_P(1) \text{ as } n \to \infty.$$

Proof. Let us consider only the first claim because the second one is proven in the similar way. Here we have

$$U_{n,0}(\widetilde{\vartheta}) - U_{\vartheta_0}(\widetilde{\vartheta}) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left\{ \log h_{\widetilde{\vartheta}}^n(\lambda) - \log\left(\widetilde{\nu}^2 f_H(\lambda)\right) \right\} d\lambda \tag{56}$$
$$+ \frac{1}{4\pi} \left\{ Q_n\left(\widetilde{\mathbf{Y}}_n, 1/h_{\widetilde{\vartheta}}^n\right) - Q_{\vartheta_0}\left(1/(\widetilde{\nu}^2 f_H)\right) \right\},$$

where Q_{ϑ_0} is defined in (45). Under the assumption $(H.1)$ and $(H.3)$, we obtain

$$\lim_{n \to \infty} \sup_{\lambda \in [-\pi, \pi] \setminus \{0\}, \widetilde{\vartheta} \in \Theta_{H,1}(\xi) \times \mathcal{K}} \left| \log h_{\widetilde{\vartheta}}^n(\lambda) - \log\left(\widetilde{\nu}^2 f_H(\lambda)\right) \right| = 0$$

so that the first term of rhs of (56) is negligible as $n \to \infty$. Note that we have

$$\sup_{n \in \mathbb{N}, \lambda \in [-\pi, \pi] \setminus \{0\}} \left\{ |\lambda|^{1-2H} \left| 1/h_{\widetilde{\vartheta}}^n(\lambda) \right| + |\lambda|^{1-2H+\iota} \left\| \nabla\left(1/h_{\widetilde{\vartheta}}^n\right)(\lambda) \right\|_{\mathbb{R}^2} \right\} < \infty$$

for any $\iota > 0$. Let us fix sufficiently small $\iota > 0$. Then we can show that the second term of rhs of (56)

38

is also negligible as $n \to \infty$ by using Proposition F.9 in the case of the function $1/h^n_{\widetilde{\vartheta}}$, $\beta_0(H) := 1 - 2H$ and $\beta_1(H) := 1 - 2H + \iota$. Therefore, the conclusion follows. $\qquad\square$

Proposition G.2. *Let ϑ_0 be an interior point of Θ. Under the conditions (H.1) – (H.3), the score function $\nabla U_{n,0}(\widetilde{\vartheta})$ has the following asymptotic behavior at the point ϑ_0: there exists a constant $\psi > 0$ such that*

$$\nabla U_{n,0}(\vartheta_0) = o_P(\delta_n^{\psi}) \ \text{ as } \ n \to \infty.$$

Proof. At first, we decompose $u^{\mathsf{T}} \nabla U_{n,0}(\vartheta_0)$ with $u \in \mathbb{R}^2$ into the following three parts:

$$u^{\mathsf{T}} \nabla U_{n,0}(\vartheta_0) = \frac{1}{4\pi} \left(A_1^n + A_2^n + A_3^n \right),$$

where

$$A_1^n := \int_{-\pi}^{\pi} u^{\mathsf{T}} \nabla \log h^n_{\vartheta_0}(\lambda) \, \mathrm{d}\lambda + E\left[Q_n\left(\widetilde{\mathbf{G}}_n, k^n_{\vartheta_0} \right) \right],$$

$$A_2^n := Q_n\left(\widetilde{\mathbf{Y}}_n, k^n_{\vartheta_0} \right) - Q_n\left(\widetilde{\mathbf{G}}_n, k^n_{\vartheta_0} \right), \quad A_3^n := Q_n\left(\widetilde{\mathbf{G}}_n, k^n_{\vartheta_0} \right) - E\left[Q_n\left(\widetilde{\mathbf{G}}_n, k^n_{\vartheta_0} \right) \right]$$

and $k^n_{\vartheta_0} := u^{\mathsf{T}} \nabla \left(1/h^n_{\vartheta_0} \right)$. Note that Lemma D.1 (3) yields that for any $\iota > 0$,

$$\sup_{n \in \mathbb{N}, \lambda \in [-\pi, \pi] \setminus \{0\}} |\lambda|^{1 - 2H_0 + \iota} \left| k^n_{\vartheta_0}(\lambda) \right| < \infty.$$

Let us fix sufficiently small $\iota > 0$. Then we can show that $A_3^n = O_P(1/\sqrt{n})$ as $n \to \infty$ and $A_2^n = o_P(\delta_n^{\psi})$ as $n \to \infty$ for a certain constant $\psi > 0$ by using Corollary F.6 and Proposition F.7 in the case of the function $k^n_{\vartheta_0}$ and $\beta_0(H) := 1 - 2H + \iota$ respectively. As a result, it suffices to prove that $A_1^n = O(1/\sqrt{n})$ as $n \to \infty$. Our proof is similar to that of Theorem 2 in Fox and Taqqu [12]. At first, we obtain

$$E\left[Q_n\left(\widetilde{\mathbf{G}}_n, k^n_{\vartheta_0} \right) \right] = \frac{1}{2\pi n} \sum_{i,j=1}^{n} \widehat{k^n_{\vartheta_0}}(j - i) \widehat{h^n_{\vartheta_0}}(j - i) = \frac{1}{2\pi} \sum_{|\tau| < n} \left(1 - \frac{|\tau|}{n} \right) \widehat{k^n_{\vartheta_0}}(\tau) \widehat{h^n_{\vartheta_0}}(\tau)$$

$$= \frac{1}{2\pi} \sum_{|\tau| < n} \widehat{k^n_{\vartheta_0}}(\tau) \widehat{h^n_{\vartheta_0}}(\tau) - \frac{1}{2\pi n} \sum_{|\tau| < n} |\tau| \widehat{k^n_{\vartheta_0}}(\tau) \widehat{h^n_{\vartheta_0}}(\tau). \tag{57}$$

Since the functions $k^n_{\vartheta_0}, h^n_{\vartheta_0} \in L^1[-\pi, \pi]$ are 2π-periodic, it is well-known that $\widehat{k^n_{\vartheta_0}}(\tau) \widehat{h^n_{\vartheta_0}}(\tau)$ is the τth Fourier coefficient of the convolution $k^n_{\vartheta_0} * h^n_{\vartheta_0}$ defined by

$$\left(k^n_{\vartheta_0} * h^n_{\vartheta_0} \right)(\lambda) := \int_{-\pi}^{\pi} k^n_{\vartheta_0}(x) h^n_{\vartheta_0}(\lambda - x) \, \mathrm{d}x = \int_{-\pi}^{\pi} u^{\mathsf{T}} \nabla \left(1/h^n_{\vartheta_0}(x) \right) h^n_{\vartheta_0}(x - \lambda) \, \mathrm{d}x,$$

where we use the property that $h^n_{\vartheta_0}$ is an even function in the above equality. Moreover, we have

$$\sup_{n \in \mathbb{N}} \left| \widehat{k^n_{\vartheta_0}}(\tau) \widehat{h^n_{\vartheta_0}}(\tau) \right| = O\left(|\tau|^{-2 + \iota} \right) \ \text{ as } |\tau| \to \infty \tag{58}$$

for any $\iota > 0$ from Lemma E.1 and Lemma E.2. As a result, for each $n \in \mathbb{N}$, $(k^n_{\vartheta_0} * h^n_{\vartheta_0})(\lambda)$ is expanded as the following Fourier series for a.e. $\lambda \in [-\pi, \pi]$:

$$\left(k^n_{\vartheta_0} * h^n_{\vartheta_0} \right)(\lambda) = \frac{1}{2\pi} \sum_{\tau \in \mathbb{Z}} \widehat{k^n_{\vartheta_0}}(\tau) \widehat{h^n_{\vartheta_0}}(\tau) e^{\sqrt{-1}\tau\lambda}. \tag{59}$$

Note that the continuity of the function $k^n_{\vartheta_0} * h^n_{\vartheta_0}$ on $[-\pi, \pi]$ implies the Fourier series expansion (59) also holds for all $\lambda \in [-\pi, \pi]$. In particular, we obtain

$$\frac{1}{2\pi} \sum_{\tau \in \mathbb{Z}} \widehat{k^n_{\vartheta_0}}(\tau) \widehat{h^n_{\vartheta_0}}(\tau) = \left(k^n_{\vartheta_0} * h^n_{\vartheta_0} \right)(0) \tag{60}$$

$$= \int_{-\pi}^{\pi} u^{\mathrm{T}} \nabla \left(1/h^n_{\vartheta_0}(\lambda) \right) h^n_{\vartheta_0}(\lambda) \, d\lambda = - \int_{-\pi}^{\pi} u^{\mathrm{T}} \nabla \log h^n_{\vartheta_0}(\lambda) \, d\lambda.$$

From the equalities (57) and (60), we obtain

$$\sqrt{n} A^n_3 = \sqrt{n} \left(\int_{-\pi}^{\pi} u^{\mathrm{T}} \nabla \log h^n_{\vartheta_0}(\lambda) \, d\lambda + \frac{1}{2\pi} \sum_{|\tau| < n} \widehat{k^n_{\vartheta_0}}(\tau) \widehat{h^n_{\vartheta_0}}(\tau) \right) - \frac{1}{2\pi \sqrt{n}} \sum_{|\tau| < n} |\tau| \widehat{k^n_{\vartheta_0}}(\tau) \widehat{h^n_{\vartheta_0}}(\tau)$$

$$= - \frac{\sqrt{n}}{2\pi} \sum_{|\tau| \geq n} \widehat{k^n_{\vartheta_0}}(\tau) \widehat{h^n_{\vartheta_0}}(\tau) - \frac{1}{2\pi \sqrt{n}} \sum_{|\tau| < n} |\tau| \widehat{k^n_{\vartheta_0}}(\tau) \widehat{h^n_{\vartheta_0}}(\tau). \tag{61}$$

Then we can show that both terms of (61) are negligible as $n \to \infty$ in the similar way to the proof of Theorem 2 in Fox and Taqqu [12] by using the relation (58). Therefore, the conclusion follows. $\square$

G.3 Proof of Theorem 2.8

Before proving Theorem 2.8, we show the following result.

Proposition G.3. *Let ϑ_0 be an interior point of Θ. Under the conditions (H.1) − (H.3),*

$$\widetilde{\vartheta}_n := (\widehat{H}_n, \delta_n^{-H_0} \widehat{v}_n) \to (H_0, \eta_0) \ \text{as} \ n \to \infty \ \text{in probability}.$$

Following the argument of Velasco and Robinson [32], we divide the proof of Proposition G.3 into the following two steps.

Step 1. *Let $\xi \in (0, 1)$ and $L > 0$. Define a random variable $\widetilde{\vartheta}_n^{(1)}(\xi, L)$ by*

$$\widetilde{\vartheta}_n^{(1)}(\xi, L) := \arg\min_{\widetilde{\vartheta} \in \Theta_1^n(\xi, L)} U_{n,0}(\widetilde{\vartheta}).$$

Then, for each $\xi \in (0, 1)$ and $L > 0$, $\widetilde{\vartheta}_n^{(1)}(\xi, L) \to \vartheta_0$ in probability as $n \to \infty$.

Proof of Step 1. Note that for any $\xi \in (0, 1)$ and $L > 0$, the parameter space $\Theta_1^n(\xi, L)$ no longer depends on the asymptotic parameter $n \in \mathbb{N}$ if n is sufficiently large. As a result, the conclusion immediately follows from Proposition G.1 and the identifiability condition of the limit function $U_{\vartheta_0}(\widetilde{\vartheta})$ on $\Theta_1^n(\xi, L)$, see (55). $\square$

Step 2. *There exist constants $\xi_0 \in (0, 1)$ and $L_0 > 0$ such that for any $\xi \in (0, \xi_0)$ and $L \geq L_0$, $\widetilde{\vartheta}_n - \widetilde{\vartheta}_n^{(1)}(\xi, L) \to 0$ in probability as $n \to \infty$.*

Proof of Step 2. Without loss of generality, we can assume $\Theta_2^n(\xi, L)$ and $\Theta_k^n(L)$, $k = 1, 2$, are non-empty sets for each $\xi \in (0, 1)$ and $L \in (0, \infty)$. Note that for any $\iota > 0$, we have

$$P \left[\|\widetilde{\vartheta}_n - \widetilde{\vartheta}_n^{(1)}(\xi, L)\|_{\mathbb{R}^2} > \iota \right] \leq P \left[\inf_{\widetilde{\vartheta} \in \Theta_{\widetilde{\vartheta}}^n \setminus \Theta_1^n(\xi, L)} U_{n,0}(\widetilde{\vartheta}) \leq \inf_{\widetilde{\vartheta} \in \Theta_1^n(\xi, L)} U_{n,0}(\widetilde{\vartheta}) \right]. \tag{62}$$

Since we can show that

$$\left\{ \inf_{\widetilde{\vartheta} \in \Theta_{\widetilde{\vartheta}}^n \setminus \Theta_1^n(\xi,L)} U_{n,0}(\widetilde{\vartheta}) \le U_{n,0}(\widetilde{\vartheta}_n^{(1)}) \right\} \subset \left\{ \inf_{\widetilde{\vartheta} \in \Theta_2^n(\xi,L)} U_{n,0}(\widetilde{\vartheta}) \le U_{n,0}(\widetilde{\vartheta}_n^{(1)}) \right\} \cup \bigcup_{k=1}^2 \left\{ \inf_{\widetilde{\vartheta} \in \Theta_k^n(L)} U_{n,0}(\widetilde{\vartheta}) \le U_{n,0}(\widetilde{\vartheta}_n^{(1)}) \right\},$$

the rhs of (62) is dominated by

$$P \left[\inf_{\widetilde{\vartheta} \in \Theta_2^n(\xi,L)} U_{n,0}(\widetilde{\vartheta}) \le U_{n,0}(\widetilde{\vartheta}_n^{(1)}) \right] + \sum_{k=1}^2 P \left[\inf_{\widetilde{\vartheta} \in \Theta_k^n(L)} U_{n,0}(\widetilde{\vartheta}) \le U_{n,0}(\widetilde{\vartheta}_n^{(1)}) \right]$$

$$\le P \left[\inf_{\widetilde{\vartheta} \in \Theta_2^n(\xi,L)} U_{n,0}(\widetilde{\vartheta}) \le U_{\vartheta_0}(\vartheta_0) + \iota \right] + \sum_{k=1}^2 P \left[\inf_{\widetilde{\vartheta} \in \Theta_k^n(L)} U_{n,0}(\widetilde{\vartheta}) \le U_{\vartheta_0}(\vartheta_0) + \iota \right] + 3P \left[\left| U_{n,0}(\widetilde{\vartheta}_n^{(1)}) - U_{\vartheta_0}(\vartheta_0) \right| \ge \iota \right] \quad (63)$$

for any $\iota > 0$. Then the first term of (63) is negligible as $n \to \infty$ from Proposition G.1 and Step 1. Therefore, it suffices to prove that the other terms are negligible as $n \to \infty$ if we take sufficiently large $L > 0$ and small $\xi \in (0,1)$ respectively. We divide the proof into the following three lemmas.

Lemma G.4. *For any $\iota, L > 0$, there exists a constant $\xi_0 \equiv \xi_0(L, \iota, \vartheta_0) \in (0,1)$ such that for any $\xi \in (0, \xi_0)$,*

$$\lim_{n \to \infty} P \left[\inf_{\widetilde{\vartheta} \in \Theta_2^n(\xi,L)} U_{n,0}(\widetilde{\vartheta}) \le U_{\vartheta_0}(\vartheta_0) + \iota \right] = 0.$$

Proof. Fix $\iota, L > 0$. Note that Lemma D.1 and the assumption (H.3) yield that there exist constants $c_1, c_2 > 0$, which are independent of $(H, \widetilde{\nu}, \lambda) \in \Theta_H \times [1/L, L] \times [-\pi, \pi] \setminus \{0\}$ and $n \in \mathbb{N}$, such that

$$c_1 |\lambda|^{-\alpha(H)} < h_{H,\widetilde{\nu}}^n(\lambda) < c_2 |\lambda|^{-\alpha(H)}$$

holds for any $(H, \widetilde{\nu}, \lambda) \in \Theta_H \times [1/L, L] \times [-\pi, \pi] \setminus \{0\}$ and $n \in \mathbb{N}$, where $h_{H,\widetilde{\nu}}^n$ is given in (14). Then for any $\widetilde{\vartheta}_1 = (H_1, \widetilde{\nu}_1)$ and $\widetilde{\vartheta}_2 = (H_2, \widetilde{\nu}_2)$ satisfying $\alpha(H_1) \ge \alpha(H_2)$, we obtain

$$U_{n,0}(\widetilde{\vartheta}_2) \ge \frac{1}{4\pi} \int_{-\pi}^{\pi} \log\left(c_1 |\lambda|^{-\alpha(H_2)} \right) d\lambda + \frac{1}{4\pi c_2} \int_{-\pi}^{\pi} I_n\left(\lambda, \widetilde{\mathbf{Y}}_n \right) |\lambda|^{\alpha(H_2)} d\lambda$$

$$\ge \frac{\log c_1}{2} - \frac{\alpha(H_1)}{2} (\log \pi - 1) + \frac{1}{4\pi^{1+2H_+} c_2} \int_{-\pi}^{\pi} I_n\left(\lambda, \widetilde{\mathbf{Y}}_n \right) |\lambda|^{\alpha(H_1)} d\lambda, \quad (64)$$

where we use the elementary inequality $|\lambda|^{\alpha(H_1)-\alpha(H_2)} \le \pi^{2H_+}$ for any $\lambda \in [-\pi, \pi]$ in the second inequality. Let $H_1^\dagger \equiv H_1^\dagger(\xi) := H_0 + (\xi - 1)/2 \in \partial \Theta_1^n(\xi, L)$, where $\partial \Theta_1^n(\xi, L)$ denotes the boundary of the set $\Theta_1^n(\xi, L)$. Since the relation $\alpha(H_1^\dagger) \ge \alpha(H_2)$ holds for any $\widetilde{\vartheta}_2 = (H_2, \widetilde{\nu}_2) \in \Theta_2^n(\xi, L)$, we can obtain the following inequality from (64):

$$\inf_{\widetilde{\vartheta}_2 \in \Theta_2^n(\xi,L)} U_{n,0}(\widetilde{\vartheta}_2) \ge \frac{\log c_1}{2} - \frac{\alpha(H_0) - 1 + \xi}{2} (\log \pi - 1) + \frac{1}{4\pi^{1+2H_+} c_2} \int_{-\pi}^{\pi} I_n\left(\lambda, \widetilde{\mathbf{Y}}_n \right) |\lambda|^{\alpha(H_0)-1+\xi} d\lambda. \quad (65)$$

Moreover, Corollary F.8 in the case of $k_\vartheta(\lambda) := |\lambda|^{\alpha(H)-1+\xi}$ and $\beta_0(H) := -\alpha(H) + 1 - \xi$ yields that the third term of the rhs of (65) converges to

$$\frac{1}{4\pi^{1+2H_+} c_2} \int_{-\pi}^{\pi} \eta_0^2 f_{H_0}(\lambda) |\lambda|^{\alpha(H_0)-1+\xi} d\lambda$$

41

in probability as $n \to \infty$ and we obtain the following inequality:

$$\frac{1}{4\pi^{1+2H_+}c_2} \int_{-\pi}^{\pi} \eta_0^2 f_{H_0}(\lambda)|\lambda|^{\alpha(H_0)-1+\xi} \, d\lambda \geq \frac{c_1}{4\pi^{1+2H_+}c_2} \int_{-\pi}^{\pi} |\lambda|^{-1+\xi} \, d\lambda \geq \frac{c_1}{2\pi^{1+2H_+}c_2} \cdot \frac{\pi^\xi}{\xi}. \tag{66}$$

Note that we can make the rhs of (66) arbitrarily large if we take $\xi > 0$ sufficiently small. In particular, if we take $\xi \equiv \xi(\iota) \in (0,1)$ such that

$$\frac{c_1}{2\pi^{1+2H_+}c_2} \cdot \frac{\pi^\xi}{\xi} > U_{\vartheta_0}(\vartheta_0) + \iota$$

holds, then we can obtain the following convergence from (65), (66) and Corollary F.8 again:

$$\lim_{n\to\infty} P\left[\inf_{\widetilde{\vartheta}\in\Theta_2^n(\xi,L)} U_{n,0}(\widetilde{\vartheta}) > U_{\vartheta_0}(\vartheta_0) + \iota\right] = 1. \tag{67}$$

Therefore, the conclusion follows from (67). □

Lemma G.5. *For any $\iota > 0$, there exists a constant $L_0 \equiv L_0(\iota, \vartheta_0) > 0$ such that for any $L \geq L_0$,*

$$\lim_{n\to\infty} P\left[\inf_{\widetilde{\vartheta}\in\Theta_2^n(L)} U_{n,0}(\widetilde{\vartheta}) \leq U_{\vartheta_0}(\vartheta_0) + \iota\right] = 0.$$

Proof. Fix $\iota > 0$. Since the inequality

$$U_{n,0}(\widetilde{\vartheta}) \geq \frac{1}{4\pi} \int_{-\pi}^{\pi} \log\left(\widetilde{\nu}^2 f_H(\lambda)\right) d\lambda \geq \log L + \min_{H\in\Theta_H}\left\{\frac{1}{4\pi}\int_{-\pi}^{\pi} \log f_H(\lambda) \, d\lambda\right\}. \tag{68}$$

holds for each $\widetilde{\vartheta} = (H, \widetilde{\nu}) \in \Theta_H \times (L, \infty)$ and we can make the rhs of (68) arbitrarily large if we take $L > 0$ sufficiently large, the conclusion immediately follows. □

Lemma G.6. *For any $\iota > 0$, there exists a constant $L_0 \equiv L_0(\iota, \vartheta_0) > 0$ such that for any $L \geq L_0$,*

$$\lim_{n\to\infty} P\left[\inf_{\widetilde{\vartheta}\in\Theta_1^n(L)} U_{n,0}(\widetilde{\vartheta}) \leq U_{\vartheta_0}(\vartheta_0) + \iota\right] = 0.$$

Proof. Fix $\iota > 0$. Without loss of generality, we can assume $\Theta_1^n(L) = \Theta_H \times [\eta_- \delta_n^{H_+ - H_0}, 1/L)$. At first, we can show that for any $\widetilde{\vartheta} = (H, \widetilde{\nu}) \in \Theta_1^n(L)$,

$$\begin{aligned}
U_{n,0}(\widetilde{\vartheta}) &\geq \frac{1}{4\pi} \int_{-\pi}^{\pi} \log\left(\widetilde{\nu}^2 f_H(\lambda)\right) d\lambda + \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{I_n\left(\lambda, \widetilde{\mathbf{Y}}_n\right)}{\widetilde{\nu}^2 f_H(\lambda) + 2(m_n\delta_n^{2H_0})^{-1}\ell(\lambda)} \, d\lambda \\
&\geq \frac{1}{\widetilde{\nu}^2}\left(\widetilde{\nu}^2\log\widetilde{\nu} + \frac{1}{4\pi}\int_{-\pi}^{\pi} \frac{I_n\left(\lambda, \widetilde{\mathbf{Y}}_n\right)}{f_H(\lambda) + 2(\widetilde{\nu}^2 m_n\delta_n^{2H_0})^{-1}\ell(\lambda)} \, d\lambda\right) + C_1 \\
&\geq L^2\left(\widetilde{\nu}^2\log\widetilde{\nu} + \frac{1}{4\pi}\int_{-\pi}^{\pi} \frac{I_n\left(\lambda, \widetilde{\mathbf{Y}}_n\right)}{f_H(\lambda) + 2(\eta_-^2 m_n\delta_n^{2H_+})^{-1}\ell(\lambda)} \, d\lambda\right) + C_1,
\end{aligned} \tag{69}$$

where

$$C_1 := \min_{H\in\Theta_H}\left\{\frac{1}{4\pi}\int_{-\pi}^{\pi} \log f_H(\lambda) \, d\lambda\right\}.$$

Then it suffices to prove that there exist a constant $r > 0$ such that

$$\lim_{n\to\infty} P\left[\inf_{H\in\Theta_H}\left\{\int_{-\pi}^{\pi}\frac{I_n\left(\lambda,\widetilde{\mathbf{Y}}_n\right)}{f_H(\lambda)+2(\eta_-^2 m_n\delta_n^{2H_+})^{-1}\ell(\lambda)}\,d\lambda\right\} > r\right] = 1. \tag{70}$$

Indeed, if we take sufficiently large $L > 0$, the inequality (69) and (70) imply

$$\lim_{n\to\infty} P\left[\inf_{\widetilde{\vartheta}\in\Theta_1^n(L)} U_{n,0}(\widetilde{\vartheta}) > U_{\vartheta_0}(\vartheta_0) + t\right] = 1$$

in the similar way to the proof of Lemma G.4. Therefore, the conclusion immediately follows. Since Lemma D.1 and the assumption (H.3) yield that there exists a constant $c > 0$ such that for any $n \in \mathbb{N}$, $\lambda \in [-\pi, \pi]\backslash\{0\}$ and $H \in \Theta_H$,

$$f_H(\lambda) + \frac{2}{\eta_-^2 m_n\delta_n^{2H_+}}\ell(\lambda) < c|\lambda|^{-\alpha(H)},$$

the similar argument in the proof of Lemma G.4 implies that it suffices to prove that there exist constants $r_0 > 0$ and $\xi \in (0,1)$ such that

$$\lim_{n\to\infty} P\left[\inf_{H\in\Theta_{H,1}(\xi)}\left\{\int_{-\pi}^{\pi} I_n\left(\lambda,\widetilde{\mathbf{Y}}_n\right)|\lambda|^{\alpha(H)}\,d\lambda\right\} > r_0\right] = 1 \tag{71}$$

instead of (70). In the rest of this proof, we will show (71). Since Proposition F.9 in the case of the function $k_\vartheta(\lambda) := |\lambda|^{\alpha(H)}$ and $\beta_0(H) := -\alpha(H)$ yields that

$$\inf_{H\in\Theta_{H,1}(\xi)}\left\{\int_{-\pi}^{\pi} I_n\left(\lambda,\widetilde{\mathbf{Y}}_n\right)k_\vartheta(\lambda)\,d\lambda\right\} = \inf_{H\in\Theta_{H,1}(\xi)}\left\{\int_{-\pi}^{\pi}\eta_0^2 f_{H_0}(\lambda)k_\vartheta(\lambda)\,d\lambda\right\} + o_P(1)$$

as $n \to \infty$ and we can take sufficiently small $r_0 > 0$ satisfying

$$\inf_{H\in\Theta_{H,1}(\xi)}\left\{\int_{-\pi}^{\pi}\eta_0^2 f_{H_0}(\lambda)k_\vartheta(\lambda)\,d\lambda\right\} > r_0,$$

the convergence (71) follows. Therefore, we finish the proof. $\qquad\square$

As a result, the conclusion of Proposition G.3 follows from (63) and the above three Lemmas. $\qquad\square$

In the rest of this section, we prove the consistency of the estimator $\widehat{\vartheta}_n = (\widehat{H}_n, \widehat{\eta}_n)$ as $n \to \infty$.

Proof of Theorem 2.8. Note that the following equality holds:

$$\log\widehat{\eta}_n - \log\eta_0 = \log\widetilde{\eta}_n - \log\eta_0 - (\log\delta_n)(\widehat{H}_n - H_0).$$

Therefore, it suffices to prove that $\widehat{H}_n - H_0 = o_P(|\log\delta_n|^{-1})$ as $n \to \infty$ from the above equality and the delta method. In the rest of this proof, we attempt to prove the following convergence:

$$(\log\delta_n)\left(\widetilde{\vartheta}_n - \vartheta_0\right) \xrightarrow{n\to\infty} 0 \tag{72}$$

in probability, where $\widetilde{\vartheta}_n := (\widehat{H}_n, \delta_n^{-H_0}\widehat{v}_n)$. At first, Taylor's theorem yields that

$$\nabla U_{n,0}(\widetilde{\vartheta}_n) - \nabla U_{n,0}(\vartheta_0) = \int_0^1 \nabla^2 U_{n,0}\left(\vartheta_0 + u\left(\widetilde{\vartheta}_n - \vartheta_0\right)\right) du \cdot \left(\widetilde{\vartheta}_n - \vartheta_0\right). \tag{73}$$

Here $\nabla U_{n,0}(\widetilde{\vartheta}_n) = o_P(\delta_n^{\psi_1})$ as $n \to \infty$ for any $\psi_1 > 0$ because $\widetilde{\vartheta}_n := (\widehat{H}_n, \delta_n^{-H_0}\widehat{v}_n)$ is a minimizer of $U_{n,0}(\vartheta)$ and Proposition G.3 yields that $\widetilde{\vartheta}_n$ converges to the interior point ϑ_0 as $n \to \infty$. Moreover, Proposition G.2 yields that $\nabla U_{n,0}(\vartheta_0) = o_P(\delta_n^{\psi_2})$ as $n \to \infty$ for a certain constant $\psi_2 > 0$. As a result, the following convergence holds from (73):

$$\int_0^1 \nabla^2 U_{n,0}\left(\vartheta_0 + u\left(\widetilde{\vartheta}_n - \vartheta_0\right)\right) du \cdot (\log \delta_n)\left(\widetilde{\vartheta}_n - \vartheta_0\right) = o_P(1) \quad \text{as} \quad n \to \infty. \tag{74}$$

Since $\nabla^2 U_{\vartheta_0}(\vartheta_0)$ is invertible, the convergence (72) follows from (74), Proposition G.1, Proposition G.3 and Slutsky's theorem. □

H   Proof of Theorem 2.1

In this appendix, we give a proof of Theorem 2.1. Actually, we will show the following limit theorem that is a stronger version of Theorem 2.1.

Theorem H.1. *Under the same assumption in Theorem 2.1, a sequence of càdlàg processes*

$$\sqrt{m_n}\left(\log \widehat{\sigma}_\cdot^2 - \log \int_{\cdot\delta_n}^{(\cdot+1)\delta_n} \sigma_u^2 \, du\right)$$

converges in law to a continuous Gaussian process $G = \{G_s\}_{s\in[0,\infty)}$ given by $G_s := \sqrt{2}(\acute{B}_{s+1} - \acute{B}_s)$, $s \in [0,\infty)$, as $n \to \infty$, where $\acute{B}$ is a standard Brownian motion independent of $\mathcal{F}$.

We recall the martingale functional central limit theorem in Section H.1, a preliminary result used in the proof of Theorem H.1 is summarized in Section H.2 and we prove Theorem H.1 in Section H.3.

H.1   Summary of Martingale Functional Central Limit Theorem

In this subsection, we recall the well-known martingale functional central limit theorem and give its concise proof in the case where local martingales are continuous.

Theorem H.2 (Martingale Functional Central Limit Theorem)**.** *Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\mathbb{F}^n = \{\mathcal{F}_s^n\}_{s\in[0,\infty)}$ be a sequence of filtrations on $(\Omega, \mathcal{F})$ satisfying the usual conditions and $\{Z^n\}_{n\in\mathbb{N}}$ be a sequence of continuous $\mathbb{F}^n$-local martingales. If there exists a continuous function $v : [0,\infty) \to [0,\infty)$ such that for any $s \in [0,\infty)$,*

$$\langle Z \rangle_s \xrightarrow{n\to\infty} v_s \quad \text{in probability}, \tag{75}$$

then a sequence of the $C_{[0,\infty)}$-valued random variables $\{Z^n\}_{n\in\mathbb{N}}$ converges in law to the time-changed Brownian motion $\acute{B}_v$, where $\acute{B}$ is a standard Brownian motion.

Proof. At first, Dambis-Dubins-Schwarz's theorem, see Karatzas and Shreve [25], Theorem 3.4.6, yields that there exists a sequence of standard Brownian motions $\{B^n\}_{n\in\mathbb{N}}$ such that for each $n \in \mathbb{N}$,

$$Z^n = B^n_{\langle Z^n \rangle} \quad P\text{-a.s.}.$$

Note that, since $\langle Z \rangle$ is non-negative and non-decreasing and v is continuous, the assumption (75) implies that for any $s \in [0, \infty)$,

$$\sup_{0 \le u \le s} |\langle Z \rangle_u - v_u| = o_P(1) \ \text{ as } \ n \to \infty \tag{76}$$

by using Theorem VI.2.15 in Jacod and Shiryaev [23]. Moreover, (76) and the Slutsky's theorem yield that $C^2_{[0,\infty)}$-valued process $(B^n, \langle Z^n \rangle)$ converges in law to $(\acute{B}, v)$ as $n \to \infty$, where $\acute{B}$ is a standard Brownian motion. Therefore, the conclusion follows from the above convergence in law and the continuous mapping theorem since $\psi : C^2_{[0,\infty)} \to C_{[0,\infty)}$ defined by $\psi(z, v) := z \circ v$ is continuous in the similar argument to Billingsley [6], p.145. □

Remark H.3. *In Theorem H.2, it is always possible to take a standard Brownian motion $\acute{B}$ independent of $\mathcal{F}$.*

H.2   Notation and Preliminaries

In this subsection, we summarize notation and a preliminary result used in the proof of Theorem H.1. In the rest of this section, we consider a sequence of filtrations $\mathbb{F}^n := \{\mathcal{F}_{s\delta_n}\}_{s\in[0,\infty)}$ and sequences of $\mathbb{F}^n$-martingales $M^n = \{M^n_s\}_{s\in[0,\infty)]}$ and $B^n = \{B^n_s\}_{s\in[0,\infty)]}$ defined by

$$M^n_s := \delta_n^{-\frac{1}{2}} M_{s\delta_n}, \ \ B^n_s := \delta_n^{-\frac{1}{2}} B_{s\delta_n}.$$

Moreover, we set $\tau^n_j := j/m_n$ for $j \in \mathbb{N} \cup \{0\}$ and $N_s[\tau^n] := \max\{j \in \mathbb{N} \cup \{0\} : \tau^n_j \le s\}$ for $s \in [0, \infty)$. In the following lemma, we will show that the assumption of the asset price process S introduced in Section 2.1 in the original article implies the similar conditions introduced in Fukasawa [14]. Note that, by localization argument, we can also assume without loss of generality that κ is bounded and so the volatility process σ^2 is the Hölder-continuous.

Lemma H.4. *For any $k, n \in \mathbb{N}$ and $s \in [0, \infty)$, as $n \to \infty$,*

$$\sup_{j=0,1,\cdots,N_s[\tau^n]} \left| E[(M^n_{\tau^n_{j+1}} - M^n_{\tau^n_j})^{2k} | \mathcal{F}^n_{\tau^n_j}] - \sigma^{2k}_{\tau^n_j \delta_n} (2k-1)!! \left(\frac{1}{m_n} \right)^k \right| = o_P\left(\left(\frac{1}{m_n} \right)^k \right),$$

$$\sup_{j=0,1,\cdots,N_s[\tau^n]} \left| E[(M^n_{\tau^n_{j+1}} - M^n_{\tau^n_j})^{2k-1} | \mathcal{F}^n_{\tau^n_j}] \right| = o_P\left(\left(\frac{1}{m_n} \right)^{k-1/2} \right).$$

Proof. Since we have

$$M^n_s - M^n_v = \sigma_{v\delta_n}(B^n_s - B^n_v) + \delta_n^{-\frac{1}{2}} \int_{v\delta_n}^{s\delta_n} (\sigma_u - \sigma_{v\delta_n}) \, dB_u, \ \ 0 \le v \le s < \infty,$$

the binomial theorem yields that for any $k \in \mathbb{N}$,

$$E[(M^n_{\tau^n_{j+1}} - M^n_{\tau^n_j})^k | \mathcal{F}^n_{\tau^n_j}] - \sigma^k_{\tau^n_j \delta_n} E[(B^n_{\tau^n_{j+1}} - B^n_{\tau^n_j})^k | \mathcal{F}^n_{\tau^n_j}]$$

$$= \sum_{r=1}^{k} {}_k C_r \sigma^{k-r}_{\tau^n_j \delta_n} E\left[(B^n_{\tau^n_{j+1}} - B^n_{\tau^n_j})^{k-r} \left(\delta_n^{-\frac{1}{2}} \int_{\tau^n_j \delta_n}^{\tau^n_{j+1} \delta_n} (\sigma_u - \sigma_{\tau^n_j \delta_n}) \, dB_u \right)^r \Bigg| \mathcal{F}^n_{\tau^n_j} \right]. \tag{77}$$

Since the Brownian motion B enjoys stationary and independent increments properties, we have

$$E[(B^n_{\tau^n_{j+1}} - B^n_{\tau^n_j})^{2k} | \mathcal{F}^n_{\tau^n_j}] = (2k-1)!! \left(\frac{1}{m_n} \right)^k, \ \ E[(B^n_{\tau^n_{j+1}} - B^n_{\tau^n_j})^{2k-1} | \mathcal{F}^n_{\tau^n_j}] = 0 \tag{78}$$

for any $k \in \mathbb{N}$. Moreover, the Burkholder-Davis-Gundy inequality and the Hölder-continuity of σ yield that for any $k \in (0, \infty)$, there exists a constant $C_k > 0$ such that for each $s \in [0, \infty)$,

$$
\sup_{j=1,\cdots,N_s[\tau^n]} E\left[\left|\delta_n^{-\frac{1}{2}} \int_{\tau_j^n \delta_n}^{\tau_{j+1}^n \delta_n} (\sigma_u - \sigma_{\tau_j^n \delta_n}) \, dB_u\right|^k \middle| \mathcal{F}_{\tau_j^n}^n\right] \tag{79}
$$

$$
\leq C_k \sup_{j=1,\cdots,N_s[\tau^n]} E\left[\left|\delta_n^{-1} \int_{\tau_j^n \delta_n}^{\tau_{j+1}^n \delta_n} (\sigma_u - \sigma_{\tau_j^n \delta_n})^2 \, du\right|^{k/2} \middle| \mathcal{F}_{\tau_j^n}^n\right] = o_P\left(\left(\frac{1}{m_n}\right)^{k/2}\right)
$$

as $n \to \infty$. Then the conclusion follows from (78) and (79) by using Cauchy-Schwarz's inequality to the rhs of (77). □

H.3   Proof of Theorem H.1

Before proving Theorem H.1, we will show the following theorem.

Theorem H.5. *Consider sequences of continuous $\mathbb{F}^n$-local martingales $Z^n = \{Z_s^n\}_{s\in[0,\infty)}$ and continuous stochastic processes $\Sigma^n = \{\Sigma_s^n\}_{s\in[0,\infty)}$ respectively given by*

$$
Z_s^n := \sqrt{m_n}\left(\sum_{j=0}^{\infty} \left(M_{\tau_{j+1}^n \wedge s}^n - M_{\tau_j^n \wedge s}^n\right)^2 - \frac{1}{\delta_n}\int_0^{s\delta_n} \sigma_u^2 \, du\right), \quad \Sigma_s^n := \frac{1}{\delta_n}\int_{s\delta_n}^{(s+1)\delta_n} \sigma_u^2 \, du.
$$

Then a sequence of the $C_{[0,\infty)}$-valued random variables $Y^n = \{Y_s^n\}_{s\in[0,\infty)}$ given by $Y_s^n := (Z_{s+1}^n - Z_s^n)/\Sigma_s^n$, $s \in [0,\infty)$, converges in law to the continuous Gaussian process $G = \{G_s\}_{s\in[0,\infty)}$ defined in Theorem H.1.

Proof. Since we have

$$
\frac{1}{\delta_n}\int_0^{s\delta_n} \sigma_u^2 \, du = \langle M^n \rangle_s, \quad s \in [0,\infty),
$$

Itô's formula yields that

$$
Z_s^n = 2\sqrt{m_n}\sum_{j=0}^{\infty} \int_{\tau_j^n \wedge s}^{\tau_{j+1}^n \wedge s} (M_u^n - M_{\tau_j^n \wedge s}^n) \, dM_u^n.
$$

Since Taylor's theorem yields that

$$
\frac{1}{\Sigma_s^n} = \frac{1}{\sigma_{s\delta_n}^2} - \int_0^1 \frac{(\Sigma_s^n - \sigma_{s\delta_n}^2)}{(\sigma_{s\delta_n}^2 + z(\Sigma_s^n - \sigma_{s\delta_n}^2))^2} \, dz
$$

$$
= \frac{1}{\sigma_{\tau_j^n \delta_n}^2} - \int_0^1 \frac{(\sigma_{s\delta_n}^2 - \sigma_{\tau_j^n \delta_n}^2)}{(\sigma_{\tau_j^n \delta_n}^2 + z(\sigma_{s\delta_n}^2 - \sigma_{\tau_j^n \delta_n}^2))^2} \, dz - \int_0^1 \frac{(\Sigma_s^n - \sigma_{s\delta_n}^2)}{(\sigma_{s\delta_n}^2 + z(\Sigma_s^n - \sigma_{s\delta_n}^2))^2} \, dz,
$$

we can decompose Y^n into the following three parts:

$$
Y_s^n = 2\sqrt{m_n}\sum_{j=0}^{\infty} \frac{1}{\Sigma_s^n} \int_{(\tau_j^n \vee s)\wedge(s+1)}^{(\tau_{j+1}^n \vee s)\wedge(s+1)} \left(M_u^n - M_{\tau_j^n \wedge s}^n\right) dM_u^n
$$

$$
= (\widetilde{Z}_{s+1}^n - \widetilde{Z}_s^n) - R_s^n - (Z_{s+1}^n - Z_s^n)\int_0^1 \frac{(\Sigma_s^n - \sigma_{s\delta_n}^2)}{(\sigma_{s\delta_n}^2 + z(\Sigma_s^n - \sigma_{s\delta_n}^2))^2} \, dz \tag{80}
$$

for each $s \in [0,\infty)$, where a sequence of continuous $\mathbb{F}^n$-local martingales $\widetilde{Z}^n = \{\widetilde{Z}_s^n\}_{s\in[0,\infty)}$ and continuous

process $R^n = \{R_s^n\}_{s \in [0, \infty)}$ are given by

$$\widetilde{Z}_s^n := 2\sqrt{m_n} \sum_{j=0}^{\infty} \int_{\tau_j^n \wedge s}^{\tau_{j+1}^n \wedge s} \left(\frac{M_u^n - M_{\tau_j^n \wedge s}^n}{\sigma_{\tau_j^n \delta_n}^2} \right) dM_u^n,$$

$$R_s^n := 2\sqrt{m_n} \sum_{j=0}^{\infty} \int_{(\tau_j^n \vee s) \wedge (s+1)}^{(\tau_{j+1}^n \vee s) \wedge (s+1)} (M_u^n - M_{\tau_j^n \wedge s}^n) dM_u^n \cdot \int_0^1 \frac{(\sigma_{s\delta_n}^2 - \sigma_{\tau_j^n \delta_n}^2)}{(\sigma_{\tau_j^n \delta_n}^2 + z(\sigma_{s\delta_n}^2 - \sigma_{\tau_j^n \delta_n}^2))^2} dz.$$

First of all, we will show that

$$\widetilde{Z}^n \overset{n \to \infty}{\to} \sqrt{2}\acute{B} \text{ in law.} \tag{81}$$

Then Theorem H.2 yields that, in order to prove (81), it suffices to prove that for each $s \in [0, \infty)$,

$$\langle \widetilde{Z}^n \rangle_s = 2s + o_P(1) \text{ as } n \to \infty. \tag{82}$$

By Itô's formula, we have

$$\langle \widetilde{Z}^n \rangle_s = 4m_n \sum_{j=0}^{\infty} \int_{\tau_j^n \wedge s}^{\tau_{j+1}^n \wedge s} \left(\frac{M_u^n - M_{\tau_j^n \wedge s}^n}{\sigma_{\tau_j^n \delta_n}^2} \right)^2 d\langle M^n \rangle_u = \sum_{j=0}^{N_s[\tau^n]} \mathcal{B}_j^n + o_P(1) \text{ as } n \to \infty,$$

where

$$\mathcal{B}_j^n := \frac{2}{3} m_n \frac{(M_{\tau_{j+1}^n}^n - M_{\tau_j^n}^n)^4}{\sigma_{\tau_j^n \delta_n}^4} - \frac{8}{3} m_n \int_{\tau_j^n}^{\tau_{j+1}^n} \frac{(M_u^n - M_{\tau_j^n}^n)^3}{\sigma_{\tau_j^n \delta_n}^4} dM_u^n.$$

Since Lemma H.4 and the Burkholder-Davis-Gundy inequality yield that as $n \to \infty$,

$$\sum_{j=0}^{N_s[\tau^n]} E[\mathcal{B}_j^n | \mathcal{F}_{\tau_j^n}^n] = \frac{2}{3} m_n \sum_{j=0}^{N_s[\tau^n]} \frac{1}{\sigma_{\tau_j^n \delta_n}^4} E[(M_{\tau_{j+1}^n}^n - M_{\tau_j^n}^n)^4 | \mathcal{F}_{\tau_j^n}^n] = 2s + o_P(1),$$

$$\sum_{j=0}^{N_s[\tau^n]} E[|\mathcal{B}_j^n|^2 | \mathcal{F}_{\tau_j^n}^n] = o_P(1)$$

hold, the convergence (82) follows from Lemma 2.3. in [14] and the above two convergences. Therefore, the convergence (81) follows.

In the rest of this proof, we would like to show that the second and third terms of (80) are negligible as $n \to \infty$. Namely, we will prove the following three convergences: for any $s \in [0, \infty)$ and $\iota > 0$,

$$\sup_{0 \le u \le s} \left| \int_0^1 \frac{(\Sigma_u^n - \sigma_{u\delta_n}^2)}{(\sigma_{u\delta_n}^2 + z(\Sigma_u^n - \sigma_{u\delta_n}^2))^2} dz \right| = o_P\left(\delta_n^{H_0 - \iota} \right) \text{ as } n \to \infty, \tag{83}$$

$$\sup_{0 \le s \le u} |Z_s^n| = O_P(1) \text{ as } n \to \infty, \tag{84}$$

$$\sup_{0 \le s \le u} |R_s^n| = o_P(1) \text{ as } n \to \infty. \tag{85}$$

Indeed, if (83), (84) and (85) hold, then the continuous processes appeared in the second and third terms of (80) converge in probability to the function that is identically zero as $n \to \infty$ so that the convergence of Y^n follows from (81) and the continuous mapping theorem.

At first, (83) immediately follows from the Hölder-continuity of the volatility process σ^2. Next, we will

prove (84). In the similar argument to the first term of (80), we can show that

$$\langle Z^n \rangle_s = 2 \int_0^s \sigma_u^4 \, du + o_P(1) \ \text{ as } \ n \to \infty. \tag{86}$$

Then (84) follows from (86) and Doob's inequality. Finally, we will prove (85). By Itô's formula, we have

$$R_s^n = \sum_{j=N_s[\tau^n]+1}^{N_{s+1}[\tau^n]} C_{j,s}^n + o_P(1) \ \text{ as } \ n \to \infty,$$

where

$$C_{j,s}^n := 2\sqrt{m_n} \int_{\tau_j^n}^{\tau_{j+1}^n} (M_u^n - M_{\tau_j^n}^n) \, dM_u^n \cdot \int_0^1 \frac{(\sigma_{s\delta_n}^2 - \sigma_{\tau_j^n \delta_n}^2)}{(\sigma_{\tau_j^n \delta_n}^2 + z(\sigma_{s\delta_n}^2 - \sigma_{\tau_j^n \delta_n}^2))^2} \, dz.$$

Since Lemma H.4 and the Burkholder-Davis-Gundy inequality yield

$$\sum_{j=N_s[\tau^n]+1}^{N_{s+1}[\tau^n]} E[C_{j,s}^n | \mathcal{F}_{\tau_j^n}^n] = 2\sqrt{m_n} \sum_{j=N_s[\tau^n]+1}^{N_{s+1}[\tau^n]} \int_0^1 \frac{(\sigma_{s\delta_n}^2 - \sigma_{\tau_j^n \delta_n}^2)}{(\sigma_{\tau_j^n \delta_n}^2 + z(\sigma_{s\delta_n}^2 - \sigma_{\tau_j^n \delta_n}^2))^2} \, dz E\left[\int_{\tau_j^n}^{\tau_{j+1}^n} (M_u^n - M_{\tau_j^n}^n) \, dM_u^n \bigg| \mathcal{F}_{\tau_j^n}^n \right] = 0,$$

$$\sum_{j=N_s[\tau^n]+1}^{N_{s+1}[\tau^n]} E[|C_{j,s}^n|^2 | \mathcal{F}_{\tau_j^n}^n]$$

$$= 4m_n \sum_{j=N_s[\tau^n]+1}^{N_{s+1}[\tau^n]} \left(\int_0^1 \frac{(\sigma_{s\delta_n}^2 - \sigma_{\tau_j^n \delta_n}^2)}{(\sigma_{\tau_j^n \delta_n}^2 + z(\sigma_{s\delta_n}^2 - \sigma_{\tau_j^n \delta_n}^2))^2} \, dz \right)^2 E\left[\left(\int_{\tau_j^n}^{\tau_{j+1}^n} (M_u^n - M_{\tau_j^n}^n) \, dM_u^n \right)^2 \bigg| \mathcal{F}_{\tau_j^n}^n \right] = o_P(1) \ \text{ as } \ n \to \infty,$$

the convergence (85) follows from an easy modification of Lemma 2.3. in [14] and the above two convergences. Therefore, we finish the proof. □

Let us embed the realized variance $\hat{\sigma}^2$ into a continuous-time stochastic process

$$\hat{\sigma}_s^2 := \sum_{j=0}^{m_n-1} \left| \log S_{\delta_n \tau_{\lfloor m_n s \rfloor + j + 1}^n} - \log S_{\delta_n \tau_{\lfloor m_n s \rfloor + j}^n} \right|^2, \ s \in [0, \infty).$$

Then we can obtain the following limit theorem.

Theorem H.6. *A sequence of càdlàg processes* $\widetilde{Y}^n = \{\widetilde{Y}_s^n\}_{s \in [0,\infty)}$ *given by*

$$\widetilde{Y}_s^n := \sqrt{m_n} \left(\frac{\hat{\sigma}_s^2 - \int_{s\delta_n}^{(s+1)\delta_n} \sigma_u^2 \, du}{\int_{s\delta_n}^{(s+1)\delta_n} \sigma_u^2 \, du} \right), \ s \in [0, \infty),$$

converges in law to the continuous Gaussian process $G = \{G_s\}_{s \in [0,\infty)}$ *defined in Theorem H.1.*

Proof. Note that we have

$$\frac{\sqrt{m_n}}{\delta_n} \left(\sum_{j=0}^{\infty} (\log S_{\tau_{j+1}^n \wedge (s\delta_n)} - \log S_{\tau_j^n \wedge (s\delta_n)})^2 - \int_0^{s\delta_n} \sigma_u^2 \, du \right)$$

$$= Z_s^n + 2\sqrt{m_n} \sum_{j=0}^{\infty} (M_{\tau_{j+1}^n \wedge s}^n - M_{\tau_j^n \wedge s}^n)(A_{\tau_{j+1}^n \wedge s}^n - A_{\tau_j^n \wedge s}^n) + \sqrt{m_n} \sum_{j=0}^{\infty} (A_{\tau_{j+1}^n \wedge s}^n - A_{\tau_j^n \wedge s}^n)^2,$$

48

where $A_s^n := \delta_n^{-1/2} A_{s\delta_n}$, $s \in [0, \infty)$. By using Lemma H.4, we can show that

$$\sqrt{m_n} \sum_{j=0}^{\infty} (A_{\tau_{j+1}^n \wedge s}^n - A_{\tau_j^n \wedge s}^n)^2 = o_P(1) \text{ as } n \to \infty,$$

$$2\sqrt{m_n} \sum_{j=0}^{\infty} (M_{\tau_{j+1}^n \wedge s}^n - M_{\tau_j^n \wedge s}^n)(A_{\tau_{j+1}^n \wedge s}^n - A_{\tau_j^n \wedge s}^n) = o_P(1) \text{ as } n \to \infty$$

uniformly in $u \in [0, s]$ for any $s > 0$ in the similar way to the proof of Lemma 3.9. and Theorem 3.10. in [14] respectively. Then we obtain

$$\frac{\sqrt{m_n}}{\delta_n}\left(\hat{\sigma}_s^2 - \int_{s\delta_n}^{(s+1)\delta_n} \sigma_u^2 \, du\right) = \frac{\sqrt{m_n}}{\delta_n}\left(\sum_{j=0}^{\infty}(\log S_{\tau_{j+1}^n \wedge \{(s+1)\delta_n\}} - \log S_{\tau_j^n \wedge \{(s+1)\delta_n\}})^2 - \int_0^{(s+1)\delta_n} \sigma_u^2 \, du\right)$$

$$- \frac{\sqrt{m_n}}{\delta_n}\left(\sum_{j=0}^{\infty}(\log S_{\tau_{j+1}^n \wedge (s\delta_n)} - \log S_{\tau_j^n \wedge (s\delta_n)})^2 - \int_0^{s\delta_n} \sigma_u^2 \, du\right) + o_P(1)$$

$$= (Z_{s+1}^n - Z_s^n) + o_P(1)$$

as $n \to \infty$ uniformly in $s \in [0, u]$ for any $u > 0$. Therefore, the conclusion follows from Theorem H.5 and the continuous mapping theorem since $1/\Sigma_s^n = O_P(1)$ as $n \to \infty$ uniformly in $u \in [0, s]$ for any $s > 0$. □

In the end of this appendix, we prove Theorem H.1 by using Theorem H.6.

Proof of Theorem H.1. By Taylor's theorem, we obtain

$$\sqrt{m_n}\left(\log \hat{\sigma}_s^2 - \log \int_{s\delta_n}^{(s+1)\delta_n} \sigma_u^2 \, du\right) = \sqrt{m_n} \log\left(1 + \frac{\hat{\sigma}_s^2 - \int_{s\delta_n}^{(s+1)\delta_n} \sigma_u^2 \, du}{\int_{s\delta_n}^{(s+1)\delta_n} \sigma_u^2 \, du}\right)$$

$$= \sqrt{m_n}\left(\frac{\hat{\sigma}_s^2 - \int_{s\delta_n}^{(s+1)\delta_n} \sigma_u^2 \, du}{\int_{s\delta_n}^{(s+1)\delta_n} \sigma_u^2 \, du}\right) + \sqrt{m_n}\left(\frac{\hat{\sigma}_s^2 - \int_{s\delta_n}^{(s+1)\delta_n} \sigma_u^2 \, du}{\int_{s\delta_n}^{(s+1)\delta_n} \sigma_u^2 \, du}\right)^2 \int_0^1 (1-z)\left\{1 + z\left(\frac{\hat{\sigma}_s^2 - \int_{s\delta_n}^{(s+1)\delta_n} \sigma_u^2 \, du}{\int_{s\delta_n}^{(s+1)\delta_n} \sigma_u^2 \, du}\right)\right\}^{-2} dz$$

for each $s \in [0, \infty)$. Since we have

$$\sup_{0 \le s \le s_0}\left|\int_0^1 (1-z)\left\{1 + z\left(\frac{\hat{\sigma}_s^2 - \int_{s\delta_n}^{(s+1)\delta_n} \sigma_u^2 \, du}{\int_{s\delta_n}^{(s+1)\delta_n} \sigma_u^2 \, du}\right)\right\}^{-2} dz\right| = O_P(1) \text{ as } n \to \infty$$

for each $s_0 \in [0, \infty)$, the conclusion follows from Theorem H.6 and the continuous mapping theorem. □

I Approximate Formula of Estimation Function $U_n(H, \nu)$

In this appendix, we derive the approximate formula of the estimation function (15) in the original article. Since the spectral density $g_{H,\nu}(\lambda)$ and the periodogram $I_n(\lambda)$ are symmetric with respect to $\lambda \in [-\pi, \pi]$, we

have

$$U_n(H, \nu) = \frac{1}{2\pi} \int_0^\pi \left(\log g_{H,\nu}(\lambda) + \frac{I_n(\lambda, \mathbf{Y}_n)}{g_{H,\nu}(\lambda)} \right) d\lambda$$

$$= \frac{1}{2\pi} \int_\psi^\pi \left(\log g_{H,\nu}(\lambda) + \frac{I_n(\lambda, \mathbf{Y}_n)}{g_{H,\nu}(\lambda)} \right) d\lambda + B_{H,\nu}^1(\psi) + B_{H,\nu}^2(\psi)$$

for any $\psi \in (0, \pi]$, where

$$B_{H,\nu}^1(\psi) := \frac{1}{2\pi} \int_0^\psi \log g_{H,\nu}(\lambda) \, d\lambda, \quad B_{H,\nu}^2(\psi) := \frac{1}{2\pi} \int_0^\psi \frac{I_n(\lambda, \mathbf{Y}_n)}{g_{H,\nu}(\lambda)} \, d\lambda.$$

In the rest of this subsection, we will show $B_{H,\nu}^1(\psi) \approx A_{H,\nu}^1(\psi)$ and $B_{H,\nu}^2(\psi) \approx A_{H,\nu}^2(\psi)$ as $\psi \downarrow 0$. At first, we consider the first approximation. Note that the Taylor expansion yields that

$$g_{H,\nu}(\lambda) = \nu^2 C_H |\lambda|^{1-2H} + \frac{|\lambda|^2}{m\pi} + O(|\lambda|^{3+2H}) \quad \text{as } |\lambda| \to 0. \tag{87}$$

Then we obtain the first approximation from the Taylor expansion as $\psi \downarrow 0$ as follows:

$$B_{H,\nu}^1(\psi) \approx \frac{1}{2\pi} \int_0^\psi \log \left(\nu^2 C_H \lambda^{1-2H} + \frac{\lambda^2}{m\pi} \right) d\lambda$$

$$= \frac{1}{2\pi} \left\{ \psi \log(\nu^2 C_H) + \psi(\log \psi - 1)(1 - 2H) + \int_0^\psi \log \left(1 + \frac{1}{\nu^2 C_H m\pi} \lambda^{1+2H} \right) d\lambda \right\}$$

$$\approx \frac{1}{2\pi} \left\{ \psi \log(\nu^2 C_H) + \psi(\log \psi - 1)(1 - 2H) + \frac{\psi^{2+2H}}{\nu^2 C_H m\pi(2 + 2H)} \right\}.$$

Next we consider the second approximation. Since $g_{H,\nu}$ is an even function, $B_{H,\nu}^2(\psi)$ is represented by

$$B_{H,\nu}^2(\psi) = \frac{1}{2\pi} \left(b_{H,\nu}(0, \psi) \widehat{\gamma}_n(0) + 2 \sum_{\tau=1}^{n-1} b_{H,\nu}(\tau, \psi) \widehat{\gamma}_n(\tau) \right),$$

where

$$b_{H,\nu}(\tau, \psi) := \frac{1}{2\pi} \int_0^\psi \frac{\cos(\tau\lambda)}{g_{H,\nu}(\lambda)} \, d\lambda.$$

Since the Taylor expansion as $\psi \downarrow 0$ yields that

$$b_{H,\nu}(\tau, \psi) = \frac{1}{2\pi} \sum_{j=0}^\infty \frac{(-1)^j \tau^{2j}}{(2j)!} \int_0^\psi \frac{\lambda^{2j}}{g_{H,\nu}(\lambda)} \, d\lambda \tag{88}$$

$$\approx \frac{1}{2\pi} \sum_{j=0}^\infty \frac{(-1)^j \tau^{2j}}{(2j)!} \int_0^\psi \frac{\lambda^{2j}}{\nu^2 C_H |\lambda|^{1-2H} + \frac{|\lambda|^2}{m\pi}} \, d\lambda$$

$$\approx \frac{1}{2\pi} \sum_{j=0}^\infty \frac{(-1)^j \tau^{2j}}{(2j)!} \int_0^\psi \frac{\lambda^{-1+2j+2H}}{\nu^2 C_H} \left(1 - \frac{1}{\nu^2 C_H m\pi} \lambda^{1+2H} \right) d\lambda$$

$$= \frac{1}{2\pi} \sum_{j=0}^\infty \frac{(-1)^j \tau^{2j}}{(2j)!} \frac{1}{\nu^2 C_H} \left(\frac{\psi^{2j+2H}}{2j + 2H} - \frac{\psi^{1+2j+4H}}{\nu^2 C_H m\pi(1 + 2j + 4H)} \right), \tag{89}$$

50

we obtain the second approximation when the series in (89) is truncated after finite terms. Note that the truncation error of the Taylor expansion in (88) is dominated as follows:

$$\sup_{\tau \in \{0,1,\cdots,n-1\}} \left| b_{H,\nu}(\tau,\psi) - \frac{1}{2\pi} \sum_{j=0}^{J} \frac{(-1)^j}{(2j)!} \int_0^{\psi} \frac{(\tau\lambda)^{2j}}{g_{H,\nu}(\lambda)} \, \mathrm{d}\lambda \right| \leq \frac{(n\psi)^{2J+1-1}}{(2J+1)!} \cdot \frac{1}{2} \int_0^{\psi} \frac{1}{g_{H,\nu}(\lambda)} \, \mathrm{d}\lambda$$

for any $J \in \mathbb{N}$ and $\psi > 0$. As a result, for fixed $n \in \mathbb{N}$, we can make the truncation error arbitrary small uniformly with respect to $\tau \in \{0,1,\cdots,n-1\}$ as $J \in \mathbb{N}$ is taken sufficiently large even in the case of the finite sample.