Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsapl/100305

Asymptotic Equivalence of Discretely Observed Fractional Randleman-Bartter Model to a Fractional Gaussian Shift

Jaya P. N. Bishwal

Department of Mathematics and Statistics, University of North Carolina at Charlotte, 376 Fretwell Building, 9201 University City Blvd., Charlotte, NC 28223, USA

Received: 22 Jun. 2023, Revised: 22 Jul. 2023, Accepted: 31 Jul. 2023
Published online: 1 Sep. 2023

Abstract: We show that the solution of the fractional Rendleman-Bartter model observed in discrete time points is asymptotically equivalent to a fractional Gaussian white noise model. Next we estimate the drift parameter of fractional diffusions after discretizing the models by using a weak perturbed random walk approximation of fractional Brownian motion. Then we study the problem of testing of the parametric form of the volatility in a stochastic differential equation driven by fractional Brownian motion. Finally we study semiparametric estimation of forward-backward stochastic differential equations.

Keywords: Fractional Rendleman-Bartter model, fractional white noise model, perturbed random walk, asymptotic equivalence, fractional Black-Scholes model, fractional Hull-White model, fractional Vasicek model, fractional Cox-Ingersoll-Ross model, volatility persistence, local time, nonparametric testing, forward-backward stochastic differential equations, semiparametric estimation

1. Fractional Geometric Brownian Motion

We study a class of statistical inference problems for several stochastic differential equation models which are useful for financial modeling. These include parametric estimation, semiparametric estimation and nonparametric testing. The nonparametric test is based on local time of the corresponding diffusion process. We consider forward-backward stochastic differential equation (FBSDE) where for a backward stochastic differential equation (BSDE), the initial condition comes from the forward equation. The forward equation is driven by fractional Brownian motion. These are inverse problems for finance where given the data one needs to make estimation and testing of the unknown parameters. There is a hidden backward process which is completely unobserved. It can be viewed as a nonparametric approximation of the observed variable. Asymptotic equivalence of models is a way where if a model is not tractable, one can find alternative equivalent models. Asymptotic equivalence of statistical experiments has received some attention, see Nussbaum (1996).

The fractional Brownian motion (fBm, in short), which provides a suitable generalization of the Brownian motion, is one of the simplest stochastic processes exhibiting long range-dependence. Consider a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ on which all random variables and processes below are defined. A normalized fractional Brownian motion $\{W_t^H, t \geq 0\}$ with Hurst parameter $H \in (0,1)$ is a centered Gaussian process with continuous sample paths whose covariance kernel is given by

$$E(W_t^H W_s^H) = \frac{1}{2}(s^{2H} + t^{2H} - |t-s|^{2H}), \quad s,t \geq 0. \tag{1.1}$$

The process is self similar (scale invariant) and it can be represented as a stochastic integral with respect to standard Brownian motion. For $H = \frac{1}{2}$, the process is a standard Brownian motion. For $H \neq \frac{1}{2}$, the fBm is not a semimartingale and not a Markov process, but a Dirichlet process. The increments of the fBm are negatively correlated for $H < \frac{1}{2}$ and positively correlated for for $H > \frac{1}{2}$ and in this case they display long-range dependence (persistence). The parameter H which is also called the self similarity parameter, measures the intensity of the long range dependence.

* Corresponding author e-mail: J.Bishwal@uncc.edu

Hence for $H \neq \frac{1}{2}$, the classical theory of stochastic integration with respect to semimartingales is not applicable to stochastic integration with respect to fBm. Now there exists several approaches to stochastic integration with respect to fBm.

The fractional Black-Scholes model for stock price is given by

$$dS_t = \mu S_t dt + \sigma S_t dW_t^H, \quad t \in [0,1]. \tag{1.2}$$

The white noise theory for fractional Brownian motion which was developed in Hu and Oksendal (2003) (see also Duncan, Hu and Pasik-Duncan (2001)) shows that if the corresponding integration theory (in the Itô sense, i.e. based on the Wick product rather than the pathwise product) is used, then the corresponding fractional Black-Scholes market is arbitrage-free and complete and explicit fractional option pricing and hedging formulae can be given.

We shall use the following *fractional Girsanov theorem* in the sequel:

Theorem 1.1 Let $T > 0$ and let $u : [0,T] \to \mathbb{R}$ be continuous. Suppose K satisfies the equation

$$\int_0^T K(s)\phi(s,t)ds = u(t); \quad 0 \leq t \leq T \tag{1.3}$$

where $\phi(s,t) = H(2H-1)|s-t|^{2H-2}, s,t \in \mathbb{R}$ and extend K to $\mathbb{R}$ by putting $K(s) = 0$ outside $[0,T]$. Define the probability measure $\widehat{P}_H$ on $\mathscr{F}_T^{(H)}$ by

$$d\widehat{P}_H(\omega) = \exp\left\{-\int_0^T K(s)dW_s^H - \frac{1}{2}|K|_\phi^2\right\} dP_H(\omega) \tag{1.4}$$

where

$$|K|_\phi^2 = \int_{\mathbb{R}}\int_{\mathbb{R}} K(s)K(t)\phi(s,t)dsdt. \tag{1.5}$$

Then

$$\widehat{W}_t^H := \int_0^t u(s)ds + W_t^H \tag{1.6}$$

is a fractional Brownian motion with respect to $\widehat{P}_H$.

2. Asymptotic Equivalence

In Black-Scholes model, volatility is assumed to be constant. We will consider time-dependent volatility function. The fractional Randleman-Barter model is given by

$$dS_t = \mu S_t dt + \sigma(t) S_t dW_t^H, \quad t \in [0,1] \tag{2.1}$$

where $\sigma(t) \geq \varepsilon > 0$ for all $t \in [0,1]$. Without loss of generality, we assume that the initial value $S_0 = 1$. Let $\mu = 0$. This is a model for short rate, though it is not mean-reverting, it still preserves the positivity of the interest rate. Recall that Ho-Lee model for short-rate has time dependent drift, but still constant volatility whereas Hull-White model for short rate has time dependent drift and volatility which are nonhomogeneous extensions of Vasicek model. Another alternative to Randleman-Barter model is the Black-Karasinski model which is nonhomogeneous extension of exponential Vasicek model.

Consider $Y_t = \log S_t$ and substitute this in equation (2.1). Then

$$dY_t = -\frac{\sigma^2(t)}{2}dt + \sigma(t)dW_t^H, \; t \geq 0. \tag{2.2}$$

Consider the normalized increments

$$Z_i^n := n^{1-H}(Y_{t_i} - Y_{t_{i-1}}) = -\frac{n^{1-H}}{2}\int_{t_{i-1}}^{t_i} \sigma^2(u)du + n^{1-H}\int_{t_{i-1}}^{t_i} \sigma(u)dW_u^H, \; i = 1,2,\cdots,n. \tag{2.3}$$

We show that the statistical experiment induced by the observations $\{Z_i^n, \; i = 1,2,\cdots,n\}$ with $t_i = \frac{i}{n}$ has the same asymptotic behavior as the fractional white noise shift model:

$$dX_t^n = \frac{1}{\sqrt{2}}\log\sigma(t)dt + \frac{1}{n^{1-H}}dW_t^H, \; t \geq 0. \tag{2.4}$$

Our approach is in the sense of Le Cam theory of asymptotic equivalence of statistical experiments in a nonparametric setup, see Le Cam and Yang (1990). This approach has been used by Brown and Low (1996), Nussbaum (1996) and Grama and Nussbaum (1998) in the classical case of short memory models.

We assume that the volatility function σ belongs to a Hölder class of functions, $R(\gamma,L)$ with $\gamma \in (1/2,1]$ of functions that are bounded away from zero:

$$R(\gamma,L) = \{\sigma : [0,1] \to [\varepsilon,\infty) \; |\varepsilon > 0, \; |\sigma(t) - \sigma(s)| \le L|t-s|^{\gamma}\}. \tag{2.5}$$

We obtain that the fractional geometric Brownian motion observed over a regular grid is asymptotically equivalent with the observations of drifted fractional white noise model in the sense of *Le Cam's deficiency measure* or the Δ-pseudo distance tending to 0 as $n \to \infty$.

Let $\mathscr{E} = (\Omega^1, \mathscr{F}^1, \{P_\theta, \theta \in \Theta\})$ and $\mathscr{G} = (\Omega^2, \mathscr{F}^2, \{Q_\theta, \theta \in \Theta\})$. Le Cam's deficiency measure Δ between the experiments $\mathscr{E}$ and $\mathscr{G}$ is defined as

$$\Delta(\mathscr{E},\mathscr{G}) = \max\{\delta(\mathscr{E},\mathscr{G}), \delta(\mathscr{G},\mathscr{E})\} \tag{2.6}$$

where

$$\delta(\mathscr{E},\mathscr{G}) = \inf_K \sup_{\theta\in\Theta} \|K \cdot P_\theta - Q_\theta\| \tag{2.7}$$

with $\|\cdot\|$ being the total variation norm and the inf is taken over the set $\mathscr{M}(\Omega^1,\mathscr{F}^2)$ of all Markov kernels K from $(\Omega^1,\mathscr{F}^1)$ to $(\Omega^2,\mathscr{F}^2)$. Recall that a Markov kernel K from $(\Omega^1,\mathscr{F}^1)$ to $(\Omega^2,\mathscr{F}^2)$ is a mapping from Ω^1 into the set of probability measure on $(\Omega^2,\mathscr{F}^2)$ such that for all $F \in \mathscr{F}^2$, $\omega \to K(\omega,F)$ is measuraable on $(\Omega^1,\mathscr{F}^1)$, and for all $\omega \in \Omega_1$, $K(\omega,dx)$ is a probability measure on $(\Omega^2,\mathscr{F}^2)$. The experiment $K\mathscr{E} = (\Omega^1,\mathscr{F}^1,\{K\cdot P_\theta, \theta\in\Theta\})$ is a called a randomization of $\mathscr{E}$ by the kernel K. When the kernel is deterministic, that is K is a random variable $K(\omega,F) = I_F(K(\omega))$, the experiment $K\mathscr{E}$ is called the image experiment by K.

Let $\mathscr{E}^n$ and $\mathscr{G}^n, n = 1,2,\ldots$ be two sequences of statistical experiments. The experiments $\mathscr{E}^n$ and $\mathscr{G}^n$ are said to be asymptotically equivalent if $\Delta(\mathscr{E}^n,\mathscr{G}^n) \to 0$ as $n\to\infty$. This in turn indicates that the asymptotically minimax risks for bounded loss functions in one model can be transferred to another model. In particular, one can compute the asymptotically minimax risk in non-Gaussian model by computing it in the accompanying Gaussian models.

Assume that there is some point $\theta_0 \in \Theta$ such that $P_\theta \ll P_{\theta_0}$ and $Q_\theta \ll Q_{\theta_0}, \theta\in\Theta$ and there are versions of Λ_θ^1 and Λ_θ^2 of the Radon-Nikodym derivatives or likelihoods (see Theorem 1.1) dP_θ/dP_{θ_0} and dQ_θ/dQ_{θ_0} on a common probability space $(\Omega,\mathscr{F},P)$. Then the Le Cam deficiency distance between $\mathscr{E}$ and $\mathscr{G}$ satisfies

$$\Delta(\mathscr{E},\mathscr{G}) \le \sup_{\theta\in\Theta} \frac{1}{2} E_P|\Lambda_\theta^1 - \Lambda_\theta^2|. \tag{2.8}$$

Define the measures $\widetilde{P}_\theta$ and $\widetilde{Q}_\theta$ by setting $d\widetilde{P}_\theta = \Lambda_\theta^1 dP$ and $d\widetilde{Q}_\theta = \Lambda_\theta^2 dQ$. Then

$$\frac{1}{2} E_P|\Lambda_\theta^1 - \Lambda_\theta^2| \le \sqrt{2} H(\widetilde{P}_\theta, \widetilde{Q}_\theta) \tag{2.9}$$

where $H(\cdot,\cdot)$ is the Hellinger distance which is defined as follows: if P and Q are probability measures on the measurable space $(\Omega,\mathscr{A})$ and $P \ll \nu$ and $Q \ll \nu$, where ν is a σ-finite measure on $(\Omega,\mathscr{A})$, then

$$H^2(P,Q) = \frac{1}{2}\int_\Omega \left(\left(\frac{dP}{d\nu}\right)^{1/2} - \left(\frac{dQ}{d\nu}\right)^{1/2} \right)^2 d\nu. \tag{2.10}$$

We first obtain the rates of convergence to 0 of the Hellinger distnace between the corresponding measures of two statistical experiments. The inequalities (2.8) and (2.9) are then used in getting upper bound for the Le Cam distance between equivalent experiments.

If Φ_μ is a normal distribution with mean μ and variance 1, then

$$H^2(\Phi_{\mu_1}, \Phi_{\mu_2}) = 1 - \exp\left(\frac{1}{8}(\mu_1-\mu_2)^2\right). \tag{2.11}$$

Let $P_1,\ldots,P_n$ and $Q_1,\ldots,Q_n$ be probability measures on $(\Omega,\mathscr{A})$. Let $P^n = P_1 \times \ldots \times P_n$ and $Q^n = Q_1\times\ldots\times Q_n$. Then

$$1 - H^2(P^n,Q^n) = \prod_{i=1}^n (1 - H^2(P_i,Q_i)) \tag{2.12}$$

and

$$H^2(P^n, Q^n) \le \sum_{i=1}^{n} H^2(P_i, Q_i). \tag{2.13}$$

The Kullback divergence of P with respect to Q is given by

$$K(P,Q) = \int \log \frac{dP}{dQ} dP = \frac{1}{2} \int \left| \left(\frac{dP}{d\nu}\right) - \left(\frac{dQ}{d\nu}\right) \right| d\nu \tag{2.14}$$

if $P \ll Q$, $= \infty$ otherwise. The following Pinsker inequality is well known that the total variation distance

$$\sup_{A \in \mathscr{A}} |P(A) - Q(A)| \le \sqrt{K(P,Q)/2}. \tag{2.15}$$

Milstein and Nussbaum (1998) studied diffusion approximation for nonparametric autoregression and obtained the bound $O(n^{-2}\varepsilon^{-2} + n^{-1})^{1/2}$ on the Le Cam distance as the intensity of diffusion noise $\varepsilon \to 0$ and $n\varepsilon \to \infty$ as $n \to \infty$ where n is the number of observations of the autoregression. Genon-Catalot and Laredo (2014) obtained asymptotic equivalence of nonparametric diffusion and Euler scheme when $T \to \infty$, $n \to \infty$ and $T^2/n \to 0$. In fact, they obtain the bound on the Le Cam distance for the two statistical experiments to be $O(T^2/n)^{1/2}$. However, they assume the nonconstant diffusion coefficient to be known. They used time change on the nonparametric diffusion and the Euler scheme which is defined as the first time the integrated squared diffusion coefficient hits a specified level. The time changed processes have unit diffusion coefficient. In the small variance case, replacing σ by $\sigma\varepsilon$, the upper bound for the rate of convergence is $O(n^{-2}\varepsilon^{-2} + n^{-1} + n^{-1}\varepsilon^{-4})^{1/2}$. Mariucci (2016) extended the results to small variance case and obtained the bound on the Le Cam distance to be $O(n^{-1}\varepsilon^{-1} + n^{-1} + \varepsilon)^{1/4}$). Asymptotic equivalence for interacting diffusions for the continuous observation of one path and for n i.i.d. paths was studied in Genon-Catalot and Laredo (2021) as $\sqrt{n}\varepsilon \to 0$.

We consider fractional diffusions. Consider the following random variables:

$$Z_i^* = -\frac{\sigma^2(t_i)}{2n^{1-H}} + \sigma(t_i) \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} K^H\left(\frac{[nt]}{n}, s\right) ds \right) \xi_i, \; i = 1, 2, \cdots, n \tag{2.16}$$

$$B_t^{H,n} = \sqrt{n} \sum_{i=1}^{[nt]} \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} K^H\left(\frac{[nt]}{n}, s\right) ds \right) \xi_i, \; i = 1, 2, \cdots, n \tag{2.17}$$

where the ξ_i are i.i.d. standard normal random variables and

$$K^H(t,s) = c_H s^{1/2-H} \int_s^t (u-s)^{H-3/2} u^{H-1/2} du, \; t > s \tag{2.18}$$

where $c_H = \left(\frac{H(2H-1)}{B(2-2H, H-1/2)}\right)^{1/2}$ and $B(\cdot,\cdot)$ is the beta function. Note that $B^{H,n}$ converges weakly as $n \to \infty$ in Skorohod topology to the fBM, see Sottinen (2001).

Consider the driftless fractional model

$$dV_t = \sigma(t) dW_t^H \tag{2.19}$$

and discrete observations of V_t at times $t_i = \frac{i}{n}$ for $i = 1, 2, \ldots, n$.

Let

$$U_i := n^{1-H}(V_{t_i} - V_{t_{i-1}}) = n^{1-H} \int_{t_{i-1}}^{t_i} \sigma(u) dW_u^H, \; i = 1, 2, \cdots, n. \tag{2.20}$$

Consider the Gaussian random variables

$$U_i^* := \sigma(t_i) \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} K^H\left(\frac{[nt]}{n}, s\right) ds \right) \xi_i, \; i = 1, 2, \cdots, n. \tag{2.21}$$

Theorem 2.1 Let $\mathscr{E}_{2,n}$ be the experiment associated with $\{U_i\}_{i=1,2,\ldots,n}$ and $\mathscr{E}_{2,n}^*$ be the experiment associated with $\{U_i^*\}_{i=1,2,\ldots,n}$ defined in (2.20) and (2.21) respectively. Then

$$\lim_{n \to \infty} \Delta(\mathscr{E}_{2,n}, \mathscr{E}_{2,n}^*) = 0.$$

Using the transitivity property of asymptotic equivalence property several times along with inequalities (2.8) and (2.9), we obtain our main results. Our main result of this section is the following:

Theorem 2.2 The experiments $\mathscr{E}_n^Y$ and $\mathscr{E}_n^X$ associated with $\{Y_i\}_{i=1,2,\ldots,n}$ and $\{X_i\}_{i=1,2,\ldots,n}$, defined in (2.3) and (2.4) respectively, are asymptotically equivalent, i.e,

$$\lim_{n\to\infty} \Delta(\mathscr{E}_n^Y, \mathscr{E}_n^X) = 0.$$

Moreover, the drift part of observations $\{Y_i\}_{i=1,2,\ldots,n}$ is asymptotically nonsignificant in this model.

3. Parameter Estimation

Parameter estimation for directly observed stochastic differential equations was studied in Bishwal (2008). Parameter estimation in partially observed stochastic differential equation models which include the stochastic volatility models was studied in Bishwal (2022). Minimum contrast estimation in fractional Ornstein-Uhlenbeck process was studied in Bishwal (2011). The approach in the previous paper uses equivalent martingale representation of the model. Although we have non-semimartingale models, the approach in this section uses random walk construction which allows to use martingale arguments to obtain asymptotic behavior of the estimators.

Drifted Fractional Brownian Motion

Consider the drifted fractional Brownian motion

$$dY_t = \theta dt + dW_t^H, \ \ t \in [0,T] \tag{3.1}$$

where (W_t^H) is a fractional Brownian motion with Hurst parameter $H \in (1/2, 1)$ and the parameter $\theta \in \mathbb{R}$ is unknown which needs to be estimated.

Since the standard Brownian motion is a weak limit of symmetric random walk, replacing the fractional Brownian motion by its associated disturbed (perturbed) random walk, Bertin, Torres and Tudor (2011) considered Euler type discretization of the above model given by

$$Y_{t_{j+1}} = Y_{t_j} + \theta(t_{j+1} - t_j) + B_{t_{j+1}}^{H,N} - B_{t_j}^{H,N}, \ j = 0,1,2,\ldots,N^{\alpha} - 1 \tag{3.2}$$

where $t_j = \frac{j}{N}, j = 0,1,2,\ldots,N^{\alpha}$ and $Y_{t_0} = Y_0 = 0$ and

$$B_t^{H,N} = \sqrt{N}\sum_{i=1}^{[Nt]}\left(\int_{\frac{i-1}{N}}^{\frac{i}{N}} K^H\left(\frac{[Nt]}{N}, s\right) ds\right)\xi_i \tag{3.3}$$

where the ξ_i are i.i.d. standard normal random variables and

$$K^H(t,s) = c_H s^{1/2-H}\int_s^t (u-s)^{H-3/2}u^{H-1/2}du, \ t > s \tag{3.4}$$

where $c_H = \left(\frac{H(2H-1)}{B(2-2H, H-1/2)}\right)^{1/2}$ and $B(\cdot,\cdot)$ is the beta function. Recall that the fBm with $H > 1/2$ can be written as

$$W_t^H = \int_0^t K^H(t,s)dW_s \tag{3.5}$$

where $(W_t, t \geq 0)$ is a standard Wiener process.

Sottinen (2001) proved that the sequence $B^{H,N}$ converges weakly as $N \to \infty$ in Skorohod topology to the fBM. The new model, replacing $B^{H,N}$ with W^H still keeps the main properties of the original process: it is a long range dependent model and the distribution of the noise is asymptotically self-similar. We have a non-semimartingale model, but martingales can be used to treat this model.

Assuming that H is known, Bertin, Torres and Tudor (2011) obtained maximum likelihood estimator (MLE) $\widehat{\theta}_N$ of θ based on discrete observations $\{Y_{t_j}, t_j = \frac{j}{N}, j = 0,1,2,\ldots,N^{\alpha}\}$. Let $\delta = t_{j+1} - t_j$. It is well known that for the drifted standard Brownian motion, the MLE is given by

$$\widehat{\theta}_N = \frac{1}{N\delta}\sum_{j=1}^{N}(Y_j - Y_{j-1}) = \frac{Y_N}{N\delta}.$$

Using Euler type scheme for the fSDE with $t_j = \frac{j}{N}$ in the drifted fractional Brownian motion

$$Y_{t_j} = Y_j = j\frac{\theta}{N} + W^H_{\frac{j}{N}} \tag{3.6}$$

the MLE is given by

$$\widehat{\theta}_N = N\frac{\sum_{i,j=1}^{N^\alpha} j\Gamma_{i,j}^{-1} Y_i}{\sum_{i,j=1}^{N^\alpha} ij\Gamma_{i,j}^{-1}}$$

where $\Gamma_{i,j}^{-1}$ are the coordinates of the matrix Γ^{-1} with $\Gamma = (\Gamma_{i,j})_{i,j=1,\dots,N^\alpha}$ where $\Gamma_{i,j} = Cov\left(B^H_{\frac{i}{N}}, B^H_{\frac{j}{N}}\right)$. They showed that the MLE is unbiased and L^2-consistent if and only if $N\delta \to \infty$ and $N \to \infty$ when $\alpha > 1$.

The MLE has the mean square error bound given by

$$E|\widehat{\theta}_N - \theta|^2 \le cN^{2-2H-3\alpha+\alpha(2H-1)} = CN^{(2-2H)(1-\alpha)}$$

where C is a positive constant. This goes to zero if and only if $\alpha > 1$. Further,

$$E|\widehat{\theta}_N - \theta|^p \le C\left(E|\widehat{\theta}_N - \theta|^2\right)^{p/2} \le CN^{p(1-H)(1-\alpha)}. \tag{3.7}$$

Using Chebychev inequality and Borel-Cantelli lemma, it can be show that $\widehat{\theta}_N \to \theta$ almost surely as $n \to \infty$ for $\alpha > 1$. Thus the MLE is L_P-consistent. Concretely, we want to estimate the drift parameter θ on the basis of observations

$$Y_{t_{j+1}} = Y_{t_j} + \theta(t_{j+1} - t_j) + B^{H,N}_{t_{j+1}} - B^{H,N}_{t_j},\ j = 0,1,2,\dots,N^\alpha - 1,\ Y_0 = 0 \tag{3.8}$$

which can be written as

$$Y_{t_{j+1}} = Y_{t_j} + \frac{\theta}{N} + \sum_{i=1}^{j} f_{ij}\xi_i + F_j\xi_{j+1} \tag{3.9}$$

where

$$F_j = \sqrt{N}\int_{\frac{i-1}{N}}^{\frac{i}{N}} K^H\left(\frac{j+1}{N}, s\right) ds,\quad f_{ij} = \sqrt{N}\int_{\frac{i-1}{N}}^{\frac{i}{N}} \left(K^H\left(\frac{j+1}{N}, s\right) - K^H\left(\frac{j}{N}, s\right)\right) ds \tag{3.10}$$

which can be further written as

$$Y_{j+1} = Y_j + \frac{\theta}{N}(1+\alpha_j) + h_j(Y_1,\dots,Y_j) + F_j\xi_{j+1} \tag{3.11}$$

where the functions h_j and α_j depend on $b_{i,j}$ and $f_{i,j}$.

The α_j satisfy

$$\alpha_0 = 0,\ \alpha_1 = -\frac{f_{11}}{F_0},\ \alpha_j = -\sum_{i=1}^{j} \frac{f_{ij}}{F_{i-1}}(1+\alpha_{i-1}). \tag{3.12}$$

The likelihood function of $Y_1,\dots,Y_j$ can be expressed as

$$L(\theta, y_1,\dots,y_{N^\alpha}) = f_{Y_1}(y_1) f_{Y_2/Y_1}(y_2/y_1)\dots f_{Y_{N^\alpha}/Y_1,\dots,Y_{N^\alpha-1}}(y_{N^\alpha-1}/y_1,\dots,y_{N^\alpha-1}) \tag{3.13}$$

$$= \prod_{j=0}^{N^\alpha-1} \frac{1}{\sqrt{2\pi F_j^2}} \exp\left(-\frac{1}{2}\frac{(y_{j+1} - y_j - h_j(Y_1,\dots,Y_j) - \theta/N(1+\alpha_j))^2}{F_j^2}\right). \tag{3.14}$$

For the discrete model with perturbed random walk, the MLE is given by

$$\widehat{\theta}_N = N\frac{\sum_{j=0}^{N^\alpha-1} \frac{(Y_{j+1} - Y_j - h_j(Y_1,\dots,Y_j))(1+\alpha_j)}{F_j}}{\sum_{j=0}^{N^\alpha-1} \frac{(1+\alpha_j)^2}{F_j^2}} \tag{3.15}$$

with

$$\widehat{\theta}_N - \theta = N\frac{\sum_{j=0}^{N^\alpha-1} \frac{\xi_{j+1}(1+\alpha_j)}{F_j}}{\sum_{j=0}^{N^\alpha-1} \frac{(1+\alpha_j)^2}{F_j^2}}. \tag{3.16}$$

By the independence of $\xi_i,\ i \geq 1$, we have

$$E|\widehat{\theta}_N - \theta|^2 = N^2 E\left(\frac{\sum_{j=0}^{N^\alpha-1}\frac{\xi_{j+1}(1+\alpha_j)}{F_j}}{\sum_{j=0}^{N^\alpha-1}\frac{(1+\alpha_j)^2}{F_j^2}}\right)^2 = N^2 \frac{1}{\sum_{j=0}^{N^\alpha-1}\frac{(1+\alpha_j)^2}{F_j^2}}. \tag{3.17}$$

Next result gives asymptotic normality. Since $\xi_{j+1}, j = 1,2\ldots,N^\alpha - 1$ are i.i.d. $\mathcal{N}(0,1)$, we have

Theorem 3.2

$$\text{a)}\quad N^{(H-1)(1-\alpha)}(\widehat{\theta}_N - \theta) \to^{\mathcal{D}} \mathcal{N}(0, I^{-1}(\theta)) \text{ as } N \to \infty.$$

$$\text{b)}\quad \frac{\sqrt{\sum_{j=0}^{N^\alpha-1}\frac{(1+\alpha_j)^2}{F_j^2}}}{N}(\widehat{\theta}_N - \theta) \to^{\mathcal{D}} \mathcal{N}(0,1) \text{ as } N \to \infty.$$

Remarks. 1. When $H = 3/4$ and $\alpha = 2$, the rate is of the order $N^{1/4}$.
2. We assumed the Hurst parameter H to be known. If H is unknown, it can be estimated as

$$\widehat{H}_N = 1 + \frac{\log \sum_{j=1}^{N}(Y_j - Y_{j-1})^2}{\log N}.$$

We introduce another estimator of θ which we call as the second order approximate maximum likelihood estimator:

$$\widetilde{\theta}_N = N\frac{\sum_{i,j=1}^{N^\alpha} j\Gamma_{i,j}^{-1}Y_i + \sum_{i,j=1}^{N^\alpha} j\Gamma_{i,j}^{-1}Y_{i+1} - H\sum_{i,j=1}^{N^\alpha} ij\Gamma_{i,j}^{-1}}{2\sum_{i,j=1}^{N^\alpha} ij\Gamma_{i,j}^{-1}}.$$

We have

$$\widetilde{\theta}_N - \theta = N\frac{\sum_{i,j=1}^{N^\alpha} j\Gamma_{i,j}^{-1}W_{i/N}^H + \sum_{i,j=1}^{N^\alpha} j\Gamma_{i,j}^{-1}W_{(i+1)/N}^H - H\sum_{i,j=1}^{N^\alpha} ij\Gamma_{i,j}^{-1}}{2\sum_{i,j=1}^{N^\alpha} ij\Gamma_{i,j}^{-1}}.$$

This being a Stratonovich type approximation of the stochastic integral, it converges to the stochastic integral in second order than the Itô type approximation, see Bishwal (2008). Using similar calculations as in Bertin, Torres and Tudor (2011), the mean square error bound is given by

$$E|\widetilde{\theta}_N - \theta|^2 \leq CN^{2-4H-3\alpha+\alpha(2H-1)} = CN^{(2-4H)(1-\alpha)}.$$

We do not need $\alpha > 1$ for the convergence of the right hand side to zero. Further

$$E|\widetilde{\theta}_N - \theta|^p \leq C\left(E|\widehat{\theta}_N - \theta|^2\right)^{p/2} \leq CN^{p(1-2H)(1-\alpha)}.$$

Hence the estimator is L_p consistent for any $p \geq 1$.
Using Chebychev inequality and Borel-Cantelli lemma, it can be show that $\widehat{\theta}_N \to \theta$ almost surely as $n \to \infty$.

Theorem 3.3

$$\text{a)}\quad N^{(2H-1)(1-\alpha)}(\widetilde{\theta}_N - \theta) \to^{\mathcal{D}} \mathcal{N}(0, I^{-1}(\theta)) \text{ as } N \to \infty.$$

$$\text{b)}\quad \frac{\sqrt{\sum_{j=0}^{N^\alpha-1}\frac{(1+\alpha_j)^2}{F_j^2}}}{N}(\widetilde{\theta}_N - \theta) \to^{\mathcal{D}} \mathcal{N}(0,1) \text{ as } N \to \infty.$$

Fractional Ornstein-Uhlenbeck Process

Then we consider fractional Ornstein-Uhlenbeck process.

$$dY_t = \theta Y_t dt + dW_t^H,\quad t \in [0,T] \tag{3.18}$$

where $\theta < 0$. Replacing the fractional Brownian motion by its associated disturbed random walk, the Euler discretization of the above model is given by

$$Y_{t_{j+1}} = Y_{t_j} + \theta Y_{t_j}(t_{j+1} - t_j) + B^{H,N}_{t_{j+1}} - B^{H,N}_{t_j}, \;\; j = 1, 2, \ldots, N^{\alpha} \tag{3.19}$$

The MLE is given by

$$\widehat{\theta}_N = \frac{1}{\delta} \log N \frac{\sum_{i,j}^{N^\alpha} j\Gamma_{i,j}^{-1} Y_i Y_j}{\sum_{i,j}^{N^\alpha} ij\Gamma_{i,j}^{-1} Y_i^2}.$$

Theorem 3.4

$$N^{(H-1)(1-\alpha)}(\widehat{\theta}_N - \theta) \to^D \mathscr{N}(0, I^{-1}(\theta)) \text{ as } N \to \infty.$$

We can improve this estimator.

$$\widetilde{\theta}_N = N \frac{\sum_{i,j}^{N^\alpha} j\Gamma_{i,j}^{-1} Y_i Y_j + \sum_{i+1,j+1}^{N^\alpha} j\Gamma_{i,j}^{-1} Y_i Y_j - H\delta \sum_{i,j}^{N^\alpha} ij\Gamma_{i,j}^{-1} Y_i^2}{2\delta \sum_{i,j}^{N^\alpha} ij\Gamma_{i,j}^{-1} Y_i^2}.$$

Theorem 3.5

$$N^{(2H-1)(1-\alpha)}(\widetilde{\theta}_N - \theta) \to^D \mathscr{N}(0, I^{-1}(\theta)) \text{ as } N \to \infty.$$

The advantage of the estimator $\widetilde{\theta}_N$ over $\widehat{\theta}_N$ is that the sampling restriction $\alpha > 1$ is not required for consistency and asymptotic normality. Due to symmetry, the convergence rate is faster.

On the stochastic basis the fractional Ornstein-Uhlenbeck process X_t is defined satisfying the Itô stochastic differential equation

$$dX_t = \theta X_t dt + dW_t^H, \;\; t \geq 0, \;\; X_0 = \xi \tag{3.20}$$

where $\{W_t^H\}$ is a fractional Brownian motion with $H > 1/2$ with the filtration $\{\mathscr{F}_t\}_{t\geq 0}$ and $\theta < 0$ is the unknown parameter.

Davydov (1970) gave an AR(1) approximation of the fOU model:

$$y_j = \rho y_{j-1} + v_j, (1-L)^{H-1/2} v_j = \varepsilon_j, \;\; \rho = 1, \; y_0 = 0, \; j = 1, 2 \ldots, T \tag{3.21}$$

where L is the lag operator, $\varepsilon_j \sim$ i.i.d. $(0, \sigma^2)$ with $E(\varepsilon_j^4) < \infty$, whereas $\{v_j\}$ is a stationary long-memory process generated by

$$v_j = (1-L)^{H-1/2} \varepsilon_j = \sum_{k=0}^{\infty} \frac{\Gamma(k + H - 1/2)}{\Gamma(H - 1/2)\Gamma(k+1)} \varepsilon_{j-k}. \tag{3.22}$$

Davydov (1970) proved that

$$\frac{\lambda_H^{1/2}}{\sigma T^H} y_{[Tt]} \to^{\mathscr{D}} W_t^H \;\; \text{as} \;\; T \to \infty.$$

Now consider the fCIR model

$$dX_t = a(b - X_t)dt + \sigma\sqrt{X_t} dW_t^H, \tag{3.23}$$

Then by Proposition 5.7 of Buchmann and Kluppelberg (2006), we have $X_t = f(Y_t)$ where

$$dY_t = a(b - Y_t)dt + dW_t^H, \;\; Y_0 = f^{-1}(X_0), \;\; t \in [0, T] \tag{3.24}$$

and $f(x) = sgn(x)\sigma^2 x^2/4$.

Fractional Cox-Ingersoll-Ross Process

Feller (1951) reached at the square-root process as the weak limit of Galton-Watson branching process with immigration while studying a problem in genetics. Using the Feller's square-root process, Cox *et al.* (1985) studied the theory of term structure of interest rates and the model is now known as the Cox-Ingersoll-Ross model. Overbeck and Ryden (1997) studied asymptotics of conditional least squares esimtimators of Cox-Ingersoll-Ross process from discrete observations using an auto-regressive type representation of the model with non-Gaussian error. Dehtiar *et al.* (2021) studied strong consistency for the maximum likelihood method and an alternative method of estimation of the drift parameters of the Cox-Ingersoll-Ross process based on continuous observations. Mishura and Yurchenko-Tytarenko (2018) studied hitting probability of fractional Cox-Ingersoll-Ross model which involves long memory. Mackevicius

(2015) used stochastic Verhulst model as an alternative to CIR model for modeling interest rate as both processes have similar behavior. Mackevicius (2011) studied weak approximation of CIR equation by discrete random variables. Lenkasas and Mackevicius (2015) obtained a second order weak approximation of Heston model by discrete random variables. Lileika and Mackevicius (2020) studied weak approximation of CKLS and CEV process (cf. Cox (1996)) by discrete random variables.

We consider fractional Cox-Ingersoll-Ross process.

$$dY_t = \theta Y_t dt + \sqrt{Y_t} dW_t^H, \quad t \in [0,T] \tag{3.25}$$

where $\theta < 0$. Replacing the fractional Brownian motion by its associated disturbed random walk, the Euler discretization of the above model is given by

$$Y_{t_{j+1}} = Y_{t_j} + \theta Y_{t_j}(t_{j+1} - t_j) + \sqrt{Y_{t_j}}(B^{H,N}_{t_{j+1}} - B^{H,N}_{t_j}), \ j = 1, 2, \ldots, N^\alpha. \tag{3.26}$$

The MLE is given by

$$\widehat{\theta}_N = \frac{1}{\delta} \log N \frac{\sum_{i,j}^{N^\alpha} j\Gamma_{i,j}^{-1} Y_i \sum_{i,j}^{N^\alpha} j\Gamma_{i,j}^{-1} Y_j^{-1}}{\sum_{i,j}^{N^\alpha} ij\Gamma_{i,j}^{-1} Y_i^2}. \tag{3.27}$$

We have

Theorem 3.6

$$N^{(H-1)(1-\alpha)}(\widehat{\theta}_N - \theta) \to^{\mathscr{D}} \mathscr{N}(0, I^{-1}(\theta)) \text{ as } N \to \infty$$

where $I(\theta)$ is the Fisher information of the model.

We can improve this estimator in terms of lower variance in simulation.

$$\widetilde{\theta}_N = N \frac{\sum_{i,j}^{N^\alpha} j\Gamma_{i,j}^{-1} Y_i \sum_{i,j}^{N^\alpha} j\Gamma_{i,j}^{-1} Y_j^{-1} + \sum_{i+1,j+1}^{N^\alpha} j\Gamma_{i,j}^{-1} Y_i \sum_{i+1,j+1}^{N^\alpha} j\Gamma_{i,j}^{-1} Y_j^{-1}}{\delta(\sum_{i,j}^{N^\alpha} ij\Gamma_{i,j}^{-1} Y_i^2 + \sum_{i+1,j+1}^{N^\alpha} ij\Gamma_{i,j}^{-1} Y_i^2)}$$
$$-HN \frac{(\sum_{i,j}^{N^\alpha} ij\Gamma_{i,j}^{-1} Y_i^2 + \sum_{i+1,j+1}^{N^\alpha} ij\Gamma_{i,j}^{-1} Y_i^2)}{\delta(\sum_{i,j}^{N^\alpha} ij\Gamma_{i,j}^{-1} Y_i^2 + \sum_{i+1,j+1}^{N^\alpha} ij\Gamma_{i,j}^{-1} Y_i^2)}. \tag{3.28}$$

We have

Theorem 3.7

$$N^{(2H-1)(1-\alpha)}(\widetilde{\theta}_N - \theta) \to^{\mathscr{D}} \mathscr{N}(0, I^{-1}(\theta)) \text{ as } N \to \infty$$

where $I(\theta)$ is the Fisher information of the model.

4. Testing the Fractional Volatility

We study the problem of testing of the parametric form of the volatility in a stochastic differential equation driven by fractional Brownian motion with Hurst parameter $H > \frac{1}{2}$. This class of models is important as it captures the long memory behavior of the log-share prices. We then extend to fractional stochastic volatility model.

Inference about volatility is an important problem in derivative pricing. A misspecification of the volatility function could lead to misspecfied derivative prices. When the price process follows the classical diffusion, Ait-Sahalia (1996a) studied the volatility testing problem using nonparametric density matching. His test is based on the comparison of parametric marginal density with its nonparametric estimator. However, the classical diffusion does not take in to account the long memory behavior of the price process. In the current work, we take the long memory behavior of the price in to account using the fractional diffusion and study the testing problem for volatility. We study the asymptotic behavior of the test statistic. Nonparametric pricing of interest rate derivatives studied in Ait-Sahalia (1996b). Basically the pricing PDE is solved by replacing the volatility by its nonparametric kernel estimator which is pointwise consistent and asymptotically normal. Due to the relationship between the diffusion coefficient and the drift coeffcient, once the drift parameter has been identified, the diffusion function can be identified from the marginal distribution. The drift parameters are estimated by OLS which are plugged in to the kernel estimator of the marginal density which are then further plugged into the diffusion-drift equation to obtain the nonparametric estimator of the diffusion coefficient. The asymptotic distribution of the nonparametric estimator of the price is normal.

We assume that the price process of an underlying is a one dimensional fractional diffusion satisfying the fractional Itô SDE

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t^H,\ 0 < t \leq T \tag{4.1}$$

where μ and σ are smooth functions such that a unique solution of the above SDE exists and $\{W_t^H, t \in [0,T]\}$ is a fractional Brownian motion with Hurst coefficient $H > 1/2$. Assume the σ has continuous derivatives up to second order.

For the pricing of contingent claims with a payoff function that depends on the evolution of the underlying process X during the time period $[0,T]$, the derivative prices are calculated under the risk-neutral quasi-martingale measure, which is not influenced by the drift function, but the diffusion function σ captures the volatility of the underlying.

In the field of interest rate modeling, the specification of μ is also important. However, a first step in evaluating a particular parametric interest rate model could be test of parametric form of its volatility coefficient.

It is often assumed that the diffusion coefficient belongs to a set of parametric functions , i.e, there exists an unknown parameter $\theta_0 \in \Theta$ such that $\sigma(x) = \sigma(\theta_0, x)$.

The null and the alternative hypotheses to be tested are:
$\mathscr{H}_0$: There exists $\theta_0 \in \Theta$ such that for every $t \in [0,T]:\ \sigma(X_t) = \sigma(\theta_0, X_t)$ P-a.s.
$\mathscr{H}_1$: For all $\theta \in \Theta$ and for every $t \in [0,T] : |\sigma(X_t) - \sigma(\theta, X_t)| \geq c_n\ d_n(X_t)$ P-a.s.

Here d_n is the local shift in the alternative, a sequence of bounded functions and c_n is the order of difference between $\mathscr{H}_0$ and $\mathscr{H}_1$, which is a deterministic sequence converging to zero at a proper rate as $n \to \infty$.

We need the following assumptions in the sequel:
(A1) The following condition holds for σ^2: $|\sigma^2(\theta,x) - \sigma^2(\theta_0,x)| \leq d(x)|\theta - \theta_0| \quad \forall x \in I_x$ where $d(x)$ is a function depending on x and the set I_x is defined by $I_x := \{x : L_T(x) \geq \varepsilon > 0\}$ with an arbitrary number ε, where $L_T(x)$ denotes the local time of X at time T.

Local time of Y in x during $[0,t]$ is defined as

$$L_t(x) = \lim_{\delta\to 0} \frac{1}{\delta} \int_0^t I_{\{|Y_s - x| < \delta\}} ds \tag{4.2}$$

and its discrete approximation is defined as

$$L_{n,t}(x) = \frac{1}{2nh} \sum_{i=1}^n I_{\{|Y_{t_{i-1}} - x| < \delta\}},\ \ L_{n,T}(x) = \sum_{i=1}^n I_{x,h}(y) = \sum_{i=1}^n I_{\left(\frac{|y-x|}{h} < 1\right)}. \tag{4.3}$$

The local time of the Dirichlet process r at a point a over the time interval $[0,t]$ the amount of time spent by the process near a and is defined as

$$l(t,a) = \lim_{\varepsilon\to 0} \frac{1}{2\varepsilon} \int_0^t I_{\{|r_s - a| < \varepsilon\}} \sigma^2(r_s) ds. \tag{4.4}$$

The discrete approximation of $l(t,a)$ is given by

$$l_n(t,a) = \frac{1}{2n\Delta_n} \sum_{i=1}^{n-1} I_{\{|r_{t_i} - a| < \Delta_n\}} \sigma^2(r_{t_i})(t_{i+1} - t_i). \tag{4.5}$$

Occupation Time Formula: For every Borel f

$$\int_0^t f(r_s) d\langle r, r\rangle = \int_{-\infty}^{\infty} f(a) l(t,a) da. \tag{4.6}$$

The kernel K has compact support and is symmetric about 0 and is continuously differentiable.

(A2) $\widehat{\theta}_n$ is a $1-H$ root consistent parametric estimator of θ within the family of the parametric model, i.e., $|\widehat{\theta}_n - \theta| = O_P(n^{-(1-H)})$.

Let T_x be the first hitting time of x, i.e., the first time the point x is visited. For all x satisfying $T_x < 1$, we introduce the following test statistic:

$$\mathscr{T}_n^H(x) := \frac{1}{2} \{nh_n L_{n,T}(x)\}^{2(1-H)} \left(\frac{S_n^2(x)}{\widetilde{\sigma}^2(\widehat{\theta}_n, x)} - 1 \right)^2 \tag{4.7}$$

where

$$S_n^2(x) = \frac{\sum_{i=1}^{n-1} I_{\{|X_{t_i} - a| < \Delta_n\}} n(X_{t_{i+1}} - X_{t_i})^2}{\sum_{i=1}^{n-1} I_{\{|X_{t_i} - a| < \Delta_n\}}}.$$

The proposed test statistic is asymptotically equivalent to the L_2-distance between $S_n(\cdot)$ and $\widetilde{\sigma}^2(\widehat{\theta}_n,\cdot)$. We will study the behavior of the above test statistic for large sample size. In the following section, we introduce fractional stochastic calculus briefly and the material needed to prove our main results.

We need the fractional Itô formula: if $f:\mathbb{R}\to\mathbb{R}$ is a twice continuously differentiable function with bounded derivatives of order two, then

$$f(W_T^H)-f(W_0^H)=\int_0^T f'(W_s^H)dW_s^H+H\int_0^T s^{2H-1}f''(W_s^H)ds \ \ a.s.$$

For $H=\frac{1}{2}$, it gives the classical Itô formula for standard Brownian motion.

We also need the local time theory for fractional Brownian motion. The local time of fractional Brownian motion is defined as

$$l_t^{(W^H)}(x)=\int_0^t \delta(W_s^H-x)ds=\lim_{\varepsilon\to 0}\left|\frac{1}{2\varepsilon}\left\{s\in[0,t]:|W_s^H-x|<\varepsilon\right\}\right| \tag{4.8}$$

where δ is the Dirac delta function and $|A|$ denotes the Lebesgue measure of the Borel set A.

The weighted local time of a fractional Brownian motion process is studied in Hu (2001), Hu and Oksendal (1999), Hu, Oksendal and Salopek (2001). The weighted local time is defined as

$$L_t^{(W^H)}(x)=\int_0^t \delta(W_s^H-x)s^{2H-1}ds \tag{4.9}$$

where δ is the Dirac delta function. For $H=1/2$, the usual local time is the same as the weighted local time. We have

$$E\left[l_t^{(W^H)}(x)\right]=(2\pi)^{-1/2}\int_0^t r^{-H}\exp(-\frac{1}{2}r^{-2H}x^2)dr, \tag{4.10}$$

$$E\left[L_t^{(W^H)}(x)\right]=(2\pi)^{-1/2}\int_0^t r^{H-1}\exp(-\frac{1}{2}r^{-2H}x^2)dr. \tag{4.11}$$

Further,

$$E\left[L_t^{(W^H)}(x)\right]^2\leq\frac{\Gamma(H)\Gamma(1-H)t^{2H}}{2H\pi\sqrt{k}} \tag{4.12}$$

where k is a constant defined below:

For $0\leq s<t\leq T$, there is a constant $k>0$ such that

$$\mathrm{Var}(\xi W_t^H-\eta W_s^H)=\mathrm{Var}(\xi(W_t^H-W_s^H)+(\xi-\eta)W_s^H)\geq k[\xi^2|t-s|^{2H}+(\eta-\xi)^2s^{2H}]. \tag{4.13}$$

The local time and the weighted local time are jointly continuous t and x and for almost all $\omega\in\Omega$.

The Meyer-Tanaka formula for fractional Brownian motion is given by

$$|W_t^H-z|=|z|+\int_0^t \mathrm{sign}(W_s^H-z)dW_s^H+2HL^{W^H}(z). \tag{4.14}$$

For the fractional Black-Scholes model

$$dX_t=\mu X_t dt+\sigma X_t dW_t^H,\ t\geq 0,\ X_0=x>0 \tag{4.15}$$

whose solution the fractional geometric Brownin motion is

$$X_t=x\exp(\sigma W_t^H+\mu t-\frac{1}{2}\sigma^2t^{2H}) \tag{4.16}$$

the following are the moments of the weighted local time:

$$E\left[L_t^X(z)\right]=\frac{1}{\sqrt{2\pi}\sigma z}\int_0^t s^{H-1}\exp\left(-\frac{h(s,z)^2}{2s^{2H}}\right)ds, \tag{4.17}$$

where

$$h(s,z)=\frac{1}{\sigma}\left[\log\frac{z}{x}-\mu t+\frac{\sigma^2}{2}t^{2H-1}\right]. \tag{4.18}$$

Further

$$E\left[L_t^X(z)\right]^2 \leq \frac{\Gamma(H)\Gamma(1-H)t^{2H}}{2H\pi\sqrt{k}\sigma^2 z^2}. \tag{4.19}$$

We present the Itô-Tanaka formula: Let $f:\mathbb{R}^+ \to \mathbb{R}$ be a convex function of polynomial growth. Then

$$f(X_t) = f(X_0) + \int_0^t D^- f(X_s)dX_s + \sigma^2 H L_t(x) d\nu_f(x) \tag{4.20}$$

where $D^- f$ denotes the left derivative of f and ν_f denotes the second derivative of the measures of f. The Meyer-Tanaka formula for fractional geometric Brownian motion is given by

$$|X_t - z| = |X_0 - z| + \int_0^t \text{sign}(X_s - z)dX_s + 2\sigma^2 H z^2 L_t^X(z) \tag{4.21}$$

for any $z > 0$.

Also,

$$(X_t - z)^+ = (X_0 - z)^+ + \int_0^t \chi_{[z,\infty]}(X_s)dX_s + \sigma^2 H z^2 L_t^X(z) \tag{4.22}$$

for any $z > 0$.

Theorem 4.1 Under $\mathscr{H}_0$, the test statistic $\mathscr{T}_n^H(x)$ converges in distribution to a χ^2- random variable with k degrees of freedom as $n \to \infty, nh \to \infty, (nh^2)^{-1}\log n \to 0$ and $nh^3 \to \infty$.

Empirical Likelihood Test About the Fractional Volatility

The main advantage of empirical likelihood methods is their ability to studentize internally and to correct test statistics and confidence intervals for empirical properties of data.

We obtain that $\mathscr{T}_n^H(x)$ converges under the alternative hypothesis to a noncentral chi-square distribution with k degrees of freedom and non-centrality parameter $\sum_{i=1}^k \{d_n^2(x_i)/\sigma^4(\theta_0, x_i)\}$.

Theorem 4.2 Under $\mathscr{H}_1$, $\mathscr{T}_n^H(x)$ converges in distribution to a non-central χ^2- random variable with k degrees of freedom and non-centrality parameter $\sum_{l=1}^k \{d_n^2(x_l)/\sigma^4(\theta_0, x_l)\}$ as $n \to \infty, nh \to \infty, (nh^2)^{-1}\log n \to 0$ and $nh^3 \to \infty$.

5. Semiparametric Estimation in Fractional FBSDE

El Karoui and Quenez (1995) stated that prices of many important derivative securities could be solved by backward stochastic differential equation (BSDE). When the randomness of the initial condition comes from a forward equation, then one gets the forward-backward stochastic differential equation (FBSDE). Su and Lin (2009) studied semi-parametric estimation of forward-backward stochastic differential equations. Lin *et al.* (2011) proposed FBSDE for an ecological problem and studied statistical inference. Zhang and Lin (2014) studied parameter estimation of FBSDEs depending on a terminal condition. Chen and Lin (2015) studied nonparametric estimation of FBSDEs models with applications in finance.

Consider the fractional forward-backward stochastic differential equation (FFBSDE)

$$dX_t = -g(Y_t, X_t, Z_t)dt + Z_t dW_t,\ X_T = \xi(Y_T),$$

$$dY_t = \mu(Y_t)\,dt + \sigma(Y_t)\,dW_t^H, \quad t \in [0,T] \tag{5.1}$$

with the forward component Y_t and backward component X_t being observed. The other backward process Z_t is totally unobserved and the function g can not be completely specified in usual financial market. However, Z_t can be viewed as a nonparametric approximation of the observable variables. Compared to ordinary SDE that contains an initial condition, the solution to FBSDE is affected by the terminal condition $X_T = \xi(Y_T)$. Using the terminal condition into account, Bishwal (2010) studied MLE in anticipative SDE. There exist a number of parametric and nonparametric methods of estimation in ordinary SDEs. However, thse method can not be directly applied to BSDE and FBSDE because they are both related to terminal conditions.

We study semiparametric estimation of g and the process Z based on observations (X_t, Y_t). The process Z_t is also the coefficient in the diffusion term, for example, within a complete market, it serves to characterise the dynamic value of a replicating portfolio X_t with a final wealth ξ and a special quantity Z_t that depends on the hedging portfolio. We assume $g = g(\beta, t, Y_t, Z_t)$ where g is a given function and β is an unknown parameter vector. Let $\beta := (a, b, c)^T$. We consider $g(\beta, t, Y_t, Z_t) = a + bY_t + cZ_t$.

For $H = 0.5$, for a fully observed model, Bandi and Phillips (2003) proposed the following two estimators when Y is observed:

$$\widehat{\mu}_{n,T}(x) = \frac{\sum_{i=1}^n K_{x,h}(Y_{t_{i-1}})\Delta Y_{t_i}}{\sum_{i=1}^n K_{x,h}(Y_{t_{i-1}})\Delta t_i}, \tag{5.2}$$

$$\widehat{\sigma}^2_{n,T}(x) = \frac{\sum_{i=1}^n K_{x,h}(Y_{t_{i-1}})(\Delta Y_{t_i})^2}{\sum_{i=1}^n K_{x,h}(Y_{t_{i-1}})\Delta t_i} \tag{5.3}$$

where the kernel

$$K_{x,h}(y) := K\left(\frac{y-x}{h}\right).$$

We propose another estimator of μ as

$$\widetilde{\mu}_{n,T}(x) = \frac{\sum_{i=1}^n [K_{x,h}(Y_{t_{i-1}}) + K_{x,h}(Y_{t_i})]\Delta Y_{t_i} - H\sum_{i=1}^n K_{x,h}(Y_{t_{i-1}})\Delta t_i}{2\sum_{i=1}^n K_{x,h}(Y_{t_{i-1}})\Delta t_i}. \tag{5.4}$$

We consider the bivariate process (X, Y). Bishwal and Pena (2009) studied inference for bivariate normal distribution model. Observe that

$$Z_t^2 = \frac{1}{\delta}E((X_{s+\delta} - X_s)^2 | Y_t) + O_P(\delta).$$

The Nadaraya-Watson estimator of Z^2 is given by

$$\widehat{Z}^2_{n,T}(x) = \frac{\sum_{i=1}^n K_{x,h}(Y_{t_{i-1}})(\Delta X_{t_i})^2}{\sum_{i=1}^n K_{x,h}(Y_{t_{i-1}})\Delta t_i}. \tag{5.5}$$

We have the following assumptions about the kernel. Consider a continuously differentiable kernel K with shrinking bandwidth $h \to 0$. Let

$$K_h(x) := \frac{1}{h}K\left(\frac{x}{h}\right) \tag{5.6}$$

where $K : \mathbb{R} \to \mathbb{R}$ is a kernel which normalizes to

$$\int_{\mathbb{R}} K(x)dx = 1. \tag{5.7}$$

For example, consider the Epanechikov kernel

$$K(x) = \begin{cases} \frac{3}{4}(1-x)^2 & : \quad |x| \le 1 \\ 0 & : \quad \text{otherwise} \end{cases} \tag{5.9}$$

and the kernel suggested by Zhang and Karunamani (1998)

$$K(x) = \begin{cases} 6(1+3x+2x^2) & : \quad -1 \le x \le 0 \\ 0 & : \quad \text{otherwise} \end{cases} \tag{5.9}$$

The kernel estimator converges to the integrated variance as the bandwith h vanishes. In order to improve the rate of convergence of kernel estimators, we consider symmetric kernel estimators.

For simplicity of notation, we will denote $K_h((t-T)) =: K(t)$. Recall that the integrated volatility is estimated by the realized volatility on the basis of discrete observations of the process $\{X_t\}$ at times $0 = t_0 < t_1 < \cdots t_n = T$ with $t_i - t_{i-1} = \frac{T}{n}, i = 1, 2 \cdots, n$. We discretize (5.1) as

$$X_{t_i} - X_{t_{i-1}} = -g(Y_{t_i}, X_{t_i}, Z_{t_i})\delta + m(Y_{t_i}, \xi) + u_{t_i}$$

where $m(Y_{t_i}, \xi) := E[Z_{t_i}(W_{t_i} - W_{t_{i-1}})|Y_{t_i}, \xi] \ne 0$ and $u_{t_i} := Z_{t_i}(W_{t_i} - W_{t_{i-1}}) - m(Y_{t_i}, \xi)$. The process u_{t_i} can be regarded as the error term with $E(u_{t_i}|Y_{t_i}, \xi) = 0$. The terminal condition ξ enters into the formula as terminal control variable. This plays the similar role of instrumental variable (IV). It is well known that IV is widely used to estimate causal relationships when controlled experiments are not feasible in statistics. So the terminal control variable can be called quasi-IV.

Suppose for each terminal data ξ_j and equally spaced time points $\{t_i = t_1 + (i-1)\Delta, i = 1,2,\ldots,n\} \subseteq [0,T]$, we record the observed time series data $\{X_{i,j}, Y_{i,j}, i = 1,2,\ldots,n, j = 1,2,\ldots,m\} = \{X_{t_i,j}, Y_{t_i,j}, i = 1,2,\ldots,n, j = 1,2,\ldots,m\}$. Denote $\Delta_i = t_{i+1} - t_i, i = 1,2,\ldots,n$.

Nadaraya-Watson estimator of m valued at (y_0, ξ_0) is given by

$$\widehat{m}(y_0,\xi_0) = \frac{\sum_{i=1}^{n-1}\sum_{j=1}^{m}(X_{i+1,j} - X_{i,j})K_{x,h}(Y_{i,j} - y_0)K_{\xi,h}(\xi_j - \xi_0)}{\sum_{i=1}^{n-1}\sum_{j=1}^{m}K_{x,h}(Y_{i,y} - y_0)K_{\xi,h}(\xi_j - \xi_0)}.$$

After replacing Z and m with the estimators, we can derive the LSE of β by minimizing

$$\sum_{i=1}^{n-1}\sum_{j=1}^{m}\left(X_{i+1,j} - X_{i,j} - (a + bX_{i,j} + c\widehat{Z}_{i,j})\Delta - \widehat{m}(Y_{i,j},\xi_j)\right)^2.$$

Denote

$$DX_{t_{i-1}}^2 := (X_{t_i} - X_{t_{i-1}})^2. \tag{5.10}$$

The realized squared volatility is defined as

$$\widehat{V}_{n,T} := \sum_{i=1}^{n} DX_{t_{i-1}}^2. \tag{5.11}$$

Note that

$$\text{P-}\lim_{n\to\infty}\widehat{V}_{n,T} = V_T. \tag{5.12}$$

We have

Theorem 5.1 As $n \to \infty$, $nh \to \infty$, $nh^5 \to 0$, $nh\Delta^2 \to 0$,

$$(nh)^{1-H}(\widehat{Z}_{n,T}^2(y_0) - Z_{n,T}^2(y_0)) \to^{\mathscr{D}} \mathscr{N}\left(0, \frac{Z^4(y_0)J_K}{p(y_0)}\right)$$

with $J_K = \int_0^1 K^2(u)du$ where the continuous kernel function $K(\cdot)$ is symmetric about 0, with a support of interval $[-1,1]$, and $\int_{-1}^1 K(u)du = 1, \int_{-1}^1 u^2K(u)du \neq 0, \int_{-1}^1 |u|^jK^k(u)du < \infty, j \leq k = 1,2$ and $p(\cdot)$ is the probability density of $Y_i, i = 1,\ldots,n$ which has continuous two derivatives in a neighborhood of y_0.

Since the distribution of ξ is supposed to be known, one can get the sample $\{\xi_i, 1 \leq i \leq k\}$ for $k > 1/\delta$ and the estimate $\widehat{\xi} = \frac{1}{k}\sum_{i=1}^k \xi_i$. Next, we obtain the conditional least squares estimator (CLSE) of β.

Minimizing

$$\sum_{i=1}^{n-1}\left(X_{t_i} - \widehat{\xi} - \sum_{j=i+1}^{n}(a + bX_{t_j} + c\widehat{Z}_{t_j})\Delta_j\right)^2$$

produces the conditional least squares estimator

$$\widehat{\beta}_n = (\widehat{U}^{\top}\widehat{U})^{-1}\widehat{U}^T V \tag{5.13}$$

where

$$V = \frac{1}{T}(X_{t_i} - \widehat{\xi}),$$

$$U = \frac{1}{T}\left(\sum_{j=i+1}^{n}\Delta_j, \sum_{j=i+1}^{n}X_{t_j}\Delta_j, \sum_{j=i+1}^{n}Z_{t_j}\Delta_j\right), \quad \widehat{U} = \frac{1}{T}\left(\sum_{j=i+1}^{n}\Delta_j, \sum_{j=i+1}^{n}X_{t_j}\Delta_j, \sum_{j=i+1}^{n}\widehat{Z}_{t_j}\Delta_j\right).$$

We have

Theorem 5.2

$$\text{(a) } \widehat{\beta}_n \to^P \beta \quad \text{as} \quad n \to \infty,$$

$$\text{(b) } n^{1-H}(\widehat{\beta}_n-\beta)\rightarrow^{\mathscr{D}}\mathscr{N}(0,I^{-1}(\beta)) \text{ as } n\rightarrow\infty$$

where $I(\beta)$ is the Fisher information of the model.

We can improve this estimator using symmetric distribution. Another Nadaraya-Watson estimator of Z^2 is given by

$$\widetilde{Z}^2_{n,T}(x)=\frac{\sum_{i=1}^n[K_{x,h}(Y_{t_{i-1}})+K_{x,h}(Y_{t_i})](\Delta X_{t_i})^2-H\sum_{i=1}^n[K'_{x,h}(Y_{t_{i-1}})+K'_{x,h}(Y_{t_i})]\Delta t_i}{\sum_{i=1}^n[K_{x,h}(Y_{t_{i-1}})+K_{x,h}(Y_{t_i})]\Delta t_i}. \tag{5.14}$$

We have

Theorem 5.3 As $n\rightarrow\infty, nh\rightarrow\infty, nh^5\rightarrow 0, nh\Delta^2\rightarrow 0$,

$$(nh)^{1-H}(\widetilde{Z}^2_{n,T}(y_0)-Z^2_{n,T}(y_0))\rightarrow^{\mathscr{D}}\mathscr{N}\left(0,\frac{Z^4(y_0)J_K}{p(y_0)}\right)$$

with $J_K=\int_0^1K^2(u)du$ where the continuous kernel function $K(\cdot)$ is symmetric about 0, with a support of interval $[-1,1]$, and $\int_{-1}^1K(u)du=1, \int_{-1}^1u^2K(u)du\neq 0, \int_{-1}^1|u|^jK^k(u)du<\infty, j\leq k=1,2$ and $p(\cdot)$ is the probability density of $Y_i, i=1,\ldots,n$ which has continuous two derivatives in a neighborhood of y_0.

Minimizing

$$\sum_{i=1}^{n-1}\left(X_{t_i}-\widehat{\xi}-\sum_{j=i+1}^{n}((a+bX_{t_j}+c\widetilde{Z}_{t_j})\Delta t_j\right)^2$$

produces a more efficient conditional least squares estimator (CLSE)

$$\widetilde{\beta}_n=(\widetilde{U}^{\mathsf{T}}\widetilde{U})^{-1}\widetilde{U}^TV \tag{5.15}$$

where

$$V=\frac{1}{T}(X_{t_i}-\widehat{\xi}),$$

$$U=\frac{1}{T}\left(\sum_{j=i+1}^n\Delta_j,\ \sum_{j=i+1}^nX_{t_j}\Delta_j,\ \sum_{j=i+1}^nZ_{t_j}\Delta_j\right),\quad \widetilde{U}=\frac{1}{T}\left(\sum_{j=i+1}^n\Delta_j,\ \sum_{j=i+1}^nX_{t_j}\Delta_j,\ \sum_{j=i+1}^n\widetilde{Z}_{t_j}\Delta_j\right).$$

We have

Theorem 5.4

$$\text{(a) } \widetilde{\beta}_n\rightarrow^P\beta \text{ as } n\rightarrow\infty,$$

$$\text{(b) } n^{1-H}(\widetilde{\beta}_n-\beta)\rightarrow^{\mathscr{D}}\mathscr{N}(0,I^{-1}(\beta)) \text{ as } n\rightarrow\infty$$

where $I(\beta)$ is the Fisher information of the model.

Example Consider the FBSDEs

$$dX_t=[rX_t+(\mu-r)\sigma^{-1}Z_t]dt+Z_tdW_t,\ \ X_T=(Y_T-K)^+ \tag{5.16}$$

$$dY_t=\mu Y_tdt+\sigma Y_tdW_t^H,\ Y_0=y. \tag{5.17}$$

We denote

$$dX_t=(bX_t+cZ_t)dt+Z_tdW_t. \tag{5.18}$$

The process Y_t is the FGBM. With $e^{-rT}X_T=e^{-rT}(Y_T-K)^+$ being the payoff of the call option with strike price K, the call option price is given by

$$C_0=e^{-rT}\int_{\mathbb{R}}\frac{1}{\sqrt{2\pi}T^H}\left(x\exp\left[\sigma y+rT-\frac{1}{2}\sigma^2T^{2H}\right]-K\right)^+\exp\left[\frac{y^2}{2T^{2H}}\right]dy. \tag{5.19}$$

Recall that Meyer-Tanaka formula gives

$$(Y_T-K)^+=(Y_0-K)^++\int_0^TI_{[K,\infty)}(Y_s)dY_s+\sigma^2HK^2L_T^Y(K) \tag{5.20}$$

Hence an alternative call valuation formula using Meyer-Tanaka formula is

$$C_0 = (Y_0 - K)^+ + \sigma^2 H K^2 \widehat{E}[L_T^Y(K)] \tag{5.21}$$

where $\widehat{E}$ denotes the expectation under risk neutral probability measure defined in the fractional Girsanov Theorem 1.1, the expectation of the middle term is zero and the mean local time can be calculated from (4.17). $(Y_0 - K)^+$ is called the intrinsic value of the option and $\sigma^2 H K^2 \widehat{E}[L_T^Y(K)]$ is called the time value of the option. As $H \to 1/2$, we get the price process of a geometric Brownian motion.

Remarks Recently Ichiba *et al.* (2021, 2022) studied generalized fractional Brownian motion (GFBM). A generalized fractional Brownian motion is a Gaussian self-similar process whose increments are not necessarily stationary. It appears in the scaling limit of a shot-noise process with a power law shape function and non-stationary noises with a power law variance function. They studied semimartingale properties of the mixed process made up of an independent Brownian motion and a GFBM for the persistent Hurst parameter. It would be interesting to extend the current paper to GFBM noise.

Conclusion Long memory in volatility is an important stylized fact. In this paper we studied many inference problems for a class of financial models. We studied asymptotic equivalence of models. We obtained efficient estimators with faster rate of convergence. We studied volatility testing in the long memory setup. We studied semi-parametric estimation and pricing of FBSDEs when the backward component is driven by fractional Brownian motion. We suggested possible extension of the model to generalized fractional Brownin motion.

References

[1] Ait-Sahalia, Y. (1996a) : Testing continuous time models of the spot interest rates, *Review of Financial Studies* **9**, 385-426.
[2] Ait-Sahalia, Y. (1996b) : Nonparametric pricing of interest rate derivative securities, *Econometrica* **64** (3), 527-560.
[3] Bandi, F.M. and Philips, P.C.B. (2003) : Fully nonparametric estimation of scalar diffusion processes, *Econometrica* **71** (1), 241-283.
[4] Bertin, K., Torres, S. and Tudor, C.A. (2010) : Maximum likelihood estimators and random walks in long memory models, *Statistics: A Journal of Theoretical and Applied Statistics* **45** (4), 361-374.
[5] Bertin, K., Torres, S. and Tudor, C.A. (2011) : Drift parameter estimation in fractional diffusions driven by perturbed random walks, *Statistics and Probability Letters* **81**, 243-249.
[6] Bishwal, J.P.N. (2008) : *Parameter Estimation in Stochastic Differential Equations*, Lecture Notes in Mathematics **1923**, Springer-Verlag, Berlin, Heidelberg, New York.
[7] Bishwal, J.P.N. (2010) : Maximum likelihood estimation in Skorohod stochastic differential equations, *Proceedings of the American Mathematical Society* **138** (4), 1471-1478.
[8] Bishwal, J.P.N. (2011) : Minimum contrast estimation in fractional Ornstein-Uhlenbeck process: continuous and discrete sampling, *Fractional Calculus and Applied Analysis* **14** (3), 375-410.
[9] Bishwal, J.P.N. (2022) : *Parameter Estimation in Stochastic Volatility Models*, Springer Nature, Cham.
[10] Bishwal, J.P.N. and Pena, E. (2009) : A note on inference in a bivariate normal distribution model, *SAMSI Technical Report,* #2009-3.
[11] Brown, L.D. and Low, M.G. (1996) : Asymptotic equivalence of nonparametric regression and white noise, *Ann. Statist.* **24**, 2384-2398.
[12] Buchmann, B. and Kluppelberg, C. (2006) : Fractional integral equations and state space transform, *Bernoulli* **12**, 431-456.
[13] Chen, X. and Lin, L. (2010) : Nonparametric estimation of FBSDEs models with applications in finance, *Communications in Statistics: Theory and Methods* **39** (14), 2492-2514.
[14] Comte, F. and Renault, E. (1996) : Long memory continuous time models, *J. Econometrics* **73**, 101-149.
[15] Corradi, V. and White, H. (1999) : Specification tests for the variance of a diffusion, *J. Time Series Analysis* **20** (3), 253-270.
[16] Cox, J.C. (1996) : Notes on option pricing I: Constant elasticity of diffusions, *Journal of Portfolio Management* **23**, 15-17.
[17] Cox, J.C., Ingersoll, J.E., and Ross, S.A. (1985) : A theory of term structure of interest rates, *Econometrica* **53**, 385-407.
[18] Davydov, Y. (1970) : The invariance principle for stationary processes, *Theory of Probability and its Applications* **15**, 487-498.
[19] Dehtiar, O., Mishura, Y. and Ralchenko, K. (2021) : Two methods of estimation of the drift parameters of the Cox-Ingersoll-Ross process: Continuous observations, *Communications in Statistics-Theory and Methods* **50**, 1-16.
[20] Duncan, T.E., Hu, Y. and Pasik-Duncan, B. (2000) : Stochastic calculus for fractional Brownian motion I, *SIAM J. Control. Optim.* **38**, 582-612.
[21] El Karoui, N. and Quenez, M.C. (1995) : Dynamic programming and pricing of contingent claim in an incomplete market *SIAM J. Control. Optim.* **33**, 29-66.

[22] Feller, W. (1951) : Two singular diffusion problems, *Annals of Mathematics* **54**, 173-182.
[23] Florens-Zmirou, D. (1993) : On estimating the diffusion coefficient from discrete observations, *Journal of Applied Probability* **30** (4), 790-804.
[24] Genon-Catalot, V. and Jacod, J. (1993) : On estimation of the diffusion coefficient for multidimensional diffusion processes *Ann. Inst. Henri Poincare: Probab. Stat.* **29**, 119-151.
[25] Genon-Catalot, V. and Laredo, C. (2014) : Asymptotic equivalence of nonparametric diffusion and Euler scheme experiments, *Annals of Statistics* **42** (3), 1145-1165.
[26] Genon-Catalot, V. and Laredo, C. (2021) : Probabilistic properties and parametric inference of small variance nonlinear self-stabilizing stochastic differential equations, *Stochastic Processes and their Applications* **142**, 513-548.
[27] Grama, I.M. and Nussbaum, M. (1998) : Asymptotic equivalence of nonparametric generalized linear models, *Probab. Th. Rel. Fields* **111**, 167-214.
[28] Hu, Y. (2001) : Self-intersection local time for fractional Brownian motions via chaos expansion, *J. Math. Kyoto Univ.* **41** (2), 233-250.
[29] Hu, Y. and Oksendal, B. (2003) : Fractional white noise calculus and application to finance, *Infinite Dimensional Analysis, Quantum Probability and Related Topics* **6** (1), 1-32.
[30] Hu, Y. and Oksendal, B. (2002) : Chaos expansion of local time of fractional Brownian motion, *Stochastic Analysis and Applications* **20** (4), 815-837.
[31] Hu, Y., Oksendal, B. and Salopek, D. (2005) : Weighted local time for fractional Brownian motion and applications to finance, *Stochastic Analysis and Applications* **23** (1), 15-30.
[32] Hull, J. and White, A. (1990) : Pricing interest rate derivative securities, *Review of Financial Studies* **4**, 573-592.
[33] Ichiba, T., Pang, G., Taqqu, M.S. (2021) : Semimartingale properties of generalized fractional Brownian motion and its mixtures with applications in finance, *ArXiv:2012.00975v2*.
[34] Ichiba, T., Pang, G., Taqqu, M.S. (2022) : Path properties of a genaralized fractional Brownian motion, *J. Theoretical Probability*, **35**, 550-574.
[35] Le Cam, L. and Yang, G.L. (2000) : *Asymptotics in Statistics: Some Basic Concepts*, Second Edition, *Springer*, New York.
[36] Lenkasas, A. and Mackevicius, V. (2015) : A second order weak approximation of Heston model by discrete random variables, *Lithuanian Mathematical Journal* **55** (4), 555-572.
[37] Lileika, G. and Mackevicius, V. (2020) : Weak approximation of CKLS and CEV process by discrete random variables, *Lithuanian Mathematical Journal* **60** (2), 208-224.
[38] Lin, L., Li, F., Zhu, L. and Hardle, W.K. (2011) : Mean volatility regressions, *Technical Report, SFB 469*, Humboldt University, Berlin.
[39] Mackevicius, V. (2011): Weak approximation of CIR equation by discrete random variables, *Lithuanian Mathematical Journal* **51** (3), 385-401.
[40] Mackevicius, V. (2015) : Verhulst versus CIR, *Lithuanian Mathematical Journal* **55** (1), 119-133.
[41] Mariucci, E. (2016) : Asymptotic equivalence of discretely observed diffusion processes and their Euler scheme: small variance case, *Stat. Inference Stoch Process* **19**, 71-91.
[42] Milstein, G. and Nussbaum, M. (1998) : Diffusion approximation for nonparametric autoregression, *Probability Theory and Related Fields* **112** (4), 535-543.
[43] Mishura, Y. and Yurchenko-Tytarenko, A. (2018) : Fractional Cox-Ingersoll-Ross model with non-zero mean, *Modern Stochastics: Theory and Methods* **5** (1), 99-111.
[44] Nussbaum, M. (1996) : Asymptotic equivalence of density estimation and Gaussian white noise, *Ann. Statist.* **24**, 2399-2430.
[45] Overbeck, L. and Ryden, T. (1997) : Estimation in the Cox-Ingersoll-Ross model, *Econometric Theory* **13**, 430-461.
[46] Rendleman, R. and Bartter, B. (1980) : The pricing of options on debt securities, *Journal of Financial and Quantitative Analysis*, **15**, 11-24.
[47] Sottinen, T. (2001) : Fractional Brownian motion, random walks and binary market models, *Finance and Stochastics* **5**, 343-355.
[48] Su, Y. and Lin, L. (2009) : Semi-parametric estimation of forward-backward stochastic differential equations, *Communications in Statistics: Theory and Methods* **38** (11), 1759-1775.
[49] Zhang, S. and Karunamuni, R.J. (1988) : On kernel density estimation near endpoints, *Journal of Statistical Planning and Inference* **70**, 301-316.
[50] Zhang, Q and Lin, L. (2014) : Terminal dependent statistical inferences for FBSDE, *Stochastic Analysis and Applications* **32** (1), 128-151.