Fractional Brownian Motions in Financial Econometrics

WEILIN XIAO *Zhejiang University*
XILI ZHANG *Zhejiang University*

Fractional Brownian motion is a continuous-time zero mean Gaussian process with stationary increments. It has gained much attention in empirical finance and asset pricing. For example, it has been used to model the time series of volatility and interest rates. This chapter first introduces the basic properties of fractional Brownian motions and then reviews the statistical models driven by the fractional Brownian motions that have been used in financial econometrics such as the fractional Ornstein-Uhlenbeck model and the fractional stochastic volatility models. We also review the parameter estimation methods proposed in the literature. These methods are based on either continuous-time observations or discrete-time observations.

7.1 Introduction

Continuous-time processes driven by the standard Brownian motion are widely used to describe the price dynamics of financial assets or portfolios since the celebrated works of Black and Scholes (1973); Merton (1973). Even though the option pricing model from Black and Scholes (1973) remains one of the most successful methods for pricing European options, the assumptions of the independence of asset returns (e.g., Greene and Fielitz (1977); Willinger et al. (1999)) and the constant volatility of stock returns (e.g., Comte and Renault (1998); Corlay et al. (2014)) are too strong to be true.

To explain the dependence properties of asset returns, Mandelbrot and Van Ness (1968) introduced fractional Brownian motion (fBm) that models the dynamics of stock prices. If the standard Brownian motion is replaced with the fBm in the Black-Scholes model, the resulting model is a geometric fractional Brownian motion (gfBm) that was shown to be more in line with the behavior of stock markets in Cont (2001). Unfortunately, fBm is considered inadequate for modeling stock returns because it is not a semimartingale when the Hurst parameter $H \neq \frac{1}{2}$. Consequently, the arbitrage opportunity in the fBm-driven Black-Scholes model and approaches to remove arbitrage in the context of fBm have been extensively investigated in the literature, including Rogers (1997); Sottinen (2001); Cheridito et al. (2003); Elliott and Van Der Hoek (2003); Björk and Hult (2005); Rostek (2009). Additionally, when using the fBm in financial modeling, one must define an Itô-type formula and a risk-neutral measure, which is also true for Brownian motion.

Dedicated to fBm, Duncan et al. (2000) is the first study that proposed the Wick product approach to defining fractional stochastic integrals with respect to fBm, called the Wick-Itô integral. Hu and Øksendal (2003); Elliott and Van Der Hoek (2003); Hu et al. (2003) extended the idea of the Wick product and developed a fractional white noise calculus that is applied to option pricing and portfolio optimization. However, at the same time, severe critiques arose concerning the economic meaning of Wick products (e.g., Björk and Hult (2005)).

To explain the stylized facts in the implied volatility surface (e.g., volatility smile and skew), many stochastic volatility models have been developed in the literature, including Hull and White (1987); Scott (1987); Stein and Stein (1991); Heston (1993); Bates (1996); Duffie et al. (2000); Schöbel and Zhu (1999). More details regarding these stochastic volatility models can be found in Fouque et al. (2000). Both academics and practitioners have recognized the importance of the stochastic volatility models mentioned above for pricing options; see Bakshi et al. (1997) for an example. Nevertheless, none of these stochastic volatility models are problem-free. One of the key disadvantages among all the aforementioned stochastic volatility models is that the volatility process is driven by the standard Brownian motion. Therefore, the autocorrelation function of the volatility exponentially decays. However, many empirical studies argue that the decay in the autocorrelation function is better modelled by a power function. Unsurprisingly, in the discrete-time volatility literature, Baillie (1996); Baillie et al. (1996); Andersen et al. (2003); Shi and Yu (2022) proposed the autoregressive fractionally integrated moving average model for volatility. In the continuous-time volatility literature, Comte and Renault (1998); Aït-Sahalia and Mancini (2008); Comte et al. (2012); Bayer et al. (2016); Gatheral et al. (2018); Bennedsen et al. (2021); Liu et al. (2020) proposed models based on the fBm for volatility.

As a non-stationary process, the fBm is not suitable for modeling stationary time series. Consequently, the fractional Ornstein–Uhlenbeck process (fOUp) and the fractional Vasicek model (fVm) were developed to model stationary financial time series such as volatility and interest rates (see, for example, Comte and Renault (1998); Fink et al. (2013); Aït-Sahalia and Mancini (2008); Comte et al. (2012); Bayer et al. (2016); Gatheral et al. (2018); Bennedsen et al. (2021); Bolko et al. (2022)). Recently, an econometric analysis of the fOUp and fVm has received considerable attentions, including parameter estimation and asymptotic theory.

This chapter first introduces the basic properties of fractional Brownian motion and then reviews the fOUp, fVm, and stochastic volatility models driven by the fBm. We also review the parameter estimation methods proposed in the literature and the associated asymptotic theory. The estimation methods are based on either continuous-time observations or discrete-time observations.

The remainder of the chapter is organized as follows: Section 7.2 reviews some basic properties of the fBm, fOUp, and fVm and discusses their application when modeling financial time series. Section 7.3 summarizes some well-known stochastic volatility models driven by Brownian motion and by the fBm. Section 7.4 reviews alternative estimators of the drift parameters in the fVm based on continuous-time observations. Section 7.5 reviews the generalized method of moments (GMM) for the fVm based on discrete-time observations. Section 9.5 provides conclusions.

7.2 Fractional Brownian Motion and Related Stochastic Processes

7.2.1 Fractional Brownian Motion

It is well known that standard Brownian motion has independent increments that are normally distributed. As a generalization of standard Brownian motion, the fBm is self-similar and its increments are normally distributed and stationarity.[1] In this section, we present the definition of the fBm and introduce some of its basic properties by comparing it with standard Brownian motion.

The fBm was first introduced by Kolmogorov (1940) in 1940, and then Mandelbrot and Van Ness (1968) provided its stochastic integral (moving average) representation and obtained many of its properties. The definition of the fBm is as follows:

Definition 7.2.1 *Consider probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and let* B^H *be an fBm in this space. Then, a standard fBm is a centered and continuous Gaussian process with the Hurst parameter* $H \in (0, 1)$*, which has stationary increments and the following covariance function*

$$\mathbb{E}\left(B_t^H B_s^H\right) = R_H(s,t) \equiv \frac{1}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right), \tag{7.2.1}$$

with $t,\ s > 0$.

From Samorodnitsky and Taqqu (1994), we can see that the variance-covariance matrix implied by Equation (7.2.1) is non-negative definite, and hence, Equation (7.2.1) is a proper covariance function. In Equation (7.2.1), the coefficient H is known as the Hurst parameter and is named for British climatologist H.E. Hurst, who proposed the classical test to detect the long memory phenomenon of reservoir control near the Nile River Dam (see Hurst (1951)). We can obtain the standard Brownian motion by setting $H = \frac{1}{2}$, and in this case, the covariance function becomes $\min\{t, s\}$, which is the covariance function of the standard Brownian motion. Some important properties of the fBm are provided by the following proposition.

Proposition 7.2.1 *The fBm has the following properties:*

(i) $B_t^H \sim \mathcal{N}\left(0, t^{2H}\right)$.

(ii) $\mathbb{E}\left(B_t^H - B_s^H\right)^2 = |t-s|^{2H}$ *for* $t, s > 0$*;*

(iii) *For any* $s, t \geq 0$ *we have* $\mathbb{E}\left|B_t^H - B_s^H\right|^2 = |t-s|^{2H}$*. In particular, the fBm has* δ*-Hölder continuous paths for any* $\delta < H$*.*[2]

---

[1] A process $X = (X_t)_{t\geq 0}$ has stationary increments if we have $X_{t+h} - X_t \overset{d}{=} X_h - X_0$ for every $h, t \geq 0$.

[2] For a process $X : [0, +\infty) \times \Omega \to \mathbb{R}$ if for all $T > 0$, there exist $\alpha, \beta, C > 0$ such that $\mathbb{E}\left[|X_t - X_s|^\alpha\right] \leq C|t-s|^{1+\beta}, \forall 0 \leq t, s \leq T$, then there exists a version of X, which is Hölder continuous of order $\gamma \in \left[0, \frac{\beta}{\alpha}\right)$ almost surely.

(iv) *The fBm is not a Markov process except that* $H = \frac{1}{2}$.[3]

(v) *The fBm exhibits long-range dependence if* $H > \frac{1}{2}$ *and is a short-memory process if* $H < \frac{1}{2}$.[4]

(vi) *Let* $\pi_n = \left\{t_k = \frac{k}{2^n} : k = 0, \dots, 2^n\right\}$ *be a partition of* $[0,1]$ *and* $n \in \mathbb{N}$. *Then, we have*

$$\lim_{n\to\infty} \sum_{k=1}^{n} \left|B_{t_k}^H - B_{t_{k-1}}^H\right|^p = \begin{cases} \infty, & \text{if} \quad pH < 1\,, \\ \mathbb{E}\left|B_1^H\right|^p, & \text{if} \quad pH = 1\,, \\ 0, & \text{if} \quad pH > 1\,, \end{cases} \tag{7.2.2}$$

in probability.

(vii) *The fBm is not a semimartingale.*

(viii) *The fBm is nowhere differentiable. That is, for every* $s \in [0, \infty]$, *we have* $\mathbb{P}\left(\limsup_{t\to s}\left|\frac{B_t^H - B_s^H}{t-s}\right| = \infty\right) = 1$.

Using the properties provided in Proposition 7.2.1, the fBm can be represented as an integral of a deterministic kernel with respect to standard Brownian motion. In the literature, there exist several representations of the fBm as a Wiener integral, which can be described as follows:

(i) Mandelbrot and Van Ness (1968):

$$\begin{aligned} B_t^H &= \int_{-\infty}^{\infty} \left\{(t-u)_+^{H-\frac{1}{2}} - (-u)_+^{H-\frac{1}{2}}\right\} dB_u \\ &= \frac{1}{c_H}\left\{\int_{-\infty}^{0} \left[(t-u)^{H-\frac{1}{2}} - (-u)^{H-\frac{1}{2}}\right] dB_u + \int_0^t (t-u)^{H-\frac{1}{2}} dB_u\right\}\,, \end{aligned}$$

where $(x)_+ = \max(x, 0)$, $c_H = \left[\frac{1}{2H} + \int_0^\infty \left((1+s)^{H-1/2} - s^{H-\frac{1}{2}}\right)^2 ds\right]^{\frac{1}{2}}$ and B_u is a standard Brownian motion.

(ii) Samorodnitsky and Taqqu (1994):

$$B_t^H = \left(\frac{1}{C_2(H)}\int_{\mathbb{R}} \frac{e^{itx} - 1}{ix} |x|^{-\left(H-\frac{1}{2}\right)} dB_x\right)_{t\in\mathbb{R}}\,,$$

where $i = \sqrt{-1}$, $C_2(H) = \sqrt{\frac{\pi}{H\Gamma(2H)\sin(H\pi)}}$ and $\Gamma(\cdot)$ denote the gamma function.

(iii) Norros et al. (1999):

$$B_t^H = \int_0^t z_H(t,s) dB_s\,,$$

where B_s is a standard Brownian motion and the deterministic kernel is

$$z_H(t,s) = \begin{cases} \alpha_H\Big(-\left(H-\frac{1}{2}\right) s^{\frac{1}{2}-H} \int_s^t u^{H-\frac{3}{2}} (u-s)^{H-\frac{1}{2}} du \\ \qquad + \left(\frac{t}{s}\right)^{H-\frac{1}{2}} (t-s)^{H-\frac{1}{2}}\Big) \quad for\ H \in (0, \frac{1}{2}) \\ (H-\frac{1}{2})\alpha_H s^{\frac{1}{2}-H} \int_s^t (v-s)^{H-\frac{3}{2}} v^{H-\frac{1}{2}} dv \quad for\ H \in (\frac{1}{2}, 1) \end{cases}\,,$$

and $\alpha_H = \left(\frac{2H\Gamma(\frac{3}{2}-H)}{\Gamma(H+\frac{1}{2})\Gamma(2-2H)}\right)^{\frac{1}{2}}$.

[3] A Gaussian process X with covariance function $R(s,t)$ is Markovian if and only if $R(s,u) = \frac{R(s,t)R(t,u)}{R(t,t)}$ for every $s \leq t \leq u$.

[4] A process X exhibits long-range dependence (or it is a long-memory process) if $\sum_{n\geq 0} \rho_H(n) = \infty$, where $\rho_H(n) = \mathbb{E}\left(X_1 - X_0\right)\left(X_{n+1} - X_n\right)$. Otherwise, if $\sum_{n\geq 0} r_n < \infty$, we can say that X is a short-memory process.

A more detailed description of the representations can be found in Nourdin (2012). It is well known that the Hurst parameter H determines the main properties of the fBm such as self-similarity, regularity of sample paths, and long memory. Therefore, the fBm has been employed in different fields exhibiting long-range dependence and anti-persistence, including hydrology, biology, medicine, traffic networks, economics, and finance, among others. In financial mathematics, the fBm was first introduced to model stock price since a large number of empirical studies of financial time series have found long-range dependence properties in asset returns (e.g., Greene and Fielitz (1977); Lo and MacKinlay (1988); Willinger et al. (1999); Cont (2005)). It has been suggested as a replacement for Brownian motion in the geometric Brownian motion model for stock prices. Consequently, some option pricing models have been proposed to extend the classical Black-Scholes model, such as Hu and Øksendal (2003); Elliott and Van Der Hoek (2003); Elliott and Chan (2004); Mishura (2008); Rostek (2009). However, since the fBm is not a semimartingale when the Hurst parameter $H \neq \frac{1}{2}$, there exists an arbitrage opportunity in the fractional Black-Scholes model. In the literature, Rogers (1997) first stated that the fBm is an unsuitable candidate for usage in financial models. In fact, for all Hurst parameters $H \neq \frac{1}{2}$, Rogers (1997) derived the existence of arbitrage possibilities in a fractional Bachelier type model. Then, Sottinen (2001); Cheridito (2001); Bender and Elliott (2004); Björk and Hult (2005) demonstrated that the Black-Scholes model driven by the fBm allows arbitrage in a number of ways.

To exclude arbitrage, a new stochastic integral of the fBm, namely the Wick integral, can be used (see Duncan et al. (2000); Hu and Øksendal (2003); Hu et al. (2003)). From a purely mathematical point of view, the Wick integral, which was proposed by Hu and Øksendal (2003); Hu et al. (2003), is formally correct and accurate. However, until now, no reasonable economic interpretation for the Wick integral has been provided.

Fractional Gaussian noise (fGn), which is the stationary increment of the fBm with mean zero, can be defined as $G_t^H = B_{t+1}^H - B_t^H$ with $t \in \mathbb{Z}^+$. Using (7.2.1), we can easily obtain that the covariance of G_t^H and G_{t+k}^H is

$$\operatorname{Cov}\left(G_t^H, G_{t+k}^H\right) = \gamma_G(k) = \frac{1}{2}\left(|k+1|^{2H} + |k-1|^{2H} - 2|k|^{2H}\right), \quad k \in \mathbb{Z}, \tag{7.2.3}$$

where $H \in (0,1)$ is referred to as the Hurst exponent and is a measure of the long-term correlation between the discrete time points. Using the Taylor expansion, we can obtain the asymptotic behavior of $\gamma_G(k)$. Let $\gamma_G(k) = \frac{1}{2}k^{2H} g\left(k^{-1}\right)$ with $g(x) = (1+x)^{2H} - 2 + (1-x)^{2H}$. If $0 < H < 1$ and $H \neq \frac{1}{2}$, then the first non-zero term in the Taylor expansion of $g(x)$, expanded at the origin, is equal to $2H(2H-1)x^2$. Therefore, as k tends to infinity, $\gamma_G(k)$ is equivalent to $H(2H-1)k^{2H-2}$, which implies that the covariance function of the fGn has a power-law decay. The result of (7.2.3) implies that the fGn reduces to uncorrelated white noise when $H = \frac{1}{2}$. When $H > \frac{1}{2}$, the fGn has a positive correlation reflecting a persistent autocorrelation structure. Similarly, the covariance is negative when $H < \frac{1}{2}$, and the resulting process is then referred to as being anti-persistent. Note that, for $H = 1$, (7.2.3) implies $\gamma_G(k) \equiv 1$. Thus, all correlations are equal to 1 no matter how far apart in time the observations are. This case is hardly of any practical importance. For $H > 1$, we can see that $g\left(k^{-1}\right)$ diverges to infinity. This behavior contradicts the fact that $\rho(k)$ must be between -1 and 1. As a consequence, if covariances exist and $\lim_{k\to\infty} \gamma_G(k) = 0$, then $0 < H < 1$. For $\frac{1}{2} < H < 1$, the process has long-range dependence; for $H = \frac{1}{2}$, the observations are uncorrelated; and for $0 < H < \frac{1}{2}$, the process has short-range dependence and the correlations sum to zero. Some important properties of the fGn are provided by the following proposition, whose proofs can be found in Beran (1994).

Proposition 7.2.2 *The fGn has the following properties:*

(i) *The spectral density is*

$$f_G(\lambda; H) = \frac{\sin(\pi H)\Gamma(2H+1)(1-\cos(\lambda))}{\pi} \sum_{j\in\mathbb{Z}} |\lambda + 2j\pi|^{-2H-1},$$

which has a pole at $\lambda = 0$ *and asymptotic behavior*

$$f_G(\lambda; H) \sim \mathcal{O}\left(|\lambda|^{1-2H}\right) \quad \text{as} \quad \lambda \to 0\,.$$

In particular, if $H \in (0, 1/2)$*, then* $f_G(\lambda; H) \to 0$ *as* $\lambda \to 0$ *and if* $H \in (\frac{1}{2}, 1)$*,* $f_G(\lambda; H) \to \infty$ *as* $\lambda \to 0$*.*

(ii) *For large lags* k*, the behavior of the autocorrelation function (7.2.3) is*

$$\gamma_G(k) \sim H\,(2H-1)\,k^{2H-2} \tag{7.2.4}$$

as $k \to \infty$*, which implies that*

$$\sum_{k\in\mathbb{Z}} \gamma_G(k) = \infty$$

for $H \in (\frac{1}{2}, 1)$*.*

(iii) *For all* $H \in (0, 1)$*, we have* $G_t^H \sim N(0, 1)$ *and*

$$G_t^H - G_0^H \sim \mathcal{N}\left(0, 2 - |t-1|^{2H} - |t+1|^{2H} + 2|t|^{2H}\right)$$

Moreover, we have $G_t^H - G_0^H \Rightarrow \mathcal{N}(0, 2)$ *as* $t \to \infty$*.*

7.2.2 The Fractional Ornstein—Uhlenbeck Process and the Fractional Vasicek Model

Based on the discussion in the subsection above, we can see that difficulties arise when using the fBm to model financial asset returns. However, using the fBm to model the time series of volatility or interest rates is fruitful. In fact, a popular statistical model is the fVm, which extends the fOUp. Both the fOUp and fVm are driven by the fBm.

As a solution to the Langevin stochastic differential equation with the fBm, the fOUp was proposed by Cheridito et al. (2003) and can be defined as follows:

$$dX_t = -\kappa X_t dt + \sigma dB_t^H, \tag{7.2.5}$$

where $\kappa, \sigma > 0$.

Using the path-wise Riemann-Stieltjes integral, Cheridito et al. (2003) obtained a unique path-wise solution to Model (7.2.5) as

$$X_t = e^{-\kappa t} X_0 + \sigma \int_0^t e^{-\kappa(t-s)} dB_s^H, \tag{7.2.6}$$

where $\int_0^t e^{-\kappa(t-s)} dB_s^H$ is a Wiener integral with respect to the fBm. One can also rewrite the unique solution to (7.2.5) as

$$X_t = X_0 e^{-\kappa t} + \sigma\kappa e^{-\kappa t} \int_0^t e^{\kappa s} B_s^H ds + \sigma B_t^H. \tag{7.2.7}$$

The representation is obtained via integration by parts, that is, $\int_0^t e^{\kappa s} dB_s^H = -\kappa \int_0^t e^{\kappa s} B_s^H ds + e^{\kappa t} B_t^H$.

Here, we review some important properties of the fOUp using the following proposition:

Proposition 7.2.3 *The fOUp has the following properties:*

(i) X_t *is H*´*older continuous with H*´*older exponents* $\gamma \in (0, H)$*.*

(ii) Let $\mathfrak{Z}$ *be the class of nonnegative random variables* ζ*. Then there exists* $C > 0$ *independent of* M *such that* $\mathbb{E}\exp\left\{x\zeta^2\right\} < \infty$ *for any* $0 < x < C$ *such that* $\sup_{0\le s\le t}|X_s| \le \left(Ce^{\alpha s} + s^H \log^2 s\right)\zeta$*.*

(iii) The random variable X_t *has a normal distribution with mean* $X_0 e^{-\kappa t}$ *and variance* $\operatorname{Var} X_t = H\int_0^t s^{2H-1}\left(e^{-\kappa s}+e^{-\kappa(2t-s)}\right)ds$. *Moreover, if* $\kappa<0$, *then* $\operatorname{Var} X_t \sim \frac{H\Gamma(2H)}{(-\kappa)^{2H}}e^{-2\kappa t}$ *as* $t\to\infty$. *If* $\kappa>0$, *then* $\operatorname{Var} X_t \to \frac{H\Gamma(2H)}{\kappa^{2H}}$ *as* $t\to\infty$. *When* $\kappa=0$, $\operatorname{Var} X_t = t^{2H}$ *for* $t\geq 0$.

(iv) For $\kappa<0$, $e^{\kappa t}X_t \to X_0 - \kappa\int_0^\infty e^{\theta s}B_s^H ds \simeq \mathcal{N}\left(X_0, \frac{H\Gamma(2H)}{(-\kappa)^{2H}}\right)$ *a.s. as* $t\to\infty$. *For* $\kappa>0$, $\frac{1}{T}\int_0^T X_t^2 dt \to \frac{H\Gamma(2H)}{\kappa^{2H}}$ *a.s. as* $T\to\infty$.

(v) Let $\kappa>0$. *Then, for any* $p\geq 1$, *there exist positive constants* c_p *and* C_p *such that* $\mathbb{E}\,|X_t|^p \leq c_p$ *for* $t\geq 0$ *and* $\mathbb{E}\,|X_t-X_s|^p \leq C_p|t-s|^{pH}$ *for* $|t-s|\leq 1$. *Moreover, for any* $s,t\in[0,T]$, *we have* $\mathbb{E}X_tX_s \leq C|t-s|^{2H-2}$.

(vi) Let $H\in(0,1)$ *and* $t\geq s\geq 0$. *Then, the covariance function of the fOUp is*

$$\begin{aligned}\operatorname{Cov}(X_t,X_s) &= \frac{H\sigma^2}{2}\Big(-e^{-\kappa t+\kappa s}\int_0^{t-s}e^{\kappa z}z^{2H-1}dz + e^{\kappa t-\kappa s}\int_{t-s}^{t}e^{-\kappa z}z^{2H-1}dz \\ &\quad -e^{-\kappa t-\kappa s}\int_s^t e^{\kappa z}z^{2H-1}dz + e^{-\kappa t+\kappa s}\int_0^s e^{-\kappa z}z^{2H-1}dz \\ &\quad +2e^{-\kappa t-\kappa s}\int_0^t e^{\kappa z}z^{2H-1}dz\Big).\end{aligned}$$

(vii) For $H\neq\frac{1}{2}$, $N=1,2\ldots$ *and* $s\geq 0$, *the asymptotic relation for the covariance function is provided by*

$$\begin{aligned}\operatorname{Cov}(X_s,X_{s+t}) &= \frac{\sigma^2}{2}\sum_{n=1}^{N}(-\theta)^{-2n}\left(\prod_{k=0}^{2n-1}(2H-k)\right)\left(t^{2H-2n}-e^{\theta s}(s+t)^{2H-2n}\right) \\ &\quad +O\left(t^{2H-2N-2}\right), \quad \text{as } t\to\infty\,.\end{aligned}$$

Now, we consider the stationary fOUp. When $\kappa>0$, under the assumption of the initial condition

$$X_0 = \mu + \sigma\int_{-\infty}^{0} e^{\kappa s}d\widetilde{B}_s^H \ \text{ with } \widetilde{B}_s^H = \begin{cases} B_s^H & \text{for } s\geq 0 \\ W_{|s|}^H & \text{for } s\leq 0\end{cases}, \tag{7.2.8}$$

where $W_{|s|}^H$ is another fBm that is independent of B_s^H, and $\widetilde{B}_s^H$ is a two-sided fBm. In this situation, we obtain

$$Y_t = \mu + \sigma\int_{-\infty}^{t} e^{-\kappa(t-s)}dB_s^H = X_t - \left(1-e^{-\kappa t}\right)\mu - X_0e^{-\kappa t} + e^{-\kappa t}Y_0\,. \tag{7.2.9}$$

A standard but tedious calculation yields

$$\begin{aligned}\operatorname{Cov}(Y_t,Y_s) &= \frac{H\sigma^2}{2}\Bigg(\frac{\Gamma(2H)}{\kappa^{2H}}e^{-\kappa(t-s)} - - - e^{-\kappa(t-s)}\int_0^{t-s}e^{\kappa z}z^{2H-1}dz \\ &\quad +e^{\kappa(t-s)}\int_{t-s}^{+\infty}e^{-\kappa z}z^{2H-1}dz\Bigg),\end{aligned} \tag{7.2.10}$$

which implies the stationarity of the process Y_t.

Let $\tau=t-s$. To show the ergodicity of the Gaussian stationary process Y_t, it suffices to show that its covariance function $\operatorname{Cov}(Y_t,Y_s)$ vanishes as τ tends to infinity. According to (7.2.10), the change of variables, and integration by parts, we have

$$\operatorname{Cov}(Y_t,Y_s) = \frac{H\sigma^2}{2\kappa^{2H}}\left(\Gamma(2H)e^{-\kappa\tau} - e^{-\kappa\tau}\int_0^{\kappa\tau}e^{y}y^{2H-1}dy + e^{\kappa\tau}\int_{\kappa\tau}^{+\infty}e^{-y}y^{2H-1}dy\right)$$

$$\leq \frac{H(2H-1)\sigma^2}{2\kappa^{2H}}\left(e^{\frac{-\kappa\tau}{2}} + \left(\frac{\kappa\tau}{2}\right)^{2H-2} + (\kappa\tau)^{2H-2}\right) + O\left(e^{-\kappa\tau}\right)$$

which converges to zero as $\tau \to \infty$ and implies the ergodicity of the Gaussian stationary process Y_t.

In fact, when $\kappa > 0$, Cheridito et al. (2003) have shown that when $H \neq \frac{1}{2}$, Y_t in Model (7.2.9) is an ergodic stationary Gaussian process where the autocorrelation function is

$$\text{Cov}\left(Y_t, Y_{t+s}\right) = \frac{1}{2}\sigma^2 \sum_{n=1}^{N} \kappa^{-2n} \left(\prod_{k=0}^{2n-1} (2H-k)\right) s^{2H-2n} + O\left(s^{2H-2N-2}\right), \tag{7.2.11}$$

for fixed $t \in \mathbb{R}$, $s \to \infty$, and $N = 1, 2, \ldots$.

From Equation (7.2.11), we can see that the decay of the covariance function of a stationary fOUp when $H \neq \frac{1}{2}$ is like that of a power function. Equation (7.2.11) also shows that for $H \in \left(0, \frac{1}{2}\right) \cup \left(\frac{1}{2}, 1\right]$,, the decay of the stationary fOUp is very similar to the decay of the fGn defined by (7.2.4). In particular, $(Y_t)_{t\in\mathbb{R}}$ is ergodic, and for $H \in \left(\frac{1}{2}, 1\right]$, it exhibits long-range dependence.

Since a stationary fOUp has a mean of zero, to construct a good candidate model for volatility and interest rates, a natural extension of the fOUp is the fVm, which adds the long term mean to the fOUp.

$$dX_t = \kappa\left(\mu - X_t\right)dt + \sigma dB_t^H, \tag{7.2.12}$$

where $\sigma \in \mathbb{R}^+$, $\kappa \in \mathbb{R}$, $\mu \in \mathbb{R}$, the initial condition is set at X_0, and B_t^H is an fBm with Hurst parameter $H \in (0, 1)$.

The key difference between Model (7.2.12) and Model (7.2.5) is that μ is assumed to be zero and known in Model (7.2.5), while μ is unknown in Model (7.2.12). The fVm inherits all the important properties of the fOUp reviewed in Proposition 7.2.3.

7.3 Fractional Stochastic Volatility Models

Model (7.2.12) has been used to model realized volatility (RV) in the literature (e.g., Gatheral et al. (2018); Wang et al. (2021)). However, when the volatility is latent, Model (7.2.12) alone cannot be a complete model and an observation equation is needed. This modelling strategy leads to a class of fractional stochastic volatility models.

This class of fractional stochastic volatility models extends the class of standard stochastic volatility models driven by standard Brownian motion in the volatility equation. Before discussing the fractional stochastic volatility models, we first review some famous stochastic volatility models driven by Brownian motion.

A general representation of the continuous-time stochastic volatility model driven by standard Brownian motion may be written as

$$\begin{cases} dS_t/S_t = \mu dt + S_t^{\gamma} f(V_t)[\sqrt{1-\rho^2}dW_t + \rho dZ_t], \\ dV_t/V_t = \beta(V_t)dt + g(V_t)dZ_t, \end{cases} \tag{7.3.1}$$

with independent standard Brownian motions W_t and Z_t. The probability space is denoted as before by $(\Omega, \mathcal{F}, \mathbb{P})$. Here, $S_t > 0$ denotes the price of the (traded) asset and $V_t > 0$ is the (non-traded) stochastic local return variance. Model (7.3.1) allows for level (also known as scale) dependence ($\gamma = 0$) and correlation between returns and variance ($\rho = 0$). As shown in Table 7.1, Model (7.3.1) is the generalization of the most of commonly used models (without the jump component) in research as well as in practice.

One of the key assumptions among all the aforementioned stochastic volatility models is that the drivers are standard Brownian motions. As irregular as the paths might look, Brownian motion

Author(s) (year)	Specification	Remarks
Hull and White (1987)	$f(v) = v$, $\beta(v) = 0$, $g(v) = \sigma$,$\rho = 0, \gamma = 0$.	Instantaneous variance: Geometric Brownian motion. Options priced by mixing.
Wiggins (1987)	$f(v) = e^{\frac{v}{2}}$, $\beta(v) = \frac{\kappa(\theta - v)}{v}$, $g(v) = \sigma$,$\rho = 0, \gamma = 0$.	Instantaneous volatility: The Ornstein-Uhlenbeck in logarithms.
Stein and Stein (1991)	$f(v) = \lvert v \rvert$, $\beta(v) = \frac{\kappa(\theta - v)}{v}$, $g(v) = \frac{\sigma}{v}$,$\rho = 0, \gamma = 0$.	The instantaneous volatility: Reflected Ornstein-Uhlenbeck.
Heston (1993)	$f(v) = \sqrt{v}$, $\beta(v) = \frac{\kappa(\theta - v)}{v}$, $g(v) = \frac{\sigma}{\sqrt{v}}$, $\rho \in [-1, 1], \gamma = 0$.	Instantaneous volatility: CIR process. First model with correlation. Options priced by the Fourier inversion.
Romano and Touzi (1997)	f, β and g free, $\rho \in [-1, 1], \gamma = 0$.	Extension of mixing to correlation
Schöbel and Zhu (1999)	$f(v) = \lvert v \rvert$, $\beta(v) = \frac{\kappa(\theta - v)}{v}$, $g(v) = \frac{\sigma}{v}$, $\rho \in [-1, 1], \gamma = 0$.	Stein and Stein model with correlation. Options priced by the Fourier inversion.
Hagan et al. (2002)	$f(v) = v$, $\beta(v) = 0$, $g(v) = \sigma$, $\rho \in [-1, 1], \gamma \in [-1, 1]$.	Level dependence in returns. Options priced by the perturbation technique.

Table 7.1: Specification of stochastic volatility models for Model (7.3.1)

has independent increments. As a result, the ACF of V_t decays at an exponential rate. To generate the power function for the ACF that is not absolutely summable, Comte and Renault (1998) proposed the first fractional stochastic volatility model in continuous-time

$$\begin{cases} dS_t = rS_t dt + \sigma_t S_t dW_t\,, \\ d\ln(\sigma_t) = \kappa\left(\mu - \ln(\sigma_t)\right)dt + \sigma dB_t^H\,, \end{cases} \tag{7.3.2}$$

where B_t^H is an fBm with the Hurst parameter H. To ensure the ACF of $\ln\left(\sigma_t^2\right)$ is not absolutely summable, Comte and Renault (1998) assumed $H \in (\frac{1}{2}, 1)$.

Comte and Renault (1998) considered the option pricing problem in the long memory volatility environment of Model (7.3.2). Since the closed-form solution for option pricing does not exist, Comte and Renault (1998) provided discrete approximation to fVm Model (7.3.2) and computed option prices based on Monte-Carlo simulation. Comte et al. (2012) investigated the option pricing model by introducing a long-memory extension of the Heston model via the fractional integration of a usual square root volatility process, which has the power decay feature in the large-time limit for the autocovariance function of the volatility process. Both models in Comte and Renault (1998) and Comte et al. (2012) assume that the return process is independent of the volatility process. Due to the complex structures of long memory stochastic processes, Comte and Renault (1998); Comte et al. (2012) cannot derive the analytical formulae for pricing standard options but introduce some discretization schemes and price options using Monte-Carlo simulations.

Moreover, Chronopoulou and Viens (2012a,b) introduced the following fractional stochastic models:

$$\begin{cases} \mathrm{d}S_t/S_t = r\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t\,, \\ \mathrm{d}X_t = \alpha(m - X_t)\,\mathrm{d}t + \beta dB_t^H\,, \end{cases} \tag{7.3.3}$$

where $\sigma(X_t) = \sqrt{X_t}$, $\sigma(X_t) = |X_t|$, $\sigma(X_t) = e^{X_t}$, and $H > \frac{1}{2}$. In the long memory volatility framework, Chronopoulou and Viens (2012a) used an interacting particle stochastic filtering algorithm to estimate the fractional stochastic volatility model of Comte and Renault (1998) and constructed a multinomial recombining tree to price options. Chronopoulou and Viens (2012b) implemented a particle filtering algorithm to estimate the fractional stochastic volatility model of Comte and Renault (1998) and constructed a multinomial recombining tree to price options.

Recent empirical studies have documented the roughness of volatility both in realized volatility and implied volatility (e.g., Bayer et al. (2016); Gatheral et al. (2018); Livieri et al. (2018); El Euch and Rosenbaum (2019); Bennedsen et al. (2021)). More precisely, Gatheral et al. (2018) calibrated the parameters in the fVm using the implied volatility surface of options based on the S&P 500, showing that the Hurst parameter of the volatility is close to 0.1. This result indicates extremely rough paths for the volatility process; these paths are much more irregular than those of the standard stochastic volatility models driven by Brownian motion. Livieri et al. (2018) found that at-the-money short term volatility from the S&P 500 options is also rough. More empirical studies confirmed that the roughness of the log-volatility for thousands of stocks was investigated by Bennedsen et al. (2021). Therefore, both the realized volatility and the option-implied volatility have recently been shown to be rough. Thus, modeling rough volatility is becoming increasingly popular and has important applications in finance because rough volatility models can describe the volatility skew, which is defined as the derivative of the implied volatility surface under the Black—Scholes—Merton model with respect to the log-strike price evaluated at-the-money.

Model (7.2.12) can describe both the mean reverting property and the roughness of the volatility. Consequently, the fVm has become the usual candidate for capturing some phenomena of the volatility of financial assets (e.g., Comte and Renault (1998); Aït-Sahalia and Mancini (2008); Comte et al. (2012); Bayer et al. (2016); Gatheral et al. (2018); Bennedsen et al. (2021); Bolko et al. (2022)). The fVm is an extension of the Vasicek process with the fBm driving term. In finance, it has been used as a one-factor short-term interest rate model (e.g., Fink et al. (2013)) or financial asset volatility model (e.g., Comte and Renault (1998)) for more than two decades.

In Model (7.2.12), $\kappa(\mu - X_t)$ is the drift function and contains two unknown parameters, μ and κ. Parameter κ determines the persistence in X_t. Depending on the sign of κ, the model can capture stationary, explosive, and null recurrent behavior. The fVm has been used to describe the dynamics of volatility in a large set. The development of the application for the fVm naturally led to statistical inference for this model. Consequently, estimating the drift parameter in the fVm has been of great interest in the past decade, and it is a challenging theoretical problem. In the following two sections, we consider parameter estimation for the fVm based on continuous observations and discrete-time observations individually.

7.4 Estimation Methods based on Continuous-Time Observations

Despite the many applications of the fVm in practice, the statistic inference for the fVm has received little attention. For the special case of the fVm, the so-called fOUp provided by Model (7.2.5), we try to summarize the techniques proposed to estimate κ during the past two decades based on the continuous record of observations, which are listed as follows:

(i) The maximum likelihood estimator (MLE) was originally proposed by Kleptsyna and Le Breton (2002) for $\kappa > 0$ with $H \in (\frac{1}{2}, 1)$, in which the estimator of the drift parameter in the fOUp was constructed based on the Girsanov transforms for an fBm, and the strong consistency of the MLE was also established. Tudor and Viens (2007) addressed the problem of estimating the drift parameter in nonlinear stochastic differential equations driven by the fBm with a Hurst parameter range of $H \in (0, 1)$. The strong consistency of the MLE was investigated, but the asymptotic distribution was not investigated. Then, the asymptotic law of MLE was proposed by Tanaka (2013) with $H \in (\frac{1}{2}, 1)$. In fact, the asymptotic law is also valid for $H \in (0, \frac{1}{2})$. For $\kappa = 0$, Kleptsyna and Le Breton (2002) investigated the bias of the MLE. For $\kappa \leq 0$, the asymptotic theory was investigated by Tanaka (2015) with a Hurst parameter range of $H \in (0, 1)$. In particular, the MLE has three celebrated features that are clearly attractive: consistency, asymptotic normality, and the absence of stochastic integration with respect to the fBm. However, the MLE depends on the properties of the deterministic fractional operators (determined by the Hurst parameter) related to the fBm and relies on its ability to compute stochastic integrals with respect to fBm. Moreover, approximating pathwise integrals with respect to the fBm, if they exist, is challenging. Therefore, an actual implementation of the MLE is not easily computable.

(ii) The least squares estimator (LSE) was proposed by Hu and Nualart (2010) for $\kappa > 0$ with $H \in (\frac{1}{2}, 1)$ and by Hu et al. (2019) with $H \in (0, \frac{1}{2})$. For $\kappa < 0$, the LSE was investigated by Belfadli et al. (2011) in the case of $H \in (0, \frac{1}{2})$ and by El Machkouri et al. (2016) for a Hurst parameter range of $H \in (0, 1)$.

(iii) The method of moment estimator (MME) for $\kappa > 0$ was proposed by Hu and Nualart (2010) with $H \in (\frac{1}{2}, 1)$ and by Hu et al. (2019) with $H \in (0, \frac{1}{2})$ for practical purposes. It is worth emphasizing that the MLE involves deterministic fractional operators and that the LSE relies on an unobservable Skorohod integral. Therefore, an actual implementation of the MLE or the LSE is problematic. For the sake of practice, Hu and Nualart (2010); Hu et al. (2019) proposed the MME and studied asymptotical properties for the MME.

(iv) The minimum contrast estimator (MCE) was developed by Bishwal (2011) for $\kappa > 0$ with $H \in (\frac{1}{2}, 1)$ and by Tanaka (2013) for a Hurst parameter range of $H \in (0, 1)$. Note that the MCE does not involve stochastic integrals, unlike the MLE. Bishwal (2011) studied the accuracy of normal approximation for the MCE based on uniform equally spaced sampling of the fOUp. When $\kappa = 0$, the MCE was also considered by Tanaka (2013).

Prakasa Rao (2010) proposed some alternative methods and the asymptotic theory for the fOUp. A survey of parametric estimation and inference procedures for statistical models driven by the fBm has been provided in a recent monograph Kubilius et al. (2017).

In most of the empirically relevant cases, parameter μ in the drift function of Model (7.2.12) is unknown. Thus, it is important to estimate both κ and μ in the fVm. When a continuous record of observations is available for X_t with $t \in [0, T]$, there are three general methods of parameter estimation: the MLE, LSE and MME. For a general Hurst parameter $H \in (0, 1)$, the asymptotic theory of the MLE and LSE is provided for $\kappa \in \mathbb{R}$ while the asymptotic theory of the MME is considered for $\kappa > 0$. The three estimators of κ and μ are reviewed and their asymptotic distributions are developed under the scheme of $T \to \infty$.[5]

7.4.1 MLE

The MLE perhaps is the most popular method for estimating a parametric model because it has many desirable properties: sufficiency (complete information about the parameter of interest is contained in its MLE estimator), consistency (the true parameter value that generated the data is recovered asymptotically , i.e., for data of sufficiently large samples), efficiency (lowest-possible variance of parameter estimates achieved asymptotically), and parameterization invariance (the same MLE solution is obtained independent of the parametrization used).

For illustration purposes, we extend Model (7.2.12) to be a slightly more general form as follows:

$$dX_t = (\alpha - \kappa X_t)\, dt + \sigma dB_t^H \,. \tag{7.4.1}$$

From Model (7.4.1), even when $\kappa = 0$, the drift term does not vanish and is αdt. This alternative specification for the drift term was used in Chan et al. (1992); Yu and Phillips (2001). When α in Model (7.4.1) is known (without loss of generality, it is assumed to be zero), Model (7.4.1) terms to the fOUp. A unique path-wise solution to the stochastic differential equation in Model (7.4.1) is

$$X_t = e^{-\kappa t} X_0 + \frac{\alpha}{\kappa}\left(1 - e^{-\kappa t}\right) + \sigma \int_0^t e^{-\kappa(t-s)} dB_s^H, \tag{7.4.2}$$

where the stochastic integral, $\int_0^t e^{-\kappa(t-s)} dB_s^H$, is the path-wise Riemann-Stieltjes integral, and the solution is unique (Proposition A.1 in Cheridito et al. (2003)).

When $H \in (\frac{1}{2}, 1)$, the MLEs of κ and α have been investigated by Lohvinenko and Ralchenko (2017) with $\kappa > 0$ and Lohvinenko and Ralchenko (2019) with $\kappa < 0$. Consequently, we aim to develop the asymptotic distributions for the MLE of κ and α under the following scenarios: (i) $\kappa > 0$ and $H \in (0, \frac{1}{2}]$, (ii) $\kappa = 0$ and $H \in (0, 1)$, and (iii) $\kappa < 0$ and $H \in (0, 1)$. Therefore, together with Lohvinenko and Ralchenko (2017, 2019), a complete coverage of asymptotic theory for all possible cases is provided for the MLE with κ and α. Although the assumption of a continuous-time record is practically too strong, it allows us to obtain the MLE in a closed form. Moreover, the results obtained here will serve as the benchmark for those based on discrete-time data.

Following Kleptsyna et al. (2000), by applying the Girsanov theorem for the fBm developed in Norros et al. (1999), the continuous-record log-likelihood function for Model (7.4.1) can be expressed as follows:

$$\ell(\kappa, \alpha) = \int_0^T Q_H(t) dM_t^H + \frac{1}{2}\int_0^T \left(Q_H(t)\right)^2 d\omega_t^H,$$

where

$$Q_H\,(t) = \frac{1}{\sigma}\frac{d}{d\omega_t^H}\int_0^t k_H\,(t,s)\,(\alpha - \kappa X_s)\, ds\,, \tag{7.4.3}$$

[5] When a continuous record of observations is available, the unknown parameters of both H and σ can be recovered without estimation errors.

$$k_H(t,s) = \frac{1}{k_H}\left(s(t-s)\right)^{\frac{1}{2}-H}\ ,\ k_H = 2H\Gamma\left(\frac{3}{2}-H\right)\Gamma\left(H+\frac{1}{2}\right)\ , \tag{7.4.4}$$

$$\omega_t^H = \frac{1}{\lambda_H} t^{2-2H}\,, \tag{7.4.5}$$

$$\lambda_H = \frac{2H\Gamma(3-2H)\Gamma\left(H+\frac{1}{2}\right)}{\Gamma\left(\frac{3}{2}-H\right)}\,, \tag{7.4.6}$$

$$M_t^H = \int_0^t k_H(t,s)dB_s^H\,. \tag{7.4.7}$$

Taking the derivatives of the log-likelihood function with respect to κ and α and setting them to zero, Lohvinenko and Ralchenko (2017) provided the following MLE estimators for α and κ:

$$\tilde{\alpha}_T = \frac{S_T \int_0^T P_H^2(t)\,d\omega_t^H - \int_0^T P_H(t)\,dS_t \int_0^T P_H(t)\,d\omega_t^H}{\omega_T^H \int_0^T P_H^2(t)\,d\omega_t^H - \left(\int_0^T P_H(t)\,d\omega_t^H\right)^2}\sigma\,, \tag{7.4.8}$$

$$\tilde{\kappa}_T = \frac{S_T \int_0^T P_H(t)\,d\omega_t^H - \omega_T^H \int_0^T P_H(t)\,dS_t}{\omega_T^H \int_0^T P_H^2(t)\,d\omega_t^H - \left(\int_0^T P_H(t)\,d\omega_t^H\right)^2}\,, \tag{7.4.9}$$

where

$$S_t = \frac{1}{\sigma}\int_0^t k_H(t,s)\,dX_s\,, \tag{7.4.10}$$

$$P_H(t) = \frac{1}{\sigma}\frac{d}{d\omega_t^H}\int_0^t k_H(t,s)\,X_s ds\,, \tag{7.4.11}$$

Combining Model (7.4.1) with Equations (7.4.4) and (7.4.11), we have

$$P_H(t) = \frac{1}{\sigma}\frac{\alpha}{\kappa} + \frac{1}{\sigma}\left(X_0 - \frac{\alpha}{\kappa}\right)V_H(t) + \tilde{P}_H(t)\ , \tag{7.4.12}$$

where

$$V_H(t) = \frac{d}{d\omega_t^H}\int_0^t k_H(t,s)\,e^{-\kappa s}ds\,, \tag{7.4.13}$$

$$\tilde{P}_H(t) = \frac{d}{d\omega_t^H}\int_0^t k_H(t,s)\,U_s ds\,, \tag{7.4.14}$$

$$U_t = \int_0^t e^{-\kappa(t-s)}dB_s^H\,. \tag{7.4.15}$$

Using the idea from Kleptsyna and Le Breton (2002), Lohvinenko and Ralchenko (2017) proposed the following results:

$$Q_H(t) = \frac{\alpha}{\sigma} - \kappa P_H(t)\ , \tag{7.4.16}$$

$$S_t = \int_0^t Q_H(s)\,d\omega_s^H + M_t^H = \frac{\alpha}{\sigma}\omega_t^H - \kappa\int_0^t P_H(s)\,d\omega_s^H + M_t^H\,, \tag{7.4.17}$$

$$dS_t = \frac{\alpha}{\sigma}d\omega_t^H - \kappa P_H(t)\,d\omega_t^H + dM_t^H\,. \tag{7.4.18}$$

The process M_t^H, called the fundamental martingale, is a Gaussian martingale where the variance function is ω_t^H. Moreover, the natural filtration of the martingale M^H coincides with the

	$1/2 < H < 1$	$H = 1/2$	$0 < H < 1/2$
$\kappa < 0$	$T^{1-H}(\tilde{\alpha}_T - \alpha) \Rightarrow \mathcal{N}(0, \lambda_H \sigma^2)$ $\frac{e^{-\kappa T}}{2\kappa}(\tilde{\kappa}_T - \kappa) \Rightarrow \frac{X\sqrt{\sin(\pi H)}}{Y}$	$\sqrt{T}(\tilde{\alpha}_T - \alpha) \Rightarrow \mathcal{N}(0, \sigma^2)$ $\frac{e^{-\kappa T}}{2\kappa}(\tilde{\kappa}_T - \kappa) \Rightarrow \frac{\eta_\infty}{X_0 - \frac{\alpha}{\kappa} + \xi_\infty}$	$T^{1-H}(\tilde{\alpha}_T - \alpha) \Rightarrow \mathcal{N}(0, \lambda_H \sigma^2)$ $\frac{e^{-\kappa T}}{2\kappa}(\tilde{\kappa}_T - \kappa) \Rightarrow \frac{X\sqrt{\sin(\pi H)}}{Y}$
$\kappa = 0$	$T^{1-H}(\tilde{\alpha}_T - \alpha) \Rightarrow \mathcal{N}(0, \sigma^2 \rho_H)$ $T^{2-H}(\tilde{\kappa}_T - \kappa) \Rightarrow \mathcal{N}\left(0, \frac{\sigma^2}{\alpha^2}\phi_H\right)$	$\sqrt{T}(\tilde{\alpha}_T - \alpha) \Rightarrow \mathcal{N}(0, 4\sigma^2)$ $T^{3/2}(\tilde{\kappa}_T - \kappa) \Rightarrow \mathcal{N}\left(0, \frac{12\sigma^2}{\alpha^2}\right)$	$T^{1-H}(\tilde{\alpha}_T - \alpha) \Rightarrow \mathcal{N}(0, \sigma^2 \rho_H)$ $T^{2-H}(\tilde{\kappa}_T - \kappa) \Rightarrow \mathcal{N}\left(0, \frac{\sigma^2}{\alpha^2}\phi_H\right)$
$\kappa > 0$	$\sqrt{T}(\tilde{\alpha}_T - \alpha) \Rightarrow \mathcal{N}(0, \lambda_H \sigma^2)$ $\sqrt{T}(\tilde{\kappa}_T - \kappa) \Rightarrow \mathcal{N}(0, 2\kappa)$	$\sqrt{T}(\tilde{\alpha}_T - \alpha) \Rightarrow \mathcal{N}\left(0, \sigma^2 + \frac{2\alpha^2}{\kappa}\right)$ $\sqrt{T}(\tilde{\kappa}_T - \kappa) \Rightarrow \mathcal{N}(0, 2\kappa)$	$T^{1-H}(\tilde{\alpha}_T - \alpha) \Rightarrow \mathcal{N}\left(0, \frac{2\alpha^2}{\kappa}\right)$ $\sqrt{T}(\tilde{\kappa}_T - \kappa) \Rightarrow \mathcal{N}(0, 2\kappa)$

Table 7.2: The asymptotic laws of $\tilde{\alpha}_T$ and $\tilde{\kappa}_T$ for different ranges of H and κ.

natural filtration of the fBm. Based on Models (7.4.17) and (7.4.18), the MLE of α and κ can be represented as

$$\tilde{\alpha}_T = \alpha + \frac{M_T^H \int_0^T P_H^2(t)\, d\omega_t^H - \int_0^T P_H(t)\, dM_t^H \int_0^T P_H(t)\, d\omega_t^H}{\omega_T^H \int_0^T P_H^2(t)\, d\omega_t^H - \left(\int_0^T P_H(t)\, d\omega_t^H\right)^2}\sigma\,, \tag{7.4.19}$$

$$\tilde{\kappa}_T = \kappa + \frac{M_T^H \int_0^T P_H(t)\, d\omega_t^H - \omega_T^H \int_0^T P_H(t)\, dM_t^H}{\omega_T^H \int_0^T P_H^2(t)\, d\omega_t^H - \left(\int_0^T P_H(t)\, d\omega_t^H\right)^2}\,, \tag{7.4.20}$$

When a continuous record of observations of X_t is available, Lohvinenko and Ralchenko (2017) studied the consistency and the asymptotic normality of the MLE as defined by Equations (7.4.8) and (7.4.9) when $H > \frac{1}{2}$ and $\kappa > 0$. From Tanaka et al. (2020), we can obtain the asymptotic theory for the MLE of α and κ for all other cases, including $H < \frac{1}{2}$ and $\kappa > 0$, $H \in (0,1)$ and $\kappa = 0$, and $H \in (0,1)$ and $\kappa < 0$. Table 7.2 summarizes the asymptotic distributions of $\tilde{\alpha}_T$ and $\tilde{\kappa}_T$ for different ranges of H and κ, where $\lambda_H = \frac{2H\Gamma(3-2H)\Gamma\left(H+\frac{1}{2}\right)}{\Gamma\left(\frac{3}{2}-H\right)}$, $\rho_H = \lambda_H(3-2H)^2$, $\phi_H = \frac{32H(1-H)(2-H)\Gamma(3-2H)\Gamma(H+\frac{1}{2})}{\Gamma(\frac{3}{2}-H)}$, ξ_∞, η_∞ are two independent $\mathcal{N}(0, -\sigma^2/(2\kappa))$ random variables, and X and Y are two independent $\mathcal{N}(0,1)$ random variables. Moreover, we assume $X_0 = \alpha/\kappa$ for $\kappa < 0$ and $H \in (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$.

7.4.2 LSE

The LSE, unlike the MLE, which requires no or minimal distributional assumptions, is useful for obtaining a descriptive measure for the purpose of summarizing observed data. However, there is no basis for testing hypotheses or constructing confidence intervals. Based on Model (7.2.12) and motivated by the work of Hu and Nualart (2010); Belfadli et al. (2011); El Machkouri et al. (2016), we denote the LSE of κ and μ to be the minimizers of the following (formal) quadratic Function:

$$L(\kappa, \mu) = \int_0^T \left(\dot{X}_t - \kappa(\mu - X_t)\right)^2 dt\,, \tag{7.4.21}$$

where $\dot{X}_t$ denotes the differentiation of X_t with respect to t, although $\int_0^T \dot{X}_t^2 dt$ does not exist. Consequently, we obtain the following analytical expressions for the LSE of κ and μ, denoted by $\hat{\kappa}_{LS}$ and $\hat{\mu}_{LS}$, respectively.

$$\hat{\kappa}_{LS} = \frac{(X_T - X_0)\int_0^T X_t dt - T\int_0^T X_t dX_t}{T\int_0^T X_t^2 dt - \left(\int_0^T X_t dt\right)^2}\,, \tag{7.4.22}$$

$$\hat{\mu}_{LS} = \frac{(X_T - X_0)\int_0^T X_t^2 dt - \int_0^T X_t dX_t \int_0^T X_t dt}{(X_T - X_0)\int_0^T X_t dt - T\int_0^T X_t dX_t}\,. \tag{7.4.23}$$

When $H = \frac{1}{2}$, it is well known that we can interpret the stochastic integral $\int_0^T X_t dX_t$ as an Itô integral. When $H \in (\frac{1}{2}, 1)$, X_t is no longer a semimartingale. In this case, for $\hat{\kappa}_{LS}$ and $\hat{\mu}_{LS}$ to consistently estimate κ and μ, we have to interpret the stochastic integral $\int_0^T X_t dX_t$ carefully. In fact, we interpret it differently when the sign of κ is different. If $\kappa > 0$, we interpret it as an Itô-Skorohod integral; if $\kappa < 0$, we interpret it as a Young integral; and if $\kappa = 0$, we can interpret it as either an Itô-Skorohod integral or a Young integral. In the case of a Brownian motion-driven or a Lévy process-driven Vasicek model, it is known that the asymptotic theory for κ depends on the sign of κ (see, Wang and Yu (2016)).

In the case of the fVm, here, we review the idea that the asymptotic theory for κ continues to depend on the sign of κ. The asymptotic distributions of $\hat{\kappa}_{LS}$ are different across these three cases. From Xiao and Yu (2019a,b), we can obtain the following asymptotic distributions of $\hat{\kappa}_{LS}$ and $\hat{\mu}_{LS}$, which are provided in Table 7.3. Here, $\overline{B}_u^H = B_u^H - \int_0^1 B_t^H dt$, $\delta_{LS}^2 = (4H-1) + \frac{2\Gamma(2-4H)\Gamma(4H)}{\Gamma(2H)\Gamma(1-2H)}$, $C_H = (4H-1)\left(1 + \frac{\Gamma(3-4H)\Gamma(4H-1)}{\Gamma(2H)\Gamma(2-2H)}\right)$, ν and ω are two independent standard normal variables, and $R(H)$ is the Rosenblatt random variable whose characteristic function is expressed as

$$c(s) = \exp\left(\frac{1}{2}\sum_{k=2}^{\infty}\left(2\sqrt{-1}s\sigma(H)\right)^k \frac{a_k}{k}\right), \tag{7.4.24}$$

with $\sigma(H) = \sqrt{H(H-\frac{1}{2})}$ and

$$a_k = \int_0^1 \cdots \int_0^1 |x_1 - x_2|^{H-1} \cdots |x_{k-1} - x_k|^{H-1} |x_k - x_1|^{H-1} dx_1 \cdots dx_k.$$

From Table 7.3, we can see that the LSEs of μ and κ are strongly consistent regardless of the sign of the persistence parameter κ. Moreover, the asymptotic distribution of the LSE of μ is asymptotically normal regardless of the sign of κ, while the asymptotic distribution of the LSE of κ critically depends on the sign of κ. In particular, when $\kappa > 0$ and $H \in (0, 3/4)$, we can see that the asymptotic distribution of the LSE of κ is a normal distribution with a rate of convergence of $\sqrt{T}$. When $\kappa > 0$ and $H = 3/4$, we obtain that the asymptotic distribution of the LSE of κ is also a normal distribution with a rate of convergence of $\sqrt{T/\log(T)}$. However, a non-central limit theorem for the LSE of κ is established for $H \in (3/4, 1)$. In this situation, we obtain the asymptotic law as a Rosenblatt random variable. When $\kappa < 0$, we can see that the limiting distribution is a Cauchy type with a rate of convergence of $e^{-\kappa T}$. If μ equals the initial condition, it becomes the standard Cauchy distribution. When $\kappa = 0$, the asymptotic distribution is neither normal nor a mixture of normals but a Dickey-Fuller-Phillips type of distribution. The rate of convergence is T.

7.4.3 MME

In this subsection, we consider an alternative estimation technique by exploiting the ergodic property of the fVm when $\kappa > 0$. Borrowing the idea from Xiao and Yu (2019a,b), the asymptotic properties of the MME are compared.

Using the initial condition of (7.2.8), we can show that X_t in Model (7.2.12) is covariance stationary with

$$\lim_{t\to\infty} \mathbb{E}(X_t) = \mu \quad \text{and} \quad \lim_{t\to\infty} Var(X_t) = \sigma^2\kappa^{-2H} H\Gamma(2H). \tag{7.4.25}$$

Moreover, X_t can be identically represented as

$$X_t = \mu + \sigma\int_{-\infty}^{t} e^{-\kappa(t-s)} d\widetilde{B}_s^H. \tag{7.4.26}$$

	$\kappa > 0$	$\kappa = 0$	$\kappa < 0$
$0 < H < \frac{1}{2}$	$\sqrt{T}\left(\hat{\kappa}_{LS} - \kappa\right) \Rightarrow \mathcal{N}\left(0, \kappa\delta_{LS}^2\right)$	$T\hat{\kappa}_{LS} \overset{d}{=} -\frac{\int_0^1 \overline{B}_u^H dB_u^H}{\int_0^1 \left(\overline{B}_u^H\right)^2 du}$	$\frac{e^{-\kappa T}}{2\kappa}\left(\hat{\kappa}_{LS} - \kappa\right) \Rightarrow \frac{\sigma\frac{\sqrt{H\Gamma(2H)}}{\lvert\kappa\rvert^H}\nu}{X_0-\mu+\sigma\frac{\sqrt{H\Gamma(2H)}}{\lvert\kappa\rvert^H}\omega}$
	$T^{1-H}\left(\hat{\mu}_{LS} - \mu\right) \Rightarrow \mathcal{N}\left(0, \frac{\sigma^2}{\kappa^2}\right)$	———————	$T^{1-H}\left(\hat{\mu}_{LS} - \mu\right) \Rightarrow \mathcal{N}\left(0, \frac{\sigma^2}{\kappa^2}\right)$
$\frac{1}{2} \leq H < 3/4$	$\sqrt{T}\left(\hat{\kappa}_{LS} - \kappa\right) \Rightarrow \mathcal{N}\left(0, \kappa C_H\right)$	$T\hat{\kappa}_{LS} \overset{d}{=} -\frac{\int_0^1 \overline{B}_u^H dB_u^H}{\int_0^1 \left(\overline{B}_u^H\right)^2 du}$	$\frac{e^{-\kappa T}}{2\kappa}\left(\hat{\kappa}_{LS} - \kappa\right) \Rightarrow \frac{\sigma\frac{\sqrt{H\Gamma(2H)}}{\lvert\kappa\rvert^H}\nu}{X_0-\mu+\sigma\frac{\sqrt{H\Gamma(2H)}}{\lvert\kappa\rvert^H}\omega}$
	$T^{1-H}\left(\hat{\mu}_{LS} - \mu\right) \Rightarrow \mathcal{N}\left(0, \frac{\sigma^2}{\kappa^2}\right)$	———————	$T^{1-H}\left(\hat{\mu}_{LS} - \mu\right) \Rightarrow \mathcal{N}\left(0, \frac{\sigma^2}{\kappa^2}\right)$
$H = 3/4$	$\frac{\sqrt{T}}{\sqrt{\log(T)}}\left(\hat{\kappa}_{LS} - \kappa\right) \Rightarrow \mathcal{N}\left(0, \frac{4\kappa}{\pi}\right)$	$T\hat{\kappa}_{LS} \overset{d}{=} -\frac{\int_0^1 \overline{B}_u^H dB_u^H}{\int_0^1 \left(\overline{B}_u^H\right)^2 du}$	$\frac{e^{-\kappa T}}{2\kappa}\left(\hat{\kappa}_{LS} - \kappa\right) \Rightarrow \frac{\sigma\frac{\sqrt{H\Gamma(2H)}}{\lvert\kappa\rvert^H}\nu}{X_0-\mu+\sigma\frac{\sqrt{H\Gamma(2H)}}{\lvert\kappa\rvert^H}\omega}$
	$T^{1-H}\left(\hat{\mu}_{LS} - \mu\right) \Rightarrow \mathcal{N}\left(0, \frac{\sigma^2}{\kappa^2}\right)$	———————	$T^{1-H}\left(\hat{\mu}_{LS} - \mu\right) \Rightarrow \mathcal{N}\left(0, \frac{\sigma^2}{\kappa^2}\right)$
$3/4 < H < 1$	$T^{2-2H}\left(\hat{\kappa}_{LS} - \kappa\right) \Rightarrow \frac{-\kappa^{2H-1}}{H\Gamma(2H)}R(H)$	$T\hat{\kappa}_{LS} \overset{d}{=} -\frac{\int_0^1 \overline{B}_u^H dB_u^H}{\int_0^1 \left(\overline{B}_u^H\right)^2 du}$	$\frac{e^{-\kappa T}}{2\kappa}\left(\hat{\kappa}_{LS} - \kappa\right) \xrightarrow{d} \frac{\sigma\frac{\sqrt{H\Gamma(2H)}}{\lvert\kappa\rvert^H}\nu}{X_0-\mu+\sigma\frac{\sqrt{H\Gamma(2H)}}{\lvert\kappa\rvert^H}\omega}$
	$T^{1-H}\left(\hat{\mu}_{LS} - \mu\right) \Rightarrow \mathcal{N}\left(0, \frac{\sigma^2}{\kappa^2}\right)$	———————	$T^{1-H}\left(\hat{\mu}_{LS} - \mu\right) \Rightarrow \mathcal{N}\left(0, \frac{\sigma^2}{\kappa^2}\right)$

Table 7.3: The asymptotic laws of $\hat{\kappa}_T$ and $\hat{\mu}_T$ for different ranges of H and κ.

If $\kappa > 0$, under a general initial condition $X_0 = O_p(1)$, X_t is asymptotically covariance stationary. In this situation, motivated by Hu and Nualart (2010), we can consider alternative estimators of κ and μ (denoted by $\hat{\kappa}_{HN}$ and $\hat{\mu}_{HN}$, respectively). The strong solution for the fVm in Model (7.2.12) is expressed as

$$X_t = \mu + (X_0 - \mu)\exp(-\kappa t) + \sigma \int_{-\infty}^{t} e^{-\kappa(t-s)} dB_s^H . \tag{7.4.27}$$

Moreover, when $\kappa > 0$, we can easily obtain

$$\frac{1}{T}\int_0^T X_t dt \quad \to_{a.s.} \quad \mu \,, \tag{7.4.28}$$

$$\frac{1}{T}\int_0^T X_t^2 dt \quad \to_{a.s.} \quad \sigma^2\kappa^{-2H} H\Gamma(2H) + \mu^2 \,. \tag{7.4.29}$$

According to the ergodic theorem and using Equations (7.4.28) and (7.4.29), the MME of μ and $\kappa > 0$ was introduced by Xiao and Yu (2019a,b)

$$\hat{\mu}_{HN} \quad = \quad \frac{1}{T}\int_0^T X_t dt \,. \tag{7.4.30}$$

$$\hat{\kappa}_{HN} \quad = \quad \left(\frac{T\int_0^T X_t^2 dt - \left(\int_0^T X_t dt\right)^2}{T^2\sigma^2 H\Gamma(2H)} \right)^{-\frac{1}{2H}} . \tag{7.4.31}$$

Compared with the LSE in Equations (7.4.22) and (7.4.23), which involve the stochastic integral $\int_0^T X_t dX_t$, the MME $\hat{\mu}_{HN}$ and $\hat{\kappa}_{HN}$ in Equations (7.4.30) and (7.4.31) do not contain any stochastic integrals with respect to the fBm but only involve quadratic integral functionals. Therefore, they are conceptually easier to understand and numerically easier to compute than the LSE and MLE.

From Xiao and Yu (2019a,b), we can obtain the consistency of $\hat{\kappa}_{HN}$ and $\hat{\mu}_{HN}$, which is shown by the following theorem:

Theorem 7.4.1 *Let $H \in (0,1)$, $X_0/\sqrt{T} = o_{a.s.}(1)$ and $\kappa > 0$ in Model (7.2.12). Then, we have $\hat{\kappa}_{HN} \to_{a.s.} \kappa$ and $\hat{\mu}_{HN} \to_{a.s.} \mu$.*

Using the asymptotic distributions of $\hat{\kappa}_{LS}$ and $\hat{\mu}_{LS}$, the asymptotic distributions of $\hat{\kappa}_{HN}$ and $\hat{\mu}_{HN}$ can be developed as follows:

Theorem 7.4.2 *Let $X_0/\sqrt{T} = o_p(1)$ and $\kappa > 0$ in (7.2.12). Then, we have*

$$T^{1-H}\left(\hat{\mu}_{HN} - \mu\right) \Rightarrow \mathcal{N}\left(0, \frac{\sigma^2}{\kappa^2}\right). \tag{7.4.32}$$

Moreover, let $X_0/\sqrt{T} = o_p(1)$ and $\kappa > 0$ in Model (7.2.12). Then, the following convergence results hold true:

(i) For $H \in (0, 3/4)$, we have

$$\sqrt{T}\left(\hat{\kappa}_{HN} - \kappa\right) \Rightarrow \mathcal{N}\left(0, \kappa\rho_H\right), \tag{7.4.33}$$

where $\rho_H = \frac{4H-1}{4H^2}\left(1 + \frac{\Gamma(3-4H)\Gamma(4H-1)}{\Gamma(2H)\Gamma(2-2H)}\right) = \frac{C_H}{4H^2}$.

(ii) For $H = 3/4$, we have

$$\frac{\sqrt{T}}{\sqrt{\log(T)}}\left(\hat{\kappa}_{HN} - \kappa\right) \Rightarrow \mathcal{N}\left(0, \frac{16\kappa}{9\pi}\right). \tag{7.4.34}$$

(iii) For $H \in (3/4, 1)$, we have

$$T^{2-2H}\left(\hat{\kappa}_{HN} - \kappa\right) \Rightarrow \frac{-\kappa^{2H-1}}{H\Gamma(2H+1)} R(H)\,, \tag{7.4.35}$$

where $R(H)$ is the Rosenblatt random variable defined in (7.4.24).

7.5 Estimation Method based on Discrete-Time Observations

In the previous section, the estimators $\tilde{\alpha}_T$, $\tilde{\kappa}_T$, $\hat{\kappa}_{LS}$, $\hat{\mu}_{LS}$, $\hat{\mu}_{HN}$ and $\hat{\kappa}_{HN}$ were developed based on continuous-time observations. In practice, observations of X_t are available only at discrete-time points, for example, at $n(:= T/\Delta)$ equally spaced points $\{i\Delta\}_{i=0}^{n}$ over time interval $[0, T]$ where Δ is the sampling interval and T denotes the time span. When X_t is annualized and observed daily (weekly or monthly), then $\Delta = 1/252$ ($1/52$ or $1/12$). We now consider Model (7.2.12), where $\kappa \in \mathbb{R}^+$, $\sigma \in \mathbb{R}^+$, $\mu \in \mathbb{R}$, and $H \in (0, 1)$ are constants. Let $\{X_{i\Delta}\}_{i=0}^{n}$ denote the discrete-time observations of X_t. From Equation (7.4.27), the exact discrete-time model of $\{X_{i\Delta}\}_{i=0}^{n}$ is obtained as

$$X_{i\Delta} = e^{-\kappa\Delta} X_{(i-1)\Delta} + \left(1 - e^{-\kappa\Delta}\right)\mu + \varepsilon_{i\Delta}\,, \tag{7.5.1}$$

where $\varepsilon_{i\Delta} = \sigma \int_{(i-1)\Delta}^{i\Delta} e^{-\kappa(i\Delta - s)} dB_s^H$.

We now review the two-stage approach of Wang et al. (2021) for estimating the four parameters in Model (7.2.12) based on discrete-time observations of X_t. In the first stage, following Lang and Roueff (2001) and Barndorff-Nielsen et al. (2013), H is estimated based on the ratio of the squared summations of the second-order differences of X_t obtained at different frequencies. In the second stage, the estimators of the other parameters in Model (7.2.12) are constructed based on a set of moment conditions in which the true value of H is replaced with the estimated H obtained during the first stage. Closed-form expressions are established for all the proposed estimators and are denoted by $\widehat{H}$, $\widehat{\kappa}$, $\widehat{\mu}$, and $\widehat{\sigma}$. We next review a large sample theory for the proposed estimators. In particular, we review two asymptotic schemes: (i) the in-fill scheme under which the sampling interval Δ goes to zero with a fixed time span T and (ii) the double scheme in which $\Delta \to 0$ and $T \to \infty$ simultaneously. Under both schemes, the consistency and asymptotic normality of $\widehat{H}$ and $\widehat{\sigma}$ are introduced for all $H \in (0, 1)$ and regardless of the stationarity property of the model. In addition, an explicit formula is introduced for the asymptotic variance of $\widehat{H}$, which depends only on the value of H. This feature greatly facilitates statistical inference about H. Under the double scheme, the consistency and the asymptotic distributions of $\widehat{\kappa}$ and $\widehat{\mu}$ are introduced. The convergence rate of $\widehat{\mu}$ is a function of H. Both the convergence rate and the asymptotic distribution of $\widehat{\kappa}$ depend crucially on H.

7.5.1 The Two-Stage Approach

To estimate the parameters in Model (7.2.12) based on discrete-time data, it is difficult to apply the MLE to estimate all the parameters simultaneously because the errors $\{\varepsilon_{t\Delta}\}$ in Equation (7.5.1) have a complicated dependent structure when $H \neq \frac{1}{2}$. Following Phillips and Yu (2009), we review the two-stage estimation approach introduced by Wang et al. (2021), which is straightforward to implement.

In the first stage, following Lang and Roueff (2001) and Barndorff-Nielsen et al. (2013), Wang et al. (2021) estimated the Hurst parameter H by using the change-of-frequency (COF) estimator based on the second-order differences of X_t.

$$\widehat{H} = \frac{1}{2}\log_2\left(\frac{\sum_{i=1}^{n-4}\left(X_{(i+4)\Delta} - 2X_{(i+2)\Delta} + X_{i\Delta}\right)^2}{\sum_{i=1}^{n-2}\left(X_{(i+2)\Delta} - 2X_{(i+1)\Delta} + X_{i\Delta}\right)^2}\right)\,, \tag{7.5.2}$$

where $\log_2(\cdot)$ is the base-2 logarithm, $\left\{X_{(i+4)\Delta} - 2X_{(i+2)\Delta} + X_{i\Delta}\right\}_{i=1}^{n-4}$, and $\left\{X_{(i+2)\Delta} - 2X_{(i+1)\Delta} + X_{i\Delta}\right\}_{i=1}^{n-2}$ are second-order differences of $\{X_{i\Delta}\}_{i=1}^{n}$ taken at two different frequencies.

In the second stage, Wang et al. (2021) estimated the other parameters, σ, μ, and κ, in Model (7.2.12) using the following method-of-moments estimators:

$$\widehat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n-2}\left(X_{(i+2)\Delta} - 2X_{(i+1)\Delta} + X_{i\Delta}\right)^2}{n\left(4 - 2^{2\widehat{H}}\right)\Delta^{2\widehat{H}}}}, \tag{7.5.3}$$

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_{i\Delta}, \tag{7.5.4}$$

$$\widehat{\kappa} = \left(\frac{n\sum_{i=1}^{n} X_{i\Delta}^2 - \left(\sum_{i=1}^{n} X_{i\Delta}\right)^2}{n^2\widehat{\sigma}^2\widehat{H}\Gamma\left(2\widehat{H}\right)}\right)^{-1/\left(2\widehat{H}\right)}. \tag{7.5.5}$$

Note that $\widehat{\sigma}$ depends on $\widehat{H}$, which is obtained in the first Stage, and $\widehat{\kappa}$ depends on both $\widehat{\sigma}$ and $\widehat{H}$.

The estimators $\widehat{\mu}$ and $\widehat{\kappa}$ can be regarded as the discrete-time versions of the ergodic-type estimators of κ and μ of Xiao and Yu (2019a,b). However, since Xiao and Yu (2019a,b) assumed that σ^2 and H are known and a continuous record of $\{X_t\}$ is observed, we have to modify their estimators by (i) replacing σ and H with $\widehat{\sigma}$ and $\widehat{H}$ and (ii) replacing the Riemann integrals with the summations. When discrete-time observations of $\{X_t\}$ are available, the MLE, MCE, LSE, and MME have previously been studied by Tudor and Viens (2007); Ludeña (2004); Hu et al. (2019), respectively. A critical difference from these studies is that the estimator reviewed in this section does not assume H is known when estimating κ. The asymptotic distribution of the LSE of κ has been derived in the O-U model (H is assumed to be $\frac{1}{2}$). For example, assuming $\kappa > 0$, Tang and Chen (2009) obtained the long-span and double asymptotic distributions of the MLE of κ. Assuming $\kappa < 0$, Wang and Yu (2016) obtained the double asymptotic distribution of the LSE of κ. We review the asymptotics of κ from the O-U model to the fVm.

7.5.2 Asymptotic Theory

In this subsection, we first review the consistency of $\widehat{H}$ and $\widehat{\sigma}$ as long as $T\Delta \to 0$ and $n = T/\Delta \to \infty$, a condition that is satisfied under either (i) the in-fill asymptotic scheme where $\Delta \to 0$ with a fixed T or (ii) the double asymptotic scheme where $\Delta \to 0$ and $T \to \infty$ simultaneously while T diverges at a lower rate than that of $1/\Delta$.[6] Then, we review the results from Wang et al. (2021). Thus, $T \to \infty$ is a necessary condition for the consistency of $\widehat{\mu}$ and $\widehat{\kappa}$ as defined in Equations (7.5.4) and (7.5.5). We also review the double asymptotic theory from $\widehat{\mu}$ and $\widehat{\kappa}$ using the results from Wang et al. (2021).

Theorem 7.5.1 *Let $\widehat{H}$ and $\widehat{\sigma}$ denote the MME defined in Equations (7.5.2) and (7.5.3) for Model (7.2.12). Then, if $T\Delta \to 0$ and $n = T/\Delta \to \infty$, for all $H \in (0,1)$, we have*
(a) $\widehat{H} \to_p H$, and

$$\sqrt{n}\left(\widehat{H} - H\right) \Rightarrow \mathcal{N}\left(0, \frac{\Sigma_{11} + \Sigma_{22} - 2\Sigma_{12}}{(2\log 2)^2}\right); \tag{7.5.6}$$

[6] The consistency of $\widehat{H}$ only requires $\Delta \to 0$. In other words, even when T diverges faster than $1/\Delta$, violating the condition $T\Delta \to 0$, $\widehat{H}$ is still consistent as long as $\Delta \to 0$.

(b) $\widehat{\sigma} \to_p \sigma$, *and*

$$\frac{\sqrt{n}}{\log(\Delta)}(\widehat{\sigma} - \sigma) \Rightarrow \mathcal{N}\left(0, \frac{\Sigma_{11} + \Sigma_{22} - 2\Sigma_{12}}{(2\log 2)^2}\sigma^2\right), \tag{7.5.7}$$

where

$$\Sigma_{11} = 2 + 2^{2-4H}\sum_{j=1}^{\infty}\left(\rho_{j+2} + 4\rho_{j+1} + 6\rho_j + 4\rho_{|j-1|} + \rho_{|j-2|}\right)^2, \tag{7.5.8}$$

$$\Sigma_{12} = 2^{1-2H}\left(4(\rho_1 + 1)^2 + 2\sum_{j=0}^{\infty}(\rho_{j+2} + 2\rho_{j+1} + \rho_j)^2\right), \tag{7.5.9}$$

$$\Sigma_{22} = 2 + 4\sum_{j=1}^{\infty}\rho_j^2, \tag{7.5.10}$$

with

$$\rho_j = \frac{-|j+2|^{2H} + 4|j+1|^{2H} - 6|j|^{2H} + 4|j-1|^{2H} - |j-2|^{2H}}{2(4 - 2^{2H})}. \tag{7.5.11}$$

Remark 7.5.2 *The asymptotics of $\widehat{H}$ and $\widehat{\sigma}$ hold for all $H \in (0,1)$. It is clear that the asymptotics of $\widehat{H}$ and $\widehat{\sigma}$ can apply to all κ, including $\kappa > 0$, $\kappa = 0$, and $\kappa < 0$.*

Remark 7.5.3 *Let us observe that the asymptotic variance of $\widehat{H}$ only depends on H, while the asymptotic variance of $\widehat{\sigma}$ depends on both H and σ. Neither depends on κ and μ. This feature plays an important role in statistical inference about H and σ because, when T is fixed, we can obtain consistent estimators for both H and σ, while we cannot estimate κ and μ consistently. If we use an alternative estimator for H, which has a rate of convergence $\sqrt{n}$, then we can see that $\sqrt{n}\left(\widehat{H} - H\right)$ and $\sqrt{n}(\widehat{\sigma} - \sigma)/(\sigma\log(\Delta))$ should share the same limiting distribution.*

It is worth mentioning that, for all $H \in (0,1)$ and all κ, including $\kappa > 0$, $\kappa = 0$, and $\kappa < 0$, the asymptotics of $\widehat{H}$ and $\widehat{\sigma}$ always hold. When $H = \frac{1}{2}$, Model (7.2.12) becomes the standard O-U model and enjoys the Markov property, whereas Model (7.2.12) does not have the Markov property once $H \neq \frac{1}{2}$. To test of the hypothesis $H = \frac{1}{2}$, the following Corollary 7.5.4 provides the asymptotic variance of $\sqrt{n}\left(\widehat{H} - \frac{1}{2}\right)$. By substituting $H = \frac{1}{2}$ into the formulae provided in Theorem 7.5.1, we obtain that $\rho_0 = 1$, $\rho_1 = -1/2$, $\rho_j = 0$ for $j \geq 2$, $\Sigma_{11} = 7/2$, $\Sigma_{12} = 3/2$, and $\Sigma_{22} = 3$. Consequently, we can obtain Corollary 7.5.4 directly.

Corollary 7.5.4 *Let $H = \frac{1}{2}$, $T\Delta \to 0$, and $n = T/\Delta \to \infty$. Then, we obtain*

$$\sqrt{n}\left(\widehat{H} - \frac{1}{2}\right) \Rightarrow \mathcal{N}\left(0, \frac{7}{8(\log 2)^2}\right).$$

The asymptotic theory of $\widehat{\mu}$ and $\widehat{\kappa}$ defined in Equations (7.5.4) and (7.5.5) can be obtained under the double asymptotic scheme where $T \to \infty$ and $\Delta \to 0$. Under the condition of $\kappa > 0$, Wang et al. (2021) proposed a condition that governs the relative divergence/convergence rates of T and Δ.

Theorem 7.5.5 *Let $\widehat{\mu}$ be the estimator of μ defined in Equation (7.5.4). Then, when $T \to \infty$ and $\Delta \to 0$, we can obtain $\widehat{\mu} \to_p \mu$ for all $H \in (0,1)$. Moreover, if $T^{1-H}\Delta^H \to 0$, then we have*

$$T^{1-H}(\widehat{\mu} - \mu) \Rightarrow \mathcal{N}\left(0, \sigma^2/\kappa^2\right). \tag{7.5.12}$$

Theorem 7.5.6 *Let $\widehat{\kappa}$ be the estimator of κ defined in Equation (7.5.5). Then, when $T \to \infty$ and $\Delta \to 0$, we have $\widehat{\kappa} \to_p \kappa$. Moreover, (a) for $H \in (0, 3/4)$ and $\sqrt{T}\Delta^H \to 0$, then*

$$\sqrt{T}\left(\widehat{\kappa} - \kappa\right) \Rightarrow \left(0, \kappa\phi_H\right), \tag{7.5.13}$$

with

$$\phi_H = \begin{cases} \frac{1}{4H^2}\left[(4H-1) + \frac{2\Gamma(2-4H)\Gamma(4H)}{\Gamma(2H)\Gamma(1-2H)}\right] & \text{if } \ H \in (0, \frac{1}{2}) \\ \frac{4H-1}{4H^2}\left[1 + \frac{\Gamma(3-4H)\Gamma(4H-1)}{\Gamma(2-2H)\Gamma(2H)}\right] & \text{if } \ H \in [\frac{1}{2}, \frac{3}{4}) \end{cases};$$

(b) for $H = 3/4$ and $\sqrt{T}\Delta^H / \log(T) \to 0$, then

$$\frac{\sqrt{T}}{\log(T)}\left(\widehat{\kappa} - \kappa\right) \Rightarrow \mathcal{N}\left(0, \frac{16\kappa}{9\pi}\right);$$

(c) for $H \in (3/4, 1)$ and $T^{2-2H}\Delta^H \to 0$, then

$$T^{2-2H}\left(\widehat{\kappa} - \kappa\right) \Rightarrow \frac{-\kappa^{2H-1}}{H\Gamma(2H+1)} R(H),$$

where $R(H)$ is the Rosenblatt random variable defined by (7.4.24).

Let us observe that the double asymptotic distribution of the MLE of κ is known to be $\mathcal{N}(0, 2\kappa)$ if $H = \frac{1}{2}$ (see, for example, Tang and Chen (2009)). Obviously, $\phi_H = 2$ when $H = \frac{1}{2}$. Therefore, the MME $\widehat{\kappa}$ has the same limiting distribution as the MLE for $H = \frac{1}{2}$, and $\widehat{\kappa}$ is asymptotically efficient when $H = \frac{1}{2}$.

7.6 Conclusions

Stochastic models have played a key role in the development of financial markets. The first mathematical model in finance was introduced by Bachelier (1900), who used the normal distribution to describe the behavior of asset prices. An obvious deficiency in Bachelier's model Bachelier (1900) is that stock prices can lead to negative prices at any time $t \in ([0, T])$. To overcome this problem, Osborne (1959) introduced geometric Brownian motion to describe stock prices. In parallel, Itô (1951) developed the stochastic integral method with respect to Brownian motion and a stochastic differential equation driven by Brownian motion. In 1973, Black and Scholes (1973); Merton (1973) derived the celebrated Black–Scholes option pricing formula in two separate papers, which were awarded the Nobel Prize for Economics in 1997. To date, previous work has addressed pricing equity options, but little attention has been paid to interest rates. Vasicek (1977) developed a framework for pricing interest rate options. This model was the first to use geometric Brownian motion for short-term interest rates. Similar to the well-known Black–Scholes model, the bond pricing equation was postulated as a parabolic partial differential equation in Vasicek (1977). Harrison and Pliska (1981) introduced the risk-neutral pricing formula by using martingale theory, which is an essential tool of stochastic calculus.

From an empirical point of view, a basic requirement for any good model is its ability to capture the main features in the observed prices. However, the assumption of the independence of asset returns and the constant volatility in the Black–Scholes model is clearly contrary to the empirical data, which reveal that there exists a strong dependence of asset returns and implied volatility with respect to strike prices and time to maturity.

First, to consider the dependence of asset returns, the fBm, which allows for dependent increments and long memory, was introduced into finance. In fact, Delbaen and Schachermayer (1994) showed that no-arbitrage pricing is only possible in semimartingale stochastic processes. Therefore, the basic ideas of mathematical finance in Delbaen and Schachermayer (1994) already imply that the fBm allows a certain kind of arbitrage since the fBm is not a semimartingale.

Second, to exclude the constant volatility assumption, some stochastic volatility models driven by Brownian motion that allow the volatility parameter to vary in a random fashion are provided in finance (e.g., Gatheral (2006)). These stochastic volatility models are useful since they can describe empirical observations, i.e., options with different strike prices and expirations have different values of implied volatility. However, the main drawback of stochastic volatility models driven by Brownian motion is that they fail to produce decent results for short maturities. To perform well, some fractional stochastic volatility models with the long memory property were proposed by Comte and Renault (1998). Gatheral et al. (2018) showed that the log RV is rough and behaves as an fBm where H has a value of around 0.1 at any reasonable time scale. This observation motivated them to examine the performance of the fVm for modeling the RV relative to other models. In fact, the fVm has been the subject of active research for the last two decades because this model seems to be one of the most suitable tools for capturing the phenomenon of long-range dependence (e.g., Comte and Renault (1998); Aït-Sahalia and Mancini (2008)) or the roughness of the volatility (e.g., Bayer et al. (2016); Gatheral et al. (2018)) in financial asset volatility. The development of the application for the fVm naturally led to its statistical inference. Consequently, estimating the unknown parameters in the fVm has been the subject of active research for the last decade and is a challenging theoretical problem.

This chapter considers the development and the application of the fBm in finance. First, we discussed the advantages and disadvantages of using the fBm to describe the fluctuations in financial asset returns. Then, we considered the application of the fVm in stochastic volatility models and discussed the free arbitrage property of stochastic volatility models driven by the fBm. Furthermore, we discussed the problem of estimating unknown parameters in the fVm based on both continuous-time observations and discrete-time observations. Our study considered three well-known methods for estimating the drift parameters in the fVm from a continuous record of observations, including the MLE, LSE and MME. Our study also contributes to the literature by developing an estimation method for all parameters in the fVm based on discrete-time observations for all ranges of the Hurst parameter. The application and statistic inference of the fVm with jumps are important directions for future research.

7.7 Appendix

7.7.1 Appendix A: Proof of Proposition 7.2.1

Proof. From Definition 7.2.1, we can see that the fBm is centered. Moreover, using Model (7.2.1), we can easily obtain $\operatorname{Var}\left(B_t^H\right) = \mathbb{E}\left(\left(B_t^H\right)^2\right) = t^{2H}$.

For any $c > 0$ and $s, t \geq 0$, the process B_{ct}^H is a centered Gaussian process with covariance

$$\begin{aligned}\mathbb{E}\left(B_{cs}^H B_{ct}^H\right) &= \frac{1}{2}\left[|ct|^{2H} + |cs|^{2H} - c^{2H}|t-s|^{2H}\right] \\ &= c^{2H}\mathbb{E}\left(B_t^H B_s^H\right) \\ &= \mathbb{E}\left[\left(c^H B_t^H\right)\left(c^H B_s^H\right)\right] . \end{aligned} \tag{7.7.1}$$

Since all processes are centered and Gaussian, Equation (7.7.1) implies that $\left(B_{ct}^H\right) \stackrel{d}{=} \left(|c|^H B_t^H\right)$. Then, we show that the fBm has stationary increments. Note that for $h > 0$, we have

$$\begin{aligned} & \mathbb{E}\left[\left(B_{t+h}^H - B_h^H\right)\left(B_{s+h}^H - B_h^H\right)\right] \\ = \; & \mathbb{E}\left[\left(B_{t+h}^H B_{s+h}^H\right)\right] - \mathbb{E}\left[\left(B_{t+h}^H B_h^H\right)\right] - \mathbb{E}\left[\left(B_{s+h}^H B_h^H\right)\right] + \mathbb{E}\left[\left(B_h^H\right)^2\right] \\ = \; & \frac{1}{2}\Big[\left((t+h)^{2H} + (s+h)^{2H} - |t-s|^{2H}\right) - \left((t+h)^{2H} + h^{2H} - t^{2H}\right) \end{aligned}$$

$$
\begin{aligned}
& \left. - \left((s+h)^{2H} + h^{2H} - s^{2H}\right) + 2h^{2H} \right] \\
= & \frac{1}{2}\left[t^{2H} + s^{2H} - |t-s|^{2H}\right] \\
= & \mathbb{E}\left(B_t^H B_s^H\right) ,
\end{aligned}
$$

which implies $\left(B_{t+h}^H - B_h^H\right) \stackrel{d}{=} B_t^H$.

For $s, t \geq 0$, we have

$$
\begin{aligned}
\mathbb{E}\left|B_t^H - B_s^H\right|^2 &= \mathbb{E}\left|B_t^H\right|^2 - 2\mathbb{E}\left[B_t^H B_s^H\right] + \mathbb{E}\left|B_s^H\right|^2 \\
&= t^{2H} - 2R_H(t,s) + s^{2H} \\
&= |t-s|^{2H} .
\end{aligned}
$$

Since for any $s \leq t$, the random variable $B_t^H - B_s^H$ has the distribution $\sqrt{\mathbf{E}\left|B_t^H - B_s^H\right|^2} \times Z = |t-s|^H Z$, where Z denotes a standard normal random variable, we obtain that, for any $p \geq 1$

$$
\mathbb{E}\left|B_t^H - B_s^H\right|^p = \mathbf{E}|Z|^p |t-s|^{Hp} . \tag{7.7.2}
$$

The Hölder continuity follows from Equation (7.7.2) and the Kolmogorov continuity theorem.

We can see that if $H \neq \frac{1}{2}$, B_H does not satisfy the condition of $R_H(s,u)R_H(t,t) = R_H(s,t)R_H(t,u)$.

Indeed, the covariance of its increments, also known as fractional Gaussian noise (fGn), is given by

$$
\begin{aligned}
\rho_H(n) &= \mathbb{E}\left(\left(B_i^H - B_{i-1}^H\right)\left(B_{i+n}^H - B_{i-1+n}^H\right)\right) \\
&= \frac{1}{2}\left((n+1)^{2H} + (n-1)^{2H} - 2n^{2H}\right) \\
&\sim H(2H-1)n^{2H-2} \quad \text{as } n \to \infty . \qquad (7.7.3)
\end{aligned}
$$

From Equation (7.7.3), we have $\rho_H(n) = 0$ for $H = \frac{1}{2}$, suggesting that B_t^H becomes a standard Brownian motion with independent increments and is therefore uncorrelated. When $H \in \left(\frac{1}{2}, 1\right)$, it can be seen that $\rho_H(n) > 0$, which means That the increments of B_t^H are positively correlated to each other, making $\sum_{n=1}^{\infty} \rho_H(n) = \infty$ as $n \to \infty$. However, when $H \in \left(0, \frac{1}{2}\right)$, the increments of B_t^H are negatively correlated to each other and have the short-range dependence property, i.e., $\sum_{n=1}^{\infty} \rho_H(n) < \infty$.

Set $Y_n := (2^n)^{pH-1} \sum_{k=1}^{2^n} \left|B_{t_k}^H - B_{t_{k-1}}^H\right|^p$. By self-similarity, we have

$$
Y_n \sim (2^n)^{pH-1} \sum_{k=1}^{2^n} |t_k - t_{k-1}|^{pH} \left|B_k^H - B_{k-1}^H\right|^p = 2^{-n} \sum_{k=1}^{2^n} \left|B_k^H - B_{k-1}^H\right|^p .
$$

The sequence of one step increments of $B_k^H - B_{k-1}^H$ is a stationary, centered Gaussian with covariance function $\rho_H(n)$, which tends to zero. Therefore, the fGn is ergodic. According to Birkhof's Ergodic Theorem,[7]

$$
2^{-n} \sum_{k=1}^{2^n} \left|B_k^H - B_{k-1}^H\right|^p \to \mathbb{E}\left|B_1^H\right|^p \to_{a.s} C .
$$

Therefore, $Y_n \to_p C$. Because $\left|B_{t_k}^H - B_{t_{k-1}}^H\right|^p = (2^{-n})^{pH-1} Y_n$, we obtain Equation (7.2.2).

[7] Birkhof's Ergodic Theorem: Let $\{\xi_n\}$ be a stationary (strict sense), ergodic random sequence with $\mathbb{E}|\xi_1| < \infty$. Then, $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \xi_k(\omega) = \mathbb{E}(\xi_1)$ almost surely and in L^1.

From Equation (7.2.2), we can see that the index of the p-variation of the fBm is $\frac{1}{H}$. However, for a semimartingale, the index must be either in $[0,1]$ or equal to 2, i.e., $\frac{1}{H} \in [0,1] \cup \{2\}$. Since $H \in (0,1)$, $H^{-1} \notin [0,1]$, the fBm is a semimartingale only for $H = \frac{1}{2}$.

Using the self-similarity property, we have

$$\frac{B_t^H - B_{t_0}^H}{t - t_0} \overset{d}{=} (t - t_0)^{H-1} B_1^H .$$

Let us consider the event $A(t) = \left\{ \sup_{t_0 \le s \le t} \left| \frac{B_s^H - B_{t_0}^H}{s - t_0} \right| > d \right\}$. Then, for any sequence t_n decreasing to t_0, we have $A(t_n) \supset A(t_{n+1})$ and $A(t_n) \supset \left\{ \left| \frac{B_{t_n}^H - B_{t_0}^H}{t_n - t_0} \right| > d \right\}$.

Thus, we have

$$\mathbb{P}\left(\left\{ \left| \frac{B_{t_n}^H - B_{t_0}^H}{t_n - t_0} \right| > d \right\} \right) = \mathbb{P}\left(\left\{ \left| B_1^H \right| > (t_n - t_0)^{1-H} d \right\} \right) \to 1, \text{ as } n \to \infty .$$

Note that there is a stronger form since here we prove the assertion locally. ■

7.7.2 Appendix A: Proof of Proposition 7.2.2

Proof. From Sinai (1976), we can easily obtain that the spectral density of the fGn can be written as

$$f_G(\lambda; H) = F(H)(1 - \cos\lambda) \sum_{j \in \mathbb{Z}} |\lambda + 2j\pi|^{-2H-1} ,$$

where $F(H)$ is a normalizing factor designated to ensure $\int_{-\pi}^{\pi} f_G(\lambda; H)\, d\lambda = 1$.

Consequently, we have

$$f_G(\lambda; H) = \frac{\sigma^2}{\pi} \sin(\pi H)\Gamma(2H+1)(1 - \cos\lambda) \sum_{j \in \mathbb{Z}} |\lambda + 2j\pi|^{-2H-1} .$$

The behavior of $f_G(\lambda; H)$ near the origin is followed by a Taylor expansion of $(1 - \cos\lambda)$ at zero, and it can be shown that

$$f_G(\lambda; H) \sim \mathcal{O}\left(|\lambda|^{1-2H}\right) \quad \text{as} \quad \lambda \to 0 .$$

The asymptotic behavior of $\gamma_G(k)$ is follows by the Taylor expansion . Let $\gamma_G(k) = \frac{1}{2} k^{2H} g\left(k^{-1}\right)$ with $g(x) = (1+x)^{2H} - 2 + (1-x)^{2H}$. If $0 < H < 1$ and $H \neq \frac{1}{2}$, then the first non-zero term in the Taylor expansion of $g(x)$, expanded at the origin, is equal to $2H(2H-1)x^2$. Therefore, as k tends to infinity, we have (7.2.4). For $\frac{1}{2} < H < 1$, the fGn has long memory. For $H = \frac{1}{2}$, all the correlations at non-zero lags are zero, i.e., the observations are uncorrelated, and the process has short memory. For $0 < H < \frac{1}{2}$, it can be shown that

$$\sum_{k=-\infty}^{\infty} \gamma_G(k) = 0$$

and therefore, the process is anti-persistent.

A standard calculation yields

$$\begin{aligned}
\mathbb{E}\left[G_t^H - G_0^H\right] &= \mathbb{E}\left[G_t^H\right] - \mathbb{E}\left[G_0^H\right] = 0 \\
\operatorname{Var}\left(G_t^H - G_0^H\right) &= \mathbb{E}\left(Y_H(n) - Y_H(0)\right)^2 \\
&= 2 - |t-1|^{2H} - |t+1|^{2H} + 2t^{2H}
\end{aligned}$$

as

$$\begin{aligned}\mathbb{E}\left[G_t^H G_0^H\right] &= \mathbb{E}\left[\left(G_{t+1}^H - G_t^H\right)\left(G_1^H - G_0^H\right)\right] \\ &= \frac{1}{2}\left[|t-1|^{2H} + |t+1|^{2H} - 2t^{2H}\right] .\end{aligned}$$

Moreover, $\left(G_t^H, G_0^H\right)$ has a bivariate Gaussian distribution with the correlation matrix given by

$$\begin{pmatrix} 1 & \frac{1}{2}\left[|t-1|^{2H} + |t+1|^{2H} - 2t^{2H}\right] \\ \frac{1}{2}\left[|t-1|^{2H} + |t+1|^{2H} - 2t^{2H}\right] & 1 \end{pmatrix} . \tag{7.7.4}$$

Thus, $G_t^H - G_0^H$ has a Gaussian distribution with zero mean and the variance $2 - |t-1|^{2H} - |t+1|^{2H} + 2|t|^{2H}$. Of course, for $t = 0$, we have a degenerated distribution with zero variance. Moreover, using $\lim_{t\to\infty}\left(|t-1|^{2H} + |t+1|^{2H} - 2t^{2H}\right) = 0$, we have $\mathrm{Var}\left(G_t^H - G_0^H\right) \longrightarrow 2$ as $t \to \infty$. Therefore, $G_t^H - G_0^H \Rightarrow \mathcal{N}(0,2)$ as $t \to \infty$. ■

Bibliography

Aït-Sahalia, Y. and L. Mancini (2008). Out of sample forecasts of quadratic variation. *Journal of Econometrics 147*(1), 17–33.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*(2), 579–625.

Bachelier, L. (1900). Théorie de la spéculation. In *Annales scientifiques de l'École normale supérieure*, Volume 17, pp. 21–86.

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of econometrics 73*(1), 5–59.

Baillie, R. T., T. Bollerslev, and H. O. Mikkelsen (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 74*(1), 3–30.

Bakshi, G., C. Cao, and Z. Chen (1997). Empirical performance of alternative option pricing models. *Journal of Finance 52*(5), 2003–2049.

Barndorff-Nielsen, O. E., J. M. Corcuera, and M. Podolskij (2013). Limit theorems for functionals of higher order differences of brownian semi-stationary processes. In *Prokhorov and Contemporary Probability Theory: In Honor of Yuri V. Prokhorov*, Volume 33, pp. 69–96. Springer Science & Business Media.

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *Review of financial studies 9*(1), 69–107.

Bayer, C., P. Friz, and J. Gatheral (2016). Pricing under rough volatility. *Quantitative Finance 16*(6), 887–904.

Belfadli, R., K. Es-Sebaiy, and Y. Ouknine (2011). Parameter estimation for fractional ornstein-uhlenbeck processes: non-ergodic case. *Frontiers in Science and Engineering (An International Journal Edited by Hassan II Academy of Science and Technology) 1*(1), 1–16.

Bender, C. and R. J. Elliott (2004). Arbitrage in a discrete version of the wick-fractional black-scholes market. *Mathematics of Operations Research 29*(4), 935–945.

Bennedsen, M., A. Lunde, and M. S. Pakkanen (2021, 01). Decoupling the Short- and Long-Term Behavior of Stochastic Volatility. *Journal of Financial Econometrics*. nbaa049.

Beran, J. (1994). *Statistics for long-memory processes*. New York: CRC press.

Bishwal, J. (2011). Minimum contrast estimation in fractional ornstein-uhlenbeck process: Continuous and discrete sampling. *Fractional Calculus and Applied Analysis 14*(3), 375–410.

Björk, T. and H. Hult (2005). A note on wick products and the fractional black-scholes model. *Finance and Stochastics 9*(2), 197–209.

Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy 81*(3), 637–654.

Bolko, A. E., K. Christensen, M. S. Pakkanen, and B. Veliyev (2022). A gmm approach to estimate the roughness of stochastic volatility. *Journal of Econometrics, forthcoming*.

Chan, K. C., G. A. Karolyi, F. A. Longstaff, and A. B. Sanders (1992). An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance 47*(3), 1209–1227.

Cheridito, P. (2001). Mixed fractional brownian motion. *Bernoulli 7*(6), 913–934.

Cheridito, P., H. Kawaguchi, and M. Maejima (2003). Fractional ornstein-uhlenbeck processes. *Electronic Journal of Probability 8*(3), 1–14.

Chronopoulou, A. and F. G. Viens (2012a). Estimation and pricing under long-memory stochastic volatility. *Annals of Finance 8*(2-3), 379–403.

Chronopoulou, A. and F. G. Viens (2012b). Stochastic volatility and option pricing with long-memory in discrete and continuous time. *Quantitative Finance 12*(4), 635–649.

Comte, F., L. Coutin, and É. Renault (2012). Affine fractional stochastic volatility models. *Annals of Finance 8*(2-3), 337–378.

Comte, F. and E. Renault (1998). Long memory in continuous-time stochastic volatility models. *Mathematical Finance 8*(4), 291–323.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance 1*(2), 223–236.

Cont, R. (2005). Long range dependence in financial markets. In *In Fractals in Engineering: New Trends in Theory and Application. Lutton, E. and Vehel, J., (Eds.)*, pp. 159–179. Springer, New York.

Corlay, S., J. Lebovits, and J. L. Véhel (2014). Multifractional stochastic volatility models. *Mathematical Finance 24*(2), 364–402.

Delbaen, F. and W. Schachermayer (1994). A general version of the fundamental theorem of asset pricing. *Mathematische annalen 300*(1), 463–520.

Duffie, D., J. Pan, and K. Singleton (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica 68*(6), 1343–1376.

Duncan, T., Y. Hu, and B. Pasik-Duncan (2000). Stochastic calculus for fractional brownian motion i. theory. *SIAM Journal on Control and Optimization 38*(2), 582–612.

El Euch, O. and M. Rosenbaum (2019). The characteristic function of rough heston models. *Mathematical Finance 29*(1), 3–38.

El Machkouri, M., K. Es-Sebaiy, and Y. Ouknine (2016). Least squares estimator for non-ergodic ornstein–uhlenbeck processes driven by gaussian processes. *Journal of the Korean Statistical Society 45*(3), 329–341.

Elliott, R. and L. Chan (2004). Perpetual american options with fractional brownian motion. *Quantitative Finance 4*(2), 123–128.

Elliott, R. and J. Van Der Hoek (2003). A general fractional white noise theory and applications to finance. *Mathematical Finance 13*(2), 301–330.

Fink, H., C. Klüppelberg, and M. Zähle (2013). Conditional distributions of processes related to fractional brownian motion. *Journal of applied probability 50*(1), 166–183.

Fouque, J.-P., G. Papanicolaou, and K. R. Sircar (2000). *Derivatives in financial markets with stochastic volatility*. Cambridge University Press.

Gatheral, J. (2006). *The volatility surface: a practitioner's guide*. John Wiley & Sons.

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

Greene, M. T. and B. D. Fielitz (1977). Long-term dependence in common stock returns. *Journal of Financial Economics 4*(3), 339–349.

Hagan, P. S., D. Kumar, A. S. Lesniewski, and D. E. Woodward (2002). Managing smile risk. *The Best of Wilmott 1*, 249–296.

Harrison, J. M. and S. R. Pliska (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic processes and their applications 11*(3), 215–260.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies 6*(2), 327–343.

Hu, Y. and D. Nualart (2010). Parameter estimation for fractional ornstein–uhlenbeck processes. *Statistics and Probability Letters 80*(11), 1030–1038.

Hu, Y., D. Nualart, and H. Zhou (2019). Parameter estimation for fractional ornstein–uhlenbeck processes of general hurst parameter. *Statistical Inference for Stochastic Processes 22*(1), 111–142.

Hu, Y. and B. Øksendal (2003). Fractional white noise calculus and applications to finance. *Infinite Dimensional Analysis Quantum Probability and Related Topics 6*(1), 1–32.

Hu, Y., B. Øksendal, and A. Sulem (2003). Optimal consumption and portfolio in a Black–Scholes market driven by fractional brownian motion. *Infinite Dimensional Analysis, Quantum Probability and Related Topics 6*(4), 519–536.

Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance 42*(2), 281–300.

Hurst, H. E. (1951). Long-term storage capacity of reservoirs. *Trans. Amer. Soc. Civil Eng. 116*, 770–799.

Itô, K. (1951). *On stochastic differential equations*. American Mathematical Soc.

Kleptsyna, M. and A. Le Breton (2002). Statistical analysis of the fractional ornstein–uhlenbeck type process. *Statistical Inference for Stochastic Processes 5*(3), 229–248.

Kleptsyna, M. L., A. Le Breton, and M.-C. Roubaud (2000). Parameter estimation and optimal filtering for fractional type stochastic systems. *Statistical inference for stochastic processes 3*(1), 173–182.

Kolmogorov, A. N. (1940). Wienersche spiralen und einige andere interessante kurven in hilbertscen raum, cr (doklady). *Acad. Sci. URSS (NS) 26*, 115–118.

Kubilius, K., Y. Mishura, and K. Ralchenko (2017). *Parameter estimation in fractional diffusion models*. Springer.

Lang, G. and F. Roueff (2001). Semi-parametric estimation of the hölder exponent of a stationary gaussian process with minimax rates. *Statistical inference for stochastic processes 4*(3), 283–306.

Liu, X., S. Shi, and Y. Jun (2020). Persistent and rough volatility. *Working Paper*.

Livieri, G., S. Mouti, A. Pallavicini, and M. Rosenbaum (2018). Rough volatility: evidence from option prices. *IISE Transactions 50*(9), 767–776.

Lo, A. W. and A. C. MacKinlay (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *The review of financial studies 1*(1), 41–66.

Lohvinenko, S. and K. Ralchenko (2017). Maximum likelihood estimation in the fractional vasicek model. *Lithuanian Journal of Statistics 56*(1), 77–87.

Lohvinenko, S. and K. Ralchenko (2019). Maximum likelihood estimation in the non-ergodic fractional vasicek model. *Modern Stochastics: Theory and Applications 6*(3), 377–395.

Ludeña, C. (2004). Minimum contrast estimation for fractional diffusions. *Scandinavian journal of statistics 31*(4), 613–628.

Mandelbrot, B. B. and J. W. Van Ness (1968). Fractional brownian motions, fractional noises and applications. *SIAM review 10*(4), 422–437.

Merton, R. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science 4*(1), 141–183.

Mishura, Y. (2008). *Stochastic calculus for fractional Brownian motion and related processes*. Springer.

Norros, I., E. Valkeila, J. Virtamo, et al. (1999). An elementary approach to a girsanov formula and other analytical results on fractional brownian motions. *Bernoulli 5*(4), 571–587.

Nourdin, I. (2012). *Selected aspects of fractional Brownian motion*, Volume 4. Springer.

Osborne, M. F. (1959). Brownian motion in the stock market. *Operations research 7*(2), 145–173.

Phillips, P. C. B. and J. Yu (2009). A two-stage realized volatility approach to estimation of diffusion processes with discrete data. *Journal of Econometrics 150*(2), 139–150.

Prakasa Rao, B. (2010). *Statistical Inference for Fractional Diffusion Processes*. Wiley, Chichester.

Rogers, L. C. G. (1997). Arbitrage with fractional brownian motion. *Mathematical Finance 7*(1), 95–105.

Romano, M. and N. Touzi (1997). Contingent claims and market completeness in a stochastic volatility model. *Mathematical Finance 7*(4), 399–412.

Rostek, S. (2009). *Option pricing in fractional Brownian markets*. Springer Verlag.

Samorodnitsky, G. and M. Taqqu (1994). *Stable non-gaussian random processes*. Chapman & Hall, New York.

Schöbel, R. and J. Zhu (1999). Stochastic volatility with an ornstein–uhlenbeck process: an extension. *European Finance Review 3*(1), 23–46.

Scott, L. O. (1987). Option pricing when the variance changes randomly: Theory, estimation, and an application. *Journal of Financial and Quantitative analysis 22*(04), 419–438.

Shi, S. and J. Yu (2022). Volatility puzzle. *Management Science, accepted*.

Sinai, Y. G. (1976). Self-similar probability distributions. *Theory of Probability & Its Applications 21*(1), 64–80.

Sottinen, T. (2001). Fractional brownian motion, random walks and binary market models. *Finance and Stochastics 5*(3), 343–355.

Stein, E. M. and J. C. Stein (1991). Stock price distributions with stochastic volatility: an analytic approach. *Review of Financial Studies 4*(4), 727–752.

Tanaka, K. (2013). Distributions of the maximum likelihood and minimum contrast estimators associated with the fractional ornstein–uhlenbeck process. *Statistical Inference for Stochastic Processes 16*(3), 173–192.

Tanaka, K. (2015). Maximum likelihood estimation for the non-ergodic fractional ornstein–uhlenbeck process. *Statistical Inference for Stochastic Processes 18*(3), 315–332.

Tanaka, K., W. Xiao, and J. Yu (2020). Maximum likelihood estimation for the fractional vasicek model. *Econometrics 8*(3), 32.

Tang, C. Y. and S. X. Chen (2009). Parameter estimation and bias correction for diffusion processes. *Journal of Econometrics 149*(1), 65–81.

Tudor, C. and F. Viens (2007). Statistical aspects of the fractional stochastic calculus. *Annals of Statistics 35*(3), 1183–1212.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics 5*(2), 177–188.

Wang, X., W. Xiao, and J. Yu (2021). Modeling and forecasting realized volatility with the fractional ornstein–uhlenbeck process. *Journal of Econometrics*.

Wang, X. and J. Yu (2016). Double asymptotics for explosive continuous time models. *Journal of Econometrics 193*(1), 35–53.

Wiggins, J. B. (1987). Option values under stochastic volatility: Theory and empirical estimates. *Journal of Financial Economics 19*(2), 351–372.

Willinger, W., M. S. Taqqu, and V. Teverovsky (1999). Stock market prices and long-range dependence. *Finance and stochastics 3*(1), 1–13.

Xiao, W. and J. Yu (2019a). Asymptotic theory for estimating drift parameters in the fractional vasicek model. *Econometric Theory 35*(1), 198–231.

Xiao, W. and J. Yu (2019b). Asymptotic theory for rough fractional vasicek models. *Economics Letters 177*, 26–29.

Yu, J. and P. C. B. Phillips (2001). A Gaussian approach for continuous time models of the short-term interest rate. *Econometrics Journal 4*(2), 210–224.