Bias in the Estimate of a Mean Reversion Parameter for a Fractional Ornstein-Uhlenbeck Process

by

Wai Man Ng

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Quantitative Finance

Waterloo, Ontario, Canada, 2017

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In this thesis we studied the estimation bias of the least squares estimate of the mean reversion parameter, when the underlying dynamics is governed by fractional Brownian motions. Fractional Brownian motion is a continuous-time model with long-range dependency features. Least squares estimate for the mean reversion parameter under standard Brownian motion framework has been shown to be positively biased. Using an approximate bias formula, we show that the estimation bias in the fractional Brownian case behaves differently from the standard Brownian motion case, and in fact can be negative depending on the Hurst parameter and the true value of the mean reversion. We conclude the thesis by looking into the implication of these results from the perspective of risk management.

Acknowledgements

I would like to thank Professor Tony Wirjanto for being my thesis supervisor and for providing various advices and assistance during the development of this thesis. Special thanks go to Professor Adam Kolkiewicz and Professor Chengguo Weng who shared their time to serve as the readers of this thesis and to provide valuable comments on the thesis. I would also like to express my gratitude to the Program of Master of Quantitative Finance (MQF) at the University of Waterloo, which helps shape my career in the industry.

Last but not least, I would like to thank my beloved Mabel Tsang for her love, patience and encouragement in pushing me to the completion of this thesis.

Table of Contents

List of Figures

Chapter 1

Introduction

Mean reversion is a key feature in many financial assets. Interest rates and historical volatility of a stock return tend to be mean-reverting, in the sense that each of these quantities has a tendency to revert to its unconditional mean in the long run. A mean-reverting time series is more likely to correct any deviation from this fixed value.

The Ornstein-Uhlenbeck (OU) process is one of the processes that captures the behaviour of mean reversion. Such a process can be represented by the following stochastic differential equation (SDE):

$$dS_t = k(\mu - S_t)dt + \sigma dB_t,$$

where $k > 0$ is the mean reversion speed, $\mu \in \mathbb{R}$ is the long term mean, $\sigma > 0$ is the instantaneous standard deviation (sometimes called volatility) and $\{B_t\}_{t\in\mathbb{R}}$ is the standard Brownian motion.

Solving this SDE over the time interval $[t_{i-1}, t_i]$ for some $i \in \mathbb{N}$ yields the following process:

$$S_i = e^{-kh}S_{i-1} + \mu(1 - e^{-kh}) + \sigma\sqrt{\frac{1 - e^{-2kh}}{2k}}\epsilon_i,$$

where $h := t_i - t_{i-1}$[1], $S_i := S_{t_i}$ and $\epsilon_i \sim N(0, 1)$ i.i.d. for each i. In this thesis we focus on the case where the parameter μ is assumed to be known. Under this circumstance we can

[1]By assuming that the time intervals are equally spaced.

assume for simplicity that $\mu = 0$ without loss of generality. As a result, given a time series $\{S_i\}_{i\in\mathbb{N}}$ which is known to follow an Ornstein-Uhlenbeck process (and which can also be viewed as a continuous version of an autoregressive process of order 1, i.e. AR(1) process), it is natural to consider using the method of ordinary least square (OLS) to estimate its mean reversion parameter yielding the following OLS estimator:

$$\hat{k} := -\frac{1}{h} \ln \frac{\sum_{1\leq i\leq n} S_{i-1}S_i}{\sum_{1\leq i\leq n} S_i^2}. \tag{1.1}$$

It is known that for a simple linear regression model where the explanatory variable is uncorrelated with the regression error terms, the OLS estimator for this slope coefficient is unbiased. However, in a time series $\{S_i\}_{i\in\mathbb{N}}$, the covariates in the AR(1) regression are typically lagged dependent variables, and such dependence makes the OLS estimate of the mean reversion parameter to be biased when the residuals of this regression are not independently distributed.

Indeed, it has long been known that the OLS estimate as prescribed in (1.1) is positively biased, i.e. the mean reversion paramter is over-estimated by the OLS method. However, using the bias estimation formula developed in [7], it is possible to derive a bias formula which takes into account the fact that the bias tends to 0 as the true value k tends to 0, despite the fact that such a bias is still significant relative to the magnitude of the true value [11].

Meanwhile, it has also long been observed that time series observed in the financial market are not driven by a standard Brownian motion. For instance, historical data from major equity indices such as Dow Jones Industrial Index reveal that the returns are much more peaked and have fatter tails compared to the tails of a normal distribution (Figure 1.1). Moreover, historical returns usually also present *persistence*, i.e. a large return on one trading day is often followed by large returns, and so do small returns. Such clustering of returns cannot be explained by the standard Brownian motion model, since it assumes that returns are independent of each other.

Similar concerns also arise when modeling volatility by standard Brownian motions. For instance, viewing the absolute daily return as a proxy to the volatility of the underlying, we can conclude that the volatility of commonly traded equity indices is serially correlated. The

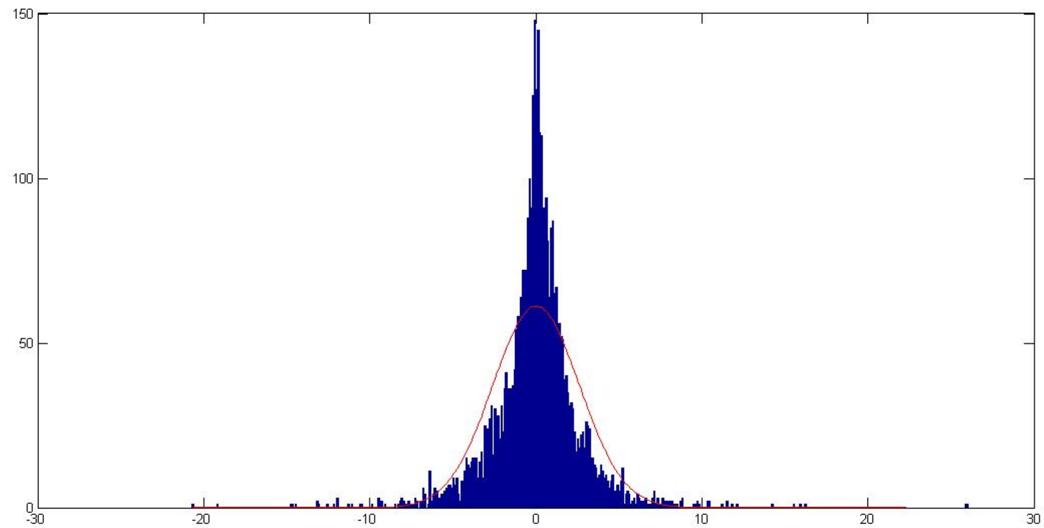

Figure 1.1: Histogram of Annualized Log Returns of DJI Index, 1/1/2000 - 1/1/2016, with Overlay by Density Function of Normal Distribution.

autocorrelation does not decay quickly enough as its lag increases (Figure 1.2). This evidence supports the argument that the volatility is not driven by standard Brownian motions either, for if otherwise we should have observed little autocorrelations in the return time series.

There are many different models which try to capture the autocorrelation feature of the return time series. One such attempt is fractional Brownian motion (fBm). Under the fBm framework, the stochastic process is still driven by some normally distributed random variables, but the increments can now be correlated with each other. The degree of correlation is governed by the so-called Hurst parameter $H \in (0, 1)$, with the increments being positively correlated when $H > \frac{1}{2}$, and negatively correlated when $H < \frac{1}{2}$. Such process nests the standard Brownian motion as a special case, when $H = \frac{1}{2}$.

There are several common features shared by standard and fractional Brownian motions. For instance, integrating a deterministic function with respect to fBm still leads to some normally distributed variable, and Ito isometry still holds under fBm. Nevertheless, applying

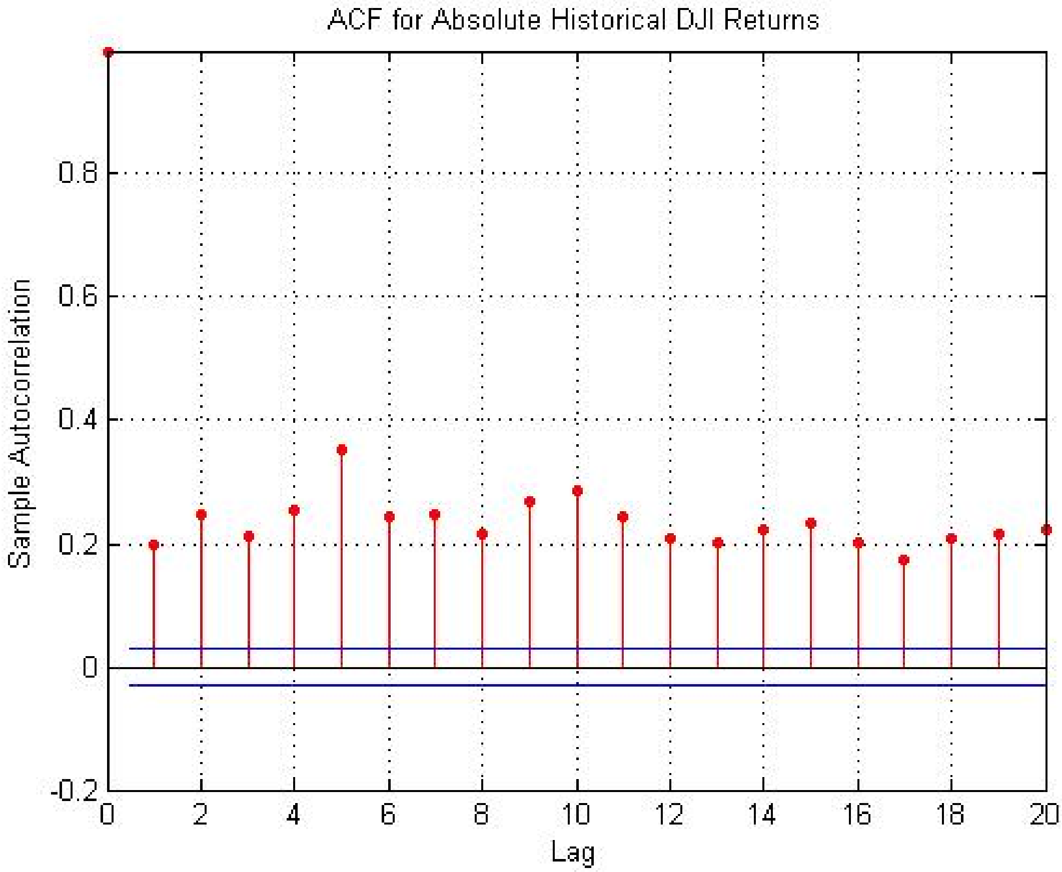

Figure 1.2: Autocorrelation of Absolute Annualized Historical Log Returns of DJI Index, 1/1/2000 - 1/1/2016.

the usual definition of a stochastic integral[2] under fBm allows for arbitrage opportunities[8], and to remedy this, it is suggested that the integral should be defined by a so-called Wick's product. The development of Ito-Wick's integrals and their applications in finance can be found for instance in [4] and [2].

In this thesis, we will generalize the results on the bias formula as presented in [11] by focusing on the estimation of the mean reversion parameter for the fractional Ornstein-

[2] In [4] and [2], such integrals are called (fractional) pathwise integrals.

Uhlenbeck (fOU) process, i.e. the mean-recersion parameter k in the following stochastic differential equation:

$$dS_t = k(\mu - S_t)dt + \sigma dB_t^H,$$

where B_t^H is a fBm with constant $H \in (\frac{1}{2}, 1)$. The result developed in this thesis covers the case in [11] by placing $H = \frac{1}{2}$. To develop the corresponding bias formula for the fOU case, we essentially need two pieces of information: (1) The bias formula for general nonlinear estimator, and (2) the moments for integrals driven by fBm.

In Chapter 2, we will review the formulation of the so-called second-order bias formula of a nonlinear estimator, based on the work by [7]. The derivation of this bias formula essentially involves a Taylor series expansion up to second order of a given stochastic expression.

In Chapter 3, we will review the theoretical background of the fractional Brownian motion. We will first briefly provide the mathematical setting for defining a fractional Brownian motion properly. Such a definition leads directly to the fact that the integral of deterministic functions with respect to fBm is also normally distributed. This normality result is essential to the later development of the thesis, for if otherwise the calculation of higher moments could be quite tedious. We will also quickly review some results regarding Wick-Ito's integration and the fBm version of Ito's lemma.

Based on the background information in the previous chapters, we will develop a new second-order bias formula for the fOU case in Chapter 4.

Chapter 2

Bias Term for a Nonlinear Estimator

Suppose that the parameter of interest is θ and its estimator is given by $\hat{\theta}$ which is dependent on the observed data. Then, to quantify how good this estimator is, we can consider an estimator bias, defined as

$$B(\hat{\theta}) := E[\hat{\theta}] - \theta,$$

where the expectation is based on the expectation of the underlying random process. In other words, the bias of an estimator is simply the expected error of using this estimator $\hat{\theta}$ to estimate the true value θ. An estimator is unbiased if the bias is zero.

This chapter is divided into three parts. In the first part the bias formula for a general nonlinear estimator is briefly revisited, based on [7]. In the second part, this bias formula is applied to the case of standard Ornstein-Uhlenbeck process, following [11], where the mean reversion of such process is estimated by the least-square estimates based on the finite sample generated from the stochastic process. In the last part, some properties regarding this estimator are discussed, together with numerical simulations to illustrate the general behaviour of the estimator bias.

2.1 Bias Formula

The class of estimators to be considered are those which are the solution to the follwing estimating equations of the form

$$\psi_n(\hat{\theta}) = \frac{1}{n}\sum_i q(\hat{\theta}) = 0, \tag{2.1}$$

given a finite sample of non-i.i.d. random variables $Z_1, \cdots, Z_n$ of size n. Here, $q(\theta) := q(Z_i; \theta)$ is a known scalar function dependent on Z_i and $\theta \in \mathbb{R}$.[1] Usually, it is also assumed that the implied estimating functions will be unbiased in the sense that

$$E[\psi_n(\theta)] = 0 \tag{2.2}$$

holds only when θ is the true value θ_0. Nevertheless, it should be emphasized that the condition that $E[\psi_n(\theta_0)] = 0$ does not necessarily hold for some estimators.

The class of estimators satisfied by (2.1) include maximum likelihood (ML) and ordinary least-square (OLS) estimators. We call such estimators *nonlinear* in the sense that unlike a multiple linear regression model, it is a general fitting procedure which encompasses both linear and nonlinear relationships among the parameter(s) to be estimated and the (independent, dependent) random variables, as long as (2.1) and some regularity conditions are satisfied.

We first recall some definitions regarding the order of magnitude in probability sense:

Definition 2.1.1 *A sequence of random variables $\{X_n\}$ is said to be*

- *at most of order n^k in probability, denoted by $O_P(n^k)$ if for each $\epsilon > 0$, there exists some positive constant $c(\epsilon) < \infty$ and integer $N(\epsilon)$, such that*

$$P\left(n^{-k}|X_n| \leq c(\epsilon)\right) \geq 1 - \epsilon, \qquad \forall n \geq N(\epsilon).$$

[1]In [7], a more general setting of multi-variate random vectors is considered instead. Since this thesis focuses on one parameter only (namely, the mean reversion parameter), we do not adopt this general framework.

- *of order smaller than n^k in probability, denoted by $o_P(n^k)$ if*

$$n^{-k}X_n \xrightarrow{P} 0,$$

or equivalently, $\lim_{n\to\infty} P[|n^{-k}X_n| > \epsilon] = 0$ *for all* $\epsilon > 0$.

By definition, $X_n \in O_P(n^0)$ is equivalent to saying that X_n is bounded in probability.

Based on [7], we state the following assumptions for the nonlinear estimator:

Assumption A

The s-th order derivative $\dfrac{\partial^s q_i}{\partial \theta^s}$ exists in a neighborhood of θ_0 and $E\left[\left|\dfrac{\partial^s q_i}{\partial \theta^s}(\theta_0)\right|^2\right] < \infty$.

Assumption B

For some neighborhood of θ_0, $\left(\dfrac{\partial \psi_n}{\partial \theta}\right)^{-1} \in O_P(1)$.

Assumption C

For some neighborhood of θ_0, we have

$$\left|\frac{\partial^s q_i}{\partial \theta^s}(\theta) - \frac{\partial^s q_i}{\partial \theta^s}(\theta_0)\right| \le M_i|\theta - \theta_0|,$$

where $E|M_i| \le C < \infty$ for each i.

One of the objectives of [7] is to derive a stochastic expansion of second order for θ

$$\hat{\theta} - \theta = A_{-1/2} + A_{-1} + o_P(n^{-1}),$$

where for each integer s, $A_{-s/2}$ represents terms of order $O_P(n^{-s/2})$, so that we have a second-order approximation for the estimation bias:

$$E[\hat{\theta}] - \theta \approx a_{-1/2} + a_{-1},$$

with $a_{-s/2} := E[A_{-s/2}]$. Hence, by "bias formula" we mean the expression $a_{-1/2} + a_{-1}$. To derive this second-order (approximated) bias, we need the above assumptions to hold for $s \ge 2$. The derivation is simply based on a Taylor series expansion, while taking care of the order of magnitude for all residual terms.

Lemma 2.1.2 ([7]) *Suppose that Assumptions A - C hold for some $s \geq 1$, then $\hat{\theta}$ has an asymptotic normal distribution.*

Lemma 2.1.3 ([7]) *Suppose that Assumptions A - C hold for $s = 2$, then*

$$\hat{\theta} - \theta = A_{-1/2} + O_P(n^{-1}),$$

where $A_{-1/2} = -\left(\frac{\partial \psi_n}{\partial \theta}(\theta_0)\right)^{-1} \psi_n(\theta_0)$.

Proof See [7] for details. ■

Lemma 2.1.4 ([7]) *Suppose that Assumptions A - C hold for $s = 2$, and that*

$$\psi_n(\hat{\theta}) = \psi_n(\theta_0) + \frac{\partial \psi_n}{\partial \theta}(\theta_0)(\hat{\theta} - \theta_0) + \frac{1}{2}\frac{\partial^2 \psi_n}{\partial \theta^2}(\theta_0)(\hat{\theta} - \theta_0)^2 + O_P(n^{-3/2}),$$

then

$$\hat{\theta} - \theta_0 = -\left(\frac{\partial \psi_n}{\partial \theta}(\theta_0)\right)^{-1} \psi_n(\theta_0) - \frac{1}{2}\left(\frac{\partial \psi_n}{\partial \theta}(\theta_0)\right)^{-1} \frac{\partial^2 \psi_n}{\partial \theta^2}(\theta_0)\left(A_{-1/2}\right)^2 + O_P(n^{-3/2}).$$

Proof See [7] for details. ■

Lemma 2.1.5 ([7]) *Suppose Assumptions A - C hold for some $s \geq 2$, then*

$$E[\hat{\theta}] - \theta_0 = a_{-1/2} + a_{-1} + O_P(n^{-3/2}), \tag{2.3}$$

where $a_{-1/2}$ and a_{-1} are defined by

$$\begin{aligned} a_{-1/2} &= -\frac{\psi_n(\theta_0)}{E\left[\psi_n'(\theta_0)\right]} \\ a_{-1} &= -\frac{\psi_n'(\theta_0) - E\left[\psi_n'(\theta_0)\right]}{E\left[\psi_n'(\theta_0)\right]} a_{-1/2} - \frac{1}{2}\frac{E\left[\psi_n''(\theta_0)\right]}{E\left[\psi_n'(\theta_0)\right]}\left(a_{-1/2}\right)^2, \end{aligned}$$

where ψ_n' and ψ_n'' are usual partials with respect to θ.

Proof See [7] for details. ■

Remark: As mentioned in [1], the proof of the above bias formula is valid for both i.i.d. and non-i.i.d. sequences of random variables $\{Z_i\}_{i=1}^n$. Indeed, in the i.i.d. case, the bias formula can be further simplified, as stated in Proposition 3.2 in [7].

2.2 Estimating the Mean Reversion Parameter for an Ornstein-Uhlenbeck (OU) Process

In this section, we directly apply the bias formula (2.3) to derive the corresponding bias formula for the mean reversion parameter for the standard Ornstein-Uhlenbeck process. This bias formula will reveal a nonlinear relationship between the bias and the true mean reversion speed. The treatment here is based on [11].

A standard Ornstein-Uhlenbeck process is governed by the following stochastic differential equation:

$$dS_t = k(\mu - S_t)dt + \sigma dB_t, \tag{2.4}$$

where $k \geq 0$ is the mean reversion parameter, $\sigma > 0$ is the volatility and B_t is the standard Brownian motion. In reality, finite sample can be extracted from a given stochastic process. We can denote such time series as

$$S_0, S_1, \cdots, S_n$$

for a sample of size n, sampled at time $0 = t_0 < t_1 < \cdots < t_n$. For simplicity, we will assume that the sampling times are equally spaced, i.e. $t_i = i \cdot h, i = 0, 1, \cdots, n$ for some fixed $h > 0$. As in the setting in [11], we assume that the initial datum is also randomly driven:

$$S_0 \sim N\left(\mu, \frac{\sigma^2}{2k}\right).$$

Note that $\frac{\sigma^2}{2k}$ is the unconditional variance of S_0. We can further assume that $\mu = 0$.[2]

Solving the SDE (2.4) over the interval $[t_{i-1}, t_i]$ gives

$$S_i = e^{-kh} S_{i-1} + \sigma \sqrt{\frac{1 - e^{-2kh}}{2k}} \epsilon_i, \tag{2.5}$$

where $\epsilon_i \sim N(0, 1)$.

[2]The assumptions of S_0 to be random and μ is known are the setting used in [11]. Following the same setup allows us to directly compare the bias behavior under standard OU and fOU processes.

Given a finite sample $\{S_i\}_{i=1}^n$, the OLS estimator of k, denoted as $\hat{k}$ is defined by

$$\hat{k} = \underset{k}{\operatorname{argmin}} \sum_{i=1}^{n} (S_i - e^{-kh} S_{i-1})^2.$$

By simple calculus, we arrive at the following equivalent expression satisfied by this $\hat{k}$:

$$\sum_{i=1}^{n} S_{i-1}(S_i - e^{-\hat{k}h} S_{i-1}) = 0. \tag{2.6}$$

Under the standard OU process, it can be shown that the OLS estimator coincides with the maximum likelihood (ML) estimator $\hat{k}_{ML}$, with the latter being as

$$\hat{k}_{ML} = \underset{k}{\operatorname{argmax}} \ln(pdf(S_i|S_{i-1})),$$

where *pdf* is the probability density function of S_i given S_{i-1}. From (2.5), we have

$$S_i|S_{i-1} \sim N\left(e^{-kh} S_{i-1}, \sigma^2 \frac{1 - e^{-2kh}}{2k}\right).$$

The objective of this section is to derive a bias formula of second order from (2.6). Let

$$\begin{aligned}
S &= [S_0, S_1, \cdots, S_n]^T \\
U_n &= \frac{1}{n} S^T C_1 S \\
V_n &= \frac{1}{n} S^T C_2 S \\
C_1 &= \frac{1}{2} \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \\
C_2 &= \begin{bmatrix} I_{n\times n} & 0_{n\times 1} \\ 0_{n\times 1}^T & 0 \end{bmatrix}
\end{aligned}$$

where $I_{n\times n}$ is the n-dimensional identity matrix and $0_{n\times 1}$ is the n-dimensional zero (column) vector. Then we can rewrite (2.6) as

$$U_n - e^{-\hat{k}h} V_n = 0.$$

To apply (2.3), we need to compute the expected values of some functions of the quadratic forms U_n and V_n.

Lemma 2.2.1 ([10]) *Suppose that $S \sim N(0, \Sigma)$, and A, A_1, A_2 are symmetric matrices, then we have*

$$E(S^T AS) = \operatorname{tr}(A\Sigma) \tag{2.7}$$
$$E[(S^T AS)^2] = (\operatorname{tr}(A\Sigma))^2 + 2\operatorname{tr}(A\Sigma A\Sigma) \tag{2.8}$$
$$E[S^T A_1 S S^T A_2 S] = \operatorname{tr}(A_1\Sigma)\operatorname{tr}(A_2\Sigma) + 2\operatorname{tr}(A_1\Sigma A_2\Sigma) \tag{2.9}$$

Remark: This lemma essentially states that expectations of quadratic forms are all dependent up to the second moment (Σ). However, it is important to note that this is not always true for random vectors with a non-normal distribution (e.g. (2.8) and (2.9) involve co-kurtosis terms). Hence, the normality assumption is crucial for the validity of this lemma.

The proof of this lemma relies on a trick used in [10] regarding quadratic forms of normally distributed random variables.

Lemma 2.2.2 ([10]) *Let f be the pdf of a n-dimensional normally distributed random vector y:*

$$f(y) = \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}(y-\mu)^T\Sigma^{-1}(y-\mu)\right)$$

Define a differential operator

$$d = \mu + \Sigma\frac{\partial}{\partial\mu}.$$

Then for any given analytic function $h : \mathbb{R}^n \to \mathbb{R}$, we have

$$h(y)f(y) = h(d)f(y).$$

Remark: It should be noted that any higher orders of d should be interpreted as recursive differential operators. For instance,

$$d^2 f(y) = d(df(y)) = d\left(\mu + \Sigma\frac{\partial}{\partial\mu}\right) f(y) \neq \left(\mu^2 + 2\Sigma\frac{\partial}{\partial\mu} + 2\left(\Sigma\frac{\partial}{\partial\mu}\right)^2\right) f(y).$$

Moreover, the shorthand $h(d)$ should be interpreted as a differential operator in the sense that if $h(y) = \sum_\alpha a_\alpha y^\alpha$ where α is an multi-index, then $h(d) = \sum_\alpha a_\alpha d^\alpha$.

Proof of Lemma 2.2.2 Here, we adopt a "less operator" approach compared to the one shown in [10]. We can start from a one-dimensional case: $y, \Sigma \in \mathbb{R}$. By definition of the differential operator d, we can easily check that

$$d^m f(y) = y^m f(y), \qquad m = 1, 2, 3, \cdots. \tag{2.10}$$

Then by a componentwise consideration, we can deduce that (2.10) also holds for $y \in \mathbb{R}^n$ and any multi-index $m = (m_1, \cdots, m_n)$. Since every real-valued analytic function h takes the form $h(y) = \sum_{m=0}^{\infty} a_m y^m$, the result follows immediately. ■

Corollary 2.2.3 ([10]) *For any analytic function $h : \mathbb{R}^n \to \mathbb{R}$ and any $Y \sim N(\mu, \Sigma)$, we have*

$$E[h(Y)] = h(d) \cdot 1.$$

In particular, we have $E[Y] = d \cdot 1 = \mu$ and $E[Y^T Y] = d^T d \cdot 1 = d^T \mu = \mu^T \mu + \Sigma$.

Proof By Lemma 2.2.2, $h(y)f(y) = h(d)f(y)$ and hence

$$E[h(Y)] = \int h(y)f(y)dy = \int h(d)f(y)dy = h(d) \cdot \int f(y)dy = h(d) \cdot 1. \qquad ■$$

It is also handy to have some simple results regarding differentiating quadratic forms:

Lemma 2.2.4 *For a constant $n \times n$ symmetric matrix M, we have*

$$\left(\frac{\partial}{\partial\mu}\right)^T (M\mu) = \mathrm{tr}(M)$$

$$\frac{\partial}{\partial\mu}(\mu^T M\mu) = 2M\mu$$

Proof This involves a straightforward calculus exercise once we recall $\frac{\partial}{\partial \mu} = \left[\frac{\partial}{\partial \mu_1}, \cdots \frac{\partial}{\partial \mu_n}\right]^T$. ■

Proof of Lemma 2.2.1 By Lemmas 2.2.3 and 2.2.4, we have

$$\begin{aligned} E\left[S^T A S\right] &= d^T A d \cdot 1 = \left(\mu + \Sigma \frac{\partial}{\partial \mu}\right)^T A \left(\mu + \Sigma \frac{\partial}{\partial \mu}\right) \cdot 1 = \left(\mu + \Sigma \frac{\partial}{\partial \mu}\right)^T A \mu \\ &= \mu^T A \mu + \left(\frac{\partial}{\partial \mu}\right)^T \Sigma^T A \mu = \mu^T A \mu + \operatorname{tr}(\Sigma^T A) \\ &= \mu^T A \mu + \operatorname{tr}(A\Sigma), \end{aligned}$$

where the last equality holds because Σ is symmetric and it is always true that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ whenever both matrix products are well-defined. Next, we have

$$\begin{aligned} E\left[S^T A_1 S S^T A_2 S\right] &= (d^T A_1 d)(d^T A_2 d) \cdot 1 = (d^T A_1 d)\left(\mu^T A_2 \mu + \operatorname{tr}(A_2 \Sigma)\right) \\ &= (d^T A_1 d)(\mu^T A_2 \mu) + \operatorname{tr}(A_2 \Sigma)(d^T A_1 d) \cdot 1 \\ &= \underbrace{(d^T A_1 d)(\mu^T A_2 \mu)}_{Expr_1} + \operatorname{tr}(A_2 \Sigma)(\mu^T A_1 \mu + \operatorname{tr}(A_1 \Sigma)) \end{aligned}$$

$$\begin{aligned} Expr_1 &= d^T A_1 \left(\mu + \Sigma \frac{\partial}{\partial \mu}\right)(\mu^T A_2 \mu) = d^T A_1 (\mu \mu^T A_2 \mu + \Sigma(2 A_2 \mu)) \\ &= \left(\mu + \Sigma \frac{\partial}{\partial \mu}\right)^T (A_1 \mu \mu^T A_2 \mu + 2 A_1 \Sigma A_2 \mu) \\ &= (\mu^T A_1 \mu)(\mu^T A_2 \mu) + 2\mu^T A_1 \Sigma A_2 \mu + \underbrace{\left(\frac{\partial}{\partial \mu}\right)^T (\Sigma A_1 \mu \mu^T A_2 \mu + 2 \Sigma A_1 \Sigma A_2 \mu)}_{Expr_2} \end{aligned}$$

Note that since $\mu^T A_2 \mu$ is a scalar, we have $\Sigma A_1 \mu \mu^T A_2 \mu = \mu^T A_2 \mu \Sigma A_1 \mu$. Now, apply the product rule of differentiation,

$$\begin{aligned} Expr_2 &= (\mu^T A_2 \mu)\left(\frac{\partial}{\partial \mu}\right)^T (\Sigma A_1 \mu) + \left(\frac{\partial}{\partial \mu}\right)^T (\mu^T A_2 \mu)(\Sigma A_1 \mu) + 2\operatorname{tr}(\Sigma A_1 \Sigma A_2) \\ &= (\mu^T A_2 \mu)\operatorname{tr}(\Sigma A_1) + (2 A_2 \mu)^T (\Sigma A_1 \mu) + 2\operatorname{tr}(\Sigma A_1 \Sigma A_2) \\ \Rightarrow\ E\left[S^T A_1 S S^T A_2 S\right] &= (\mu^T A_1 \mu)(\mu^T A_2 \mu) + 2\mu^T A_1 \Sigma A_2 \mu + (\mu^T A_2 \mu)\operatorname{tr}(\Sigma A_1) \\ &\quad + (2 A_2 \mu)^T (\Sigma A_1 \mu) + 2\operatorname{tr}(\Sigma A_1 \Sigma A_2) + \operatorname{tr}(A_2 \Sigma)(\mu^T A_1 \mu + \operatorname{tr}(A_1 \Sigma)). \end{aligned}$$

Finally, substituting $\mu = 0$ will give (2.7) and (2.9). (2.8) is then a special case of (2.9) with $A_1 = A_2 = A$. ■

Remark: Since the differential operator d involves a partial derivative with respect to μ, we cannot directly substitute $\mu = 0$ before any differentiation takes place.

With Lemma 2.2.1 in hand we can reduce the computations of $E[U_n]$, $E[V_n]$, $E[U_n^2]$ and $E[V_n^2]$ to some manipulations of traces. Computing these traces requires some algebraic identities:

Lemma 2.2.5 *For any $\phi \in \mathbb{R}$, we have*

$$\sum_{i=1}^{n-1}(n-i)\phi^{2i} = \frac{n\phi^2}{1-\phi^2} - \frac{\phi^2(1-\phi^{2n})}{(1-\phi^2)^2}$$

$$\sum_{\alpha,\beta=1}^{n}\phi^{2|\alpha-\beta|} = n + 2\sum_{i=1}^{n-1}(n-i)\phi^{2i} = n + \frac{2n\phi^2}{1-\phi^2} - \frac{2\phi^2(1-\phi^{2n})}{(1-\phi^2)^2}$$

$$\sum_{\alpha,\beta=1}^{n}\phi^{|\alpha-\beta+1|+|\alpha-\beta-1|} = n\phi^2 + 2\sum_{i=1}^{n-1}(n-i)\phi^{2i} = n\phi^2 + \frac{2n\phi^2}{1-\phi^2} - \frac{2\phi^2(1-\phi^{2n})}{(1-\phi^2)^2}$$

Proof Deriving the first equation is a standard exercise for an arithmetico-geometric series. The remaining two equations can be obtained by counting the number of ϕ^{2i} for each integer i. ■

Lemma 2.2.6 ([11]) *Denote $\phi = e^{-kh}$ and $\Sigma = [c_{\alpha\beta}]_{1\le\alpha,\beta\le n+1}$. Then we have*

$$c_{\alpha\beta} = E[S_{\alpha-1}\cdot S_{\beta-1}] = \frac{\sigma^2}{2k}\phi^{|\alpha-\beta|}$$

$$E[U_n] = \frac{\sigma^2}{2k}\phi$$

$$E[V_n] = \frac{\sigma^2}{2k}$$

$$E[U_n^2] = \frac{\sigma^4}{4k^2}\left[\phi^2 + \frac{1+4\phi^2-\phi^4}{n(1-\phi^2)} - \frac{4\phi^2(1-\phi^{2n})}{n^2(1-\phi^2)^2}\right]$$

$$E[V_n^2] = \frac{\sigma^4}{4k^2}\left[1 + \frac{2(1+\phi^2)}{n(1-\phi^2)} - \frac{4\phi^2(1-\phi^{2n})}{n^2(1-\phi^2)^2}\right],$$

where S_n, U_n and V_n are defined at the beginning of this section, preceding Lemma 2.2.1.

Proof We first compute the covariance terms for $S_i, i = 0, 1, \cdots, n$. By Ito's isometry, we have

$$
\begin{aligned}
E[S_i S_j] &= E\left[\sigma e^{-kt_i}\int_{-\infty}^{t_i} e^{ts}dB_s \cdot \sigma e^{-kt_j}\int_{-\infty}^{t_j} e^{ts}dB_s\right] = \sigma^2 e^{-k(t_i+t_j)}\int_{-\infty}^{\min(t_i,t_j)} e^{2ks}ds \\
&= \frac{\sigma^2}{2k}e^{k|t_i-t_j|} = \frac{\sigma^2}{2k}\phi^{|i-j|}.
\end{aligned}
$$

We then compute the following traces (note that due to symmetry $c_{\alpha\beta} = c_{\beta\alpha}$. Also, we adopt the shorthand notation that $c_{0\beta} = c_{\alpha 0} = 0$):

$$
\begin{aligned}
\operatorname{tr}(C_1\Sigma C_1\Sigma) &= \sum_{\alpha=1}^{n+1}\sum_{\beta=1}^{n+1}(C_1\Sigma)_{\alpha\beta}(C_1\Sigma)_{\beta\alpha} = \frac{1}{4}\sum_{\alpha=1}^{n+1}\sum_{\beta=1}^{n+1}\left(c_{\alpha-1,\beta}+c_{\alpha+1,\beta}\right)\left(c_{\beta-1,\alpha}+c_{\beta+1,\alpha}\right) \\
&= \frac{1}{4}\left[\sum_{\alpha=1}^{n}\sum_{\beta=1}^{n} c_{\alpha+1,\beta}c_{\alpha,\beta+1} + \sum_{\alpha=2}^{n+1}\sum_{\beta=2}^{n+1} c_{\alpha-1,\beta}c_{\alpha,\beta-1} + \sum_{\alpha=2}^{n+1}\sum_{\beta=1}^{n} c_{\alpha-1,\beta}c_{\alpha,\beta+1}\right. \\
&\qquad \left. + \sum_{\alpha=1}^{n}\sum_{\beta=2}^{n+1} c_{\alpha+1,\beta}c_{\alpha,\beta-1}\right] \\
&= \frac{1}{4}\sum_{\alpha,\beta=1}^{n}\left[c_{\alpha+1,\beta}c_{\alpha,\beta+1} + c_{\alpha,\beta+1}c_{\alpha+1,\beta} + c_{\alpha,\beta}c_{\alpha+1,\beta+1} + c_{\alpha+1,\beta+1}c_{\alpha,\beta}\right] \\
&= \frac{1}{2}\cdot\frac{\sigma^4}{4k^2}\sum_{\alpha,\beta=1}^{n}\left(\phi^{|\alpha-\beta+1|+|\alpha-\beta-1|} + \phi^{2|\alpha-\beta|}\right) \\
&= \frac{1}{2}\cdot\frac{\sigma^4}{4k^2}\left[n\frac{1-\phi^4+4\phi^2}{1-\phi^2} - \frac{4\phi^2(1-\phi^{2n})}{(1-\phi^2)^2}\right]
\end{aligned}
$$

The last equality is due to Lemma 2.2.5. In a similar fashion, we can obtain

$$
\begin{aligned}
\operatorname{tr}(C_1\Sigma) &= n\cdot\phi\frac{\sigma^2}{2k} \\
\operatorname{tr}(C_2\Sigma) &= n\cdot\frac{\sigma^2}{2k} \\
\operatorname{tr}(C_2\Sigma C_2\Sigma) &= \frac{\sigma^2}{4k^2}\left(n + \frac{2n\phi^2}{1-\phi^2} - \frac{2\phi^2(1-\phi^{2n})}{(1-\phi^2)^2}\right).
\end{aligned}
$$

Afterwards, a direct application of Lemma 2.2.1 yield $E[U_n]$, $E[V_n]$, $E[U_n^2]$ and $E[V_n^2]$. ■

Theorem 2.2.7 ([11]) *The second order bias for the OLS estimator of the mean reversion parameter is given by*

$$E(\hat{k}) - k \approx \frac{1}{2T}(3 + e^{2kh}) - \frac{2(1 - e^{-2nkh})}{Tn(1 - e^{-2kh})}, \tag{2.11}$$

where $T = n \cdot h$.

Proof This involves a direct application of Lemma 2.1.5 with $\psi_n(k) = U_n - e^{-kh}V_n$. It turns out that in this case $E[a_{-1/2}] = 0$ and $E[a_{-1}] = \left(\frac{2k}{\sigma^2}\right)^2 \frac{E[U_n^2] - \phi^2 E[V_n^2]}{2h\phi^2}$. See [11] for details. ■

2.3 Properties of Bias of Mean Reversion Parameter and Simulations

Based on the bias formula (2.11), we can deduce the following properties regarding the OLS estimator of the mean reversion parameter:

Corollary 2.3.1 *The estimation bias for parameter* k *is always positive.*

Corollary 2.3.2 *The OLS estimator is* T*-consistent, i.e. as* $T \to \infty$, $E[\hat{k}] - k \to 0$.

Corollary 2.3.3 ([11]) *As* $k \to 0$, *The bias for the OLS estimator tends to* 0.

Proof By L'Hospital's rule, we have $\lim_{k \to 0} \frac{1 - e^{-2nkh}}{n(1 - e^{-2kh})} = 1$ and hence $\lim_{k \to 0} E[a_{-1}] = 0$. ■

Corollary 2.3.3 is crucial in the sense that it was thought that the bias would be linear with the true value k, and is non-zero even when k is small, prior to the results by [11].

Corollary 2.3.4 ([11]) *The estimation bias for* h *does not vanish by increasing the sampling frequency. In particular, when* T *is kept fixed, we have*

$$\lim_{h \to 0} E[\hat{k}] - k = \frac{1}{T}\left[2 - \frac{1 - e^{-2Tk}}{Tk}\right] \neq 0.$$

Proof Rewrite the bias formula in terms of T and n:

$$E[\hat{k}] - k = \frac{1}{2T}\left(3 + e^{\frac{2kT}{n}}\right) - \frac{2(1-e^{-2kT})}{n(1-e^{-\frac{2kT}{n}})}.$$

Then, by L'Hospital rule, $n(1-e^{-\frac{2kT}{n}}) = -2kT$. The rest of the proof is straightforward. ■

To understand the actual bias as well as to compare this actual bias against the theoretical bias derived in previous sections, we adopt a simulation approach as described in [11]. We first fix a time horizon T and a time interval $h > 0$. This fixes the number of time steps n if we take $n = \lceil T/h \rceil$, for example. Then, 10000 simulation paths $\{S_i\}_{i=0,\cdots,n}$ are generated based on the discrete formula (2.5). For each of these paths, the mean reversion parameter estimate $\hat{k}$ is computed using (2.6). Finally, the expected value of the estimate is obtained by averaging these estimates over all paths. This process is repeated for a range of values of $k \in (0, 3]$.

Several plots of these empirical and theoretical biases are shown in Figures 2.1-2.3. The time horizon is fixed at $T = 3, 5$ or 10 years in each of the figures, with $h = 1/252, 1/52$ and $1/12$ corresponding to daily, weekly and monthly sampling. The estimation bias is shown to be always positive and nonlinear, with diminishing bias as k decreases to 0. In particular, the relative error $(E[\hat{k}] - k)/k$ is significant even when k is small. These results confirm that an adjustment based on bias formula such as (2.11) becomes necessary in order to obtain a correct estimate of the mean reversion parameter for the standard OU process.

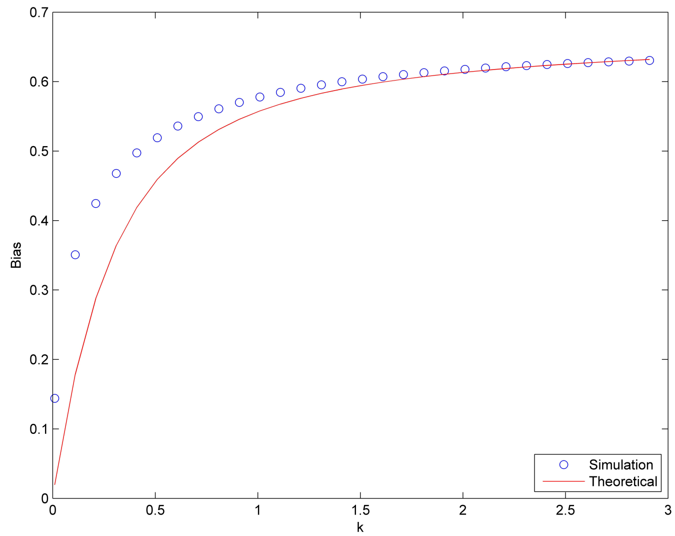

Figure 2.1: Empirical and Theoretical Bias for k, with $T = 3$ and $h = 1/252$.

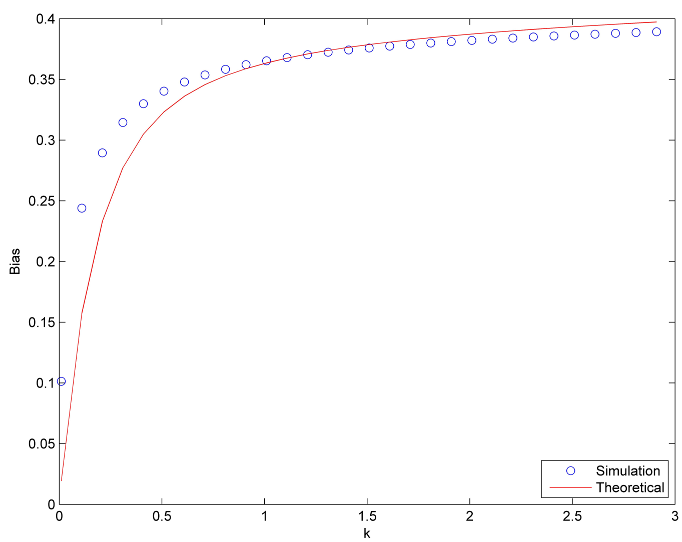

Figure 2.2: Empirical and Theoretical Bias for k, with $T = 5$ and $h = 1/52$.

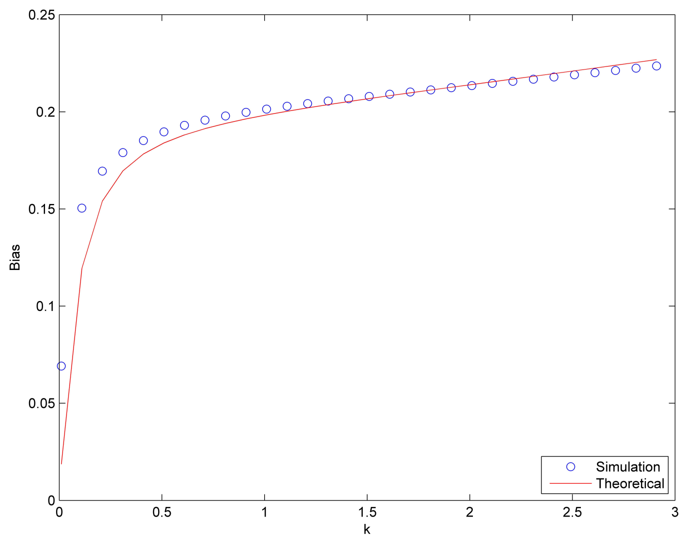

Figure 2.3: Empirical and Theoretical Bias for k, with $T = 10$ and $h = 1/12$.

Chapter 3

A Fractional Brownian Motion

In this thesis, we focus on the estimation bias of the mean reversion parameter. In last chapter, we have reviewed the bias formula when the underlying dynamic is the standard OU process. From now on, we consider the estimation bias under the fractional Ornstein-Uhlenbeck (fOU) process, i.e. estimating the parameter k in the following stochastic differential equation:

$$dS_t = k(\mu - S_t)dt + \sigma dB_t^H,$$

where $S_t = S(t)$ is the underlying asset (such as interest rate, volatility, etc.), and B_t^H with $H \in (0, 1)$ represents the fractional Brownian motion (fBm) with Hurst parameter H at time t. This is a Gaussian process satisfying

$$E(B_t^H) = 0 \tag{3.1}$$

$$E[B_t^H B_s^H] = \frac{1}{2}\{|t|^{2H} + |s|^{2H} - |s-t|^{2H}\}, \quad \forall s, t \in \mathbb{R}. \tag{3.2}$$

The above covariance requirement is equivalent to the fact the random increments are serially correlated unless $H = 1/2$:

Lemma 3.0.1 *Given $0 \le s_1 \le t_1 \le s_2 \le t_2$, the covariance of fBm increments is given by*

$$E\left((B_{t_1}^H - B_{s_1}^H) \cdot (B_{t_2}^H - B_{s_2}^H)\right) = \frac{1}{2}\{|t_1 - s_2|^{2H} + |t_2 - s_1|^{2H} - |t_1 - t_2|^{2H} - |s_1 - s_2|^{2H}\}. \tag{3.3}$$

In particular, when $H = 1/2$, the fBm increments are uncorrelated, which reduces to the standard Brownian framework.

Proof Straightforward computation. ■

In later discussion, we will focus on uniform time steps $\{t_i = i \cdot h | i = 0, \cdots, n\}$ (where $h > 0$ is fixed) and hence it is convenient to rewrite (3.3) into the following form:

$$\gamma(n) := E\left((B_{kh}^H - B_{(k+1)h}^H) \cdot (B_{(k+n)h}^H - B_{(k+n+1)h}^H)\right) = \frac{h^{2H}}{2}\{|n+1|^{2H} + |n-1|^{2H} - 2|n|^{2H}\}. \tag{3.4}$$

Note that the above covariance expression is independent of k.

The general behaviour of the fBm process is somehow different from their standard Brownian motion counterpart. Figure 3.1 shows some simulated mean-reverting paths under different values of H. First, a sequence of i.i.d. normally distributed random numbers are generated to produce the sample path for the case $H = 0.5$. These random numbers are then adjusted based on a Cholesky decomposition to produce sample paths for $H = 0.3$ and $H = 0.7$. It is obvious from Figure 3.1 that the higher the value of H, the smoother the sample path is. The increased smoothness is due to a higher level of persistence in the time series, as the time series at different time spots are more positively correlated when H increases.

In what follows, we will only consider the case where $H > \frac{1}{2}$. In such a case, it is known that the fBm exhibits long-range dependency:

Lemma 3.0.2 *A fractional Brownian motion with $H > 1/2$ exhibits long range dependence, i.e. the autovariance function $\gamma(n)$ satisfies the following asymptotic relation:*

$$\lim_{n\to\infty} \frac{\gamma(n)}{cn^{-\alpha}} = 1,$$

for some constants c and $\alpha \in (0, 1)$. In addition, the autocovariance decays slowly as $n \to \infty$ and

$$\sum_{n=1}^{\infty} \gamma(n) = \infty.$$

Proof Using L'Hopital's rule, the following equality holds:

$$\lim_{n\to\infty} \frac{(n+1)^{2H+2} + (n-1)^{2H+2} - 2n^{2H+2}}{n^{2H}} = \lim_{n\to\infty} \frac{(2H+2)(2H+1)}{2H(2H-1)} \cdot \frac{\gamma(n)}{\frac{h^{2H}}{2} n^{2H-2}}.$$

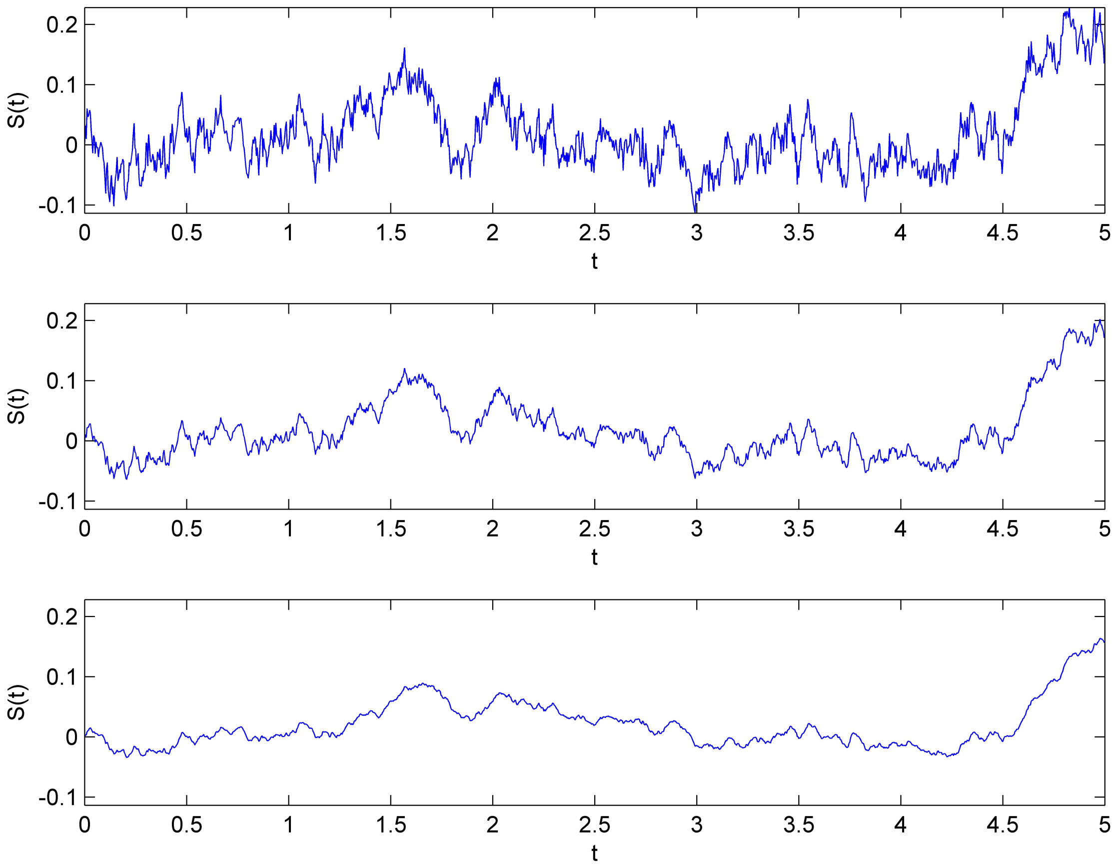

Figure 3.1: Simulation of Fractional Ornstein-Uhlenbeck process with $\mu = 0, T = 5, k = 1, \sigma = 0.1$ and $H = 0.3$ (Top), 0.5 (Middle) and 0.7 (Bottom).

Thus, we can take $c = \frac{h^{2H}}{2} \cdot \frac{2H(2H-1)}{(2H+2)(2H+1)}$ and $\alpha = 2 - 2H$, the latter of which lies within $(0, 1)$ when $H \in (\frac{1}{2}, 1)$. Using a comparison test, it is easy to conclude that the infinite sum $\sum \gamma(n)$ diverges since $\sum n^{-\alpha}$ for $\alpha \in (0, 1)$ does. $\blacksquare$

3.1 A Theoretical Setup

In this section, we briefly summarize the setup for the fractional Brownian process. The development is mainly based on the materials in [2] and [4], with some sporadic ideas borrowed from harmonic analysis (see [6]).

Given a fixed $H \in (\frac{1}{2}, 1)$, define

$$\phi(s,t) = H(2H-1)|s-t|^{2H-2}, \qquad \forall s,t \in \mathbb{R}. \tag{3.5}$$

A measurable function $f : \mathbb{R} \to \mathbb{R}$ is said to be in $L^2_\phi(\mathbb{R})$ if

$$|f|_\phi^2 := \int_{\mathbb{R}} \int_{\mathbb{R}} f(s)f(t)\phi(s,t)dsdt < \infty. \tag{3.6}$$

We can equip this space with an inner product: for all $f, g \in L^2_\phi(\mathbb{R})$,

$$(f,g)_\phi := \int_{\mathbb{R}} \int_{\mathbb{R}} f(s)g(t)\phi(s,t)dsdt. \tag{3.7}$$

Now, we want to construct a Gaussian process satisfying (3.1) and (3.2). This in turns requires us to define properly the *probability measure* μ_ϕ so that the expectations in (3.1)-(3.2) make sense. To achieve this purpose we need Bochner-Minlos theorem. Before we state the theorem, it is worthwhile to note a number of facts regarding a class of functions:

Definition 3.1.1 *A Schwarz space $\mathcal{S}(\mathbb{R})$ is the space of all rapidly decreasing smooth functions on $\mathbb{R}$. More precisely,*

$$\mathcal{S}(\mathbb{R}) := \left\{ f : \mathbb{R} \to \mathbb{R} | f \in C^\infty(\mathbb{R}), \lim_{|x|\to\infty} |x^n f^{(k)}(x)| = 0, \forall n, k = 0, 1, 2, \cdots \right\}.$$

Moreover, we can define a family of semi-norms $|\cdot|_{n,k}$ over $\mathcal{S}(\mathbb{R})$:

$$|f|_{n,k} := \left(\int_{\mathbb{R}} |x^n f^{(k)}(x)|^2 dx \right)^{\frac{1}{2}}.$$

This Schwarz space has the following nice property:

Theorem 3.1.2 $(\mathcal{S}(\mathbb{R}), |\cdot|_{n,k})$ *is a nuclear space, i.e. a topological vector space V whose topology is defined by a family of Hilbert semi-norms $\{|\cdot|_\alpha\}_{\alpha \in I}$, such that for any Hilbert semi-norm p we can find a larger Hilbert semi-norm q such that the inclusion map $\iota_{q,p} : V_q \hookrightarrow V_p$ is Hilbert-Schmidt, where V_α stands for the completion of V using $|\cdot|_\alpha$.*

To avoid going astray, we refer to [6] for the proof of the above theorem. At this moment, however, it should be emphasized that the Schwarz space being a nuclear space ensures that the probability measure to be constructed is countably additive [6]. We can now state the Bochner-Minlos theorem below:

Theorem 3.1.3 *Given a nuclear space $\mathcal{S}$, any continuous positive definite linear functional Λ on $\mathcal{S}$ satisfying $\Lambda(0) = 1$ is the Fourier transform of a countably additive positive normalized measure μ on the dual space $\mathcal{S}'$ of $\mathcal{S}$, i.e.*

$$\Lambda(f) = \int_{\mathcal{S}'} e^{i(F,f)} d\mu(F), \qquad \forall f \in \mathcal{S},$$

where (F, f) is the natural pairing of $\mathcal{S}$ and $\mathcal{S}'$.

We are now ready to apply Bochner-Minlos theorem to construct our desired probability measure. Take $\mathcal{S} = \mathcal{S}(\mathbb{R})$. Its dual $\Omega := \mathcal{S}(\mathbb{R})'$ is the space of tempered distribution ω on $\mathbb{R}$. Consider a linear functional

$$\Lambda(f) := \exp\left(-\frac{1}{2}|f|_\phi^2\right), \qquad \forall f \in \mathcal{S}(\mathbb{R}).$$

Then it is straightforward to observe that $\Lambda(0) = 1$, and Λ is continuous and positive definite. Hence, by Bochner-Minlos theorem, there exists a probability measure μ_ϕ on Ω such that

$$\int_\Omega e^{i(\omega,f)} d\mu_\phi(\omega) = \exp\left(-\frac{1}{2}|f|_\phi^2\right), \qquad \forall f \in \mathcal{S}(\mathbb{R}). \tag{3.8}$$

Now, by replacing all f in (3.8) by $t \cdot f$ where $t \in \mathbb{R}$ is a dummy variable, and by considering the resulting Taylor series expansion of (3.8), we can obtain

$$E_{\mu_\phi}[(\cdot, f)] = 0 \tag{3.9}$$
$$E_{\mu_\phi}[(\cdot, f)^2] = |f|_\phi^2, \tag{3.10}$$

where it is emphasized that the expectation is taken with respect to μ_ϕ. This allows us to define

$$\overline{B}_H(t) = \overline{B}_H(t, \omega) = (\omega, \chi_{[0,t]}(\cdot)) \tag{3.11}$$

as an element of $L^2(\mu_\phi)$ [1]for each $t \in \mathbb{R}$ where $\chi_A : \mathbb{R} \to \{0,1\}$ for a given set A stands for the usual indicator function such that

$$\chi_{[0,t]}(s) = \begin{cases} 1 & 0 \le s \le t \\ -1 & t \le s \le 0 \text{ except } t = s = 0 \\ 0 & \text{otherwise.} \end{cases}$$

Now, the picture becomes clearer if we substitute (3.11) into (3.8):

$$\int_\Omega e^{i\overline{B}_H(t)} d\mu_\phi(\omega) = \exp\left(-\frac{1}{2}|\chi_{[0,t]}(\cdot)|_\phi^2\right) = \exp\left(-\frac{1}{2}|t|^{2H}\right), \tag{3.12}$$

where the second equality can be computed directly based on the definition of $|\cdot|_\phi$ and $\chi_{[0,t]}$. Observe that LHS of (3.12) is the characteristic function of $\overline{B}_H(t)$. This means that $\overline{B}_H(t)$ by construction is a Gaussian process (with mean $= 0$ and variance $= |t|^{2H}$) for each $t \in \mathbb{R}$. By a polarization argument, we can also obtain

$$\begin{aligned} E_{\mu_\phi}[\overline{B}_H(s)\overline{B}_H(t)] &= \frac{1}{2}\left\{E_{\mu_\phi}\left[(\overline{B}_H(s) + \overline{B}_H(t))^2 - \overline{B}_H^2(s) - \overline{B}_H^2(t)\right]\right\} \\ &= \frac{1}{2}\left\{E_{\mu_\phi}\left[(\cdot, \chi_{[0,s]} + \chi_{[0,t]})^2 - (\cdot, \chi_{[0,s]})^2 - (\cdot, \chi_{[0,t]})^2\right]\right\} \\ &= \frac{1}{2}\left\{|\chi_{[0,s]} + \chi_{[0,t]}|_\phi^2 - |\chi_{[0,s]}|_\phi^2 - |\chi_{[0,t]}|_\phi^2\right\} \\ &= (\chi_{[0,s]}, \chi_{[0,t]})_\phi \\ &= \frac{1}{2}\left\{|s|^{2H} + |t|^{2H} - |s-t|^{2H}\right\}, \end{aligned}$$

where the second last equality is due to the definition of norms induced by the inner product ($|f|_\phi^2 = (f,f)_\phi$) and the last equality relies on straightforward computation of $(\chi_{[0,s]}, \chi_{[0,t]})_\phi$ based on the definition of $\chi_{[0,\alpha]}$ for different values of s and t. In other words, the requirement for being qualified as a fractional Brownian motion (equations (3.1)-(3.2)) is fulfilled by $\overline{B}_H(t)$.

Note that, however, $\overline{B}_H(t)$ constructed so far is not continuous in t. We can apply the classical Kolmogorov argument to modify it to a continuous process:

[1]It should be reminded that $L^2(\mu_\phi)$ is L^2 space with respect to μ_ϕ, which is *different* from $L^2_\phi(\mathbb{R})$.

Theorem 3.1.4 (Kolmogorov Continuity Theorem) *Let $(B, ||\cdot||)$ be a Banach space equipped with norm $||\cdot||$, and $(x_t, t \in \mathbb{R})$ be a stochastic process such that $x_t \in B$. Suppose that there exist positive p, δ, C, such that*

$$E\left[||x_t - x_s||^p\right] \leq C|t-s|^{1+\delta}, \qquad \forall s, t \in \mathbb{R},$$

then there is a continuous modification $(\hat{x}_t, t \in \mathbb{R})$ of $(x_t, t \in \mathbb{R})$ which is locally Hölder continuous with exponent $\alpha \in (0, \delta/p)$, i.e.

$$P(x_t = \hat{x}_t) = 1, \qquad \forall t \in \mathbb{R}$$

$$\sup_{s \neq t, s, t \in [a,b]} \frac{||\hat{x}(t) - \hat{x}(s)||}{|t-s|^{\alpha}} < \infty,$$

where the supremum is taken over all compact subintervals $[a, b] \subseteq \mathbb{R}$.

Proof See for example, [9]. ■

Theorem 3.1.5 *There exists a continuous modification B_t^H for $\overline{B}_H(t)$ such that B_t^H is Gaussian and (3.1)-(3.2) hold, i.e. B_t^H is a fractional Brownian motion.*

Proof Essentially we only need to check if the Kolmogorov criterion is satisfied. Indeed, since (3.1)-(3.2) hold for $\overline{B}_H(t)$, a direct computation shows that

$$E_{\mu_\phi}\left[(\overline{B}_H(s) - \overline{B}_H(t))^2\right] = |s-t|^{2H}.$$

With $H \in (\frac{1}{2}, 1)$, we can take $p = 2, C = 1$ and $\delta = 2H - 1 (> 0)$ to satisfy the criterion. ■

We are at the stage of defining the integrals with respect to a fBm:

Definition 3.1.6 *Given a non-random function $f \in L^2_\phi(\mathbb{R})$, we can define the integral $\int_{\mathbb{R}} f(t) dB_t^H$ by passing the limit to the integrals $\int_{\mathbb{R}} f_n(t) dB_t^H$, with $f_n(t) \to f(t)$ being a sequence of functions constructed from the following step functions:*

$$f_n(t) = \sum_i a_i^{(n)} \chi_{[t_i, t_{i+1})}(t),$$

and setting

$$\int_{\mathbb{R}} f_n(t) dB_t^H := \sum_i a_i^{(n)} (B_{t_{i+1}}^H - B_{t_i}^H)$$

$$\int_{\mathbb{R}} f(t)dB_t^H := \lim_{n\to\infty} \int_{\mathbb{R}} f_n(t)dB_t^H.$$

In this sense, the dual pairing is the integral of such an f:

$$(\omega, f) = \int_{\mathbb{R}} f(t)dB_t^H.$$

3.2 Integrals with respect to the fBm Process

Here we present some preliminary facts about $\int_{\mathbb{R}} f(t)dB_t^H$, where f is non-random, which are useful for a later discussion about the solutions for a fOU process.

Lemma 3.2.1 (Ito's Isometry) *Given deterministic $f \in L^2_\phi(\mathbb{R})$, we have*

$$E_{\mu_\phi}\left[\left(\int_{\mathbb{R}} f(t)dB_t^H\right)^2\right] = |f|_\phi^2.$$

Proof This is a result that can be obtained immediately from the definition of the probability measure μ_ϕ, i.e. (3.10) after passing the limit to a sequence of simple functions $f_n \to f$. ■

Lemma 3.2.2 *Given $f, g \in L^2_\phi(\mathbb{R})$, the covariance of integrals $\int_{\mathbb{R}} f(t)dB_t^H$ and $\int_{\mathbb{R}} g(t)dB_t^H$ is given by*

$$E_{\mu_\phi}\left[\int_{\mathbb{R}} f(t)dB_t^H \cdot \int_{\mathbb{R}} g(t)dB_t^H\right] = \iint_{\mathbb{R}^2} f(s)g(t)\phi(s,t)dsdt = (f,g)_\phi. \tag{3.13}$$

Proof Since the LHS of (3.13) is simply $E_{\mu_\phi}[(\omega, f)\cdot(\omega, g)]$, the result follows immediately by a polarization argument again:

$$\begin{aligned} E_{\mu_\phi}[(\omega, f)\cdot(\omega, g)] &= \frac{1}{2}E_{\mu_\phi}[((\omega, f+g)^2 - (\omega, f)^2 - (\omega, g)^2] \\ &= \frac{1}{2}\left[|f+g|_\phi^2 - |f|_\phi^2 - |g|_\phi^2\right] = (f,g)_\phi. \end{aligned}$$

■

Recall that Ito's integrals with deterministic integrands under a *standard* Brownian motion are still normally distributed. Usually this is proved by checking if the characteristic functions of the integrals match with that of a normal distribution with a zero mean. The same logic can apply to the integrals under the fBm process:

Lemma 3.2.3 *Given a deterministic function $f \in L^2_\phi(\mathbb{R})$, the Ito's integral $\int_{\mathbb{R}} f(t)dB^H_t$ with respect to a fBm process, as defined in Definition 3.1.6, is normally distributed with zero mean and variance $|f|^2_\phi$.*

Proof Since (3.8) holds for $(\omega, f) = \int_{\mathbb{R}} f(t)dB^H_t$ (by passing the limit for a sequence of functions $f_n \to f$), we can conclude that the characteristic function of $\int_{\mathbb{R}} f(t)dB^H_t$ is simply $\exp(-\frac{1}{2}|f|^2_\phi)$, the latter of which corresponds to the characteristic function of $N(0, |f|^2_\phi)$. ■

The normality feature saves us a lot of work for the bias estimation calculation, for if otherwise we would need to calculate higher order multi-variate moments including co-kurtosis terms.

3.3 A Brief Note on Ito-Wick Calculus

The mathematical treatment becomes delicate when it comes to integrating a stochastic function with respect to a general fBm. Under a standard Brownian motion, an Ito integral, say $\int F(t)dB_t$, can be defined using the following Riemann sum:

$$\sum_i F(t_i) \cdot (B(t_{i+1}) - B(t_i))$$

and such definition will lead to the properties such as

$$E\left[\int F(t)dB_t\right] = 0.$$

However, it is known that under a general fBm, the expected value $E\left[\int F(t)dB^H_t\right]$ is usually **NOT** equal to zero if we simply copy the definition of the standard Brownian motion based on some Riemann sums. Moreover, it is proved in [2] that such a definition is equivalent to the Stratonovich integrals[2] for a large class of functions F.

To ensure that the zero expectation property is still preserved for stochastic integrals under fBm, [2] introduces the so-called Wick-Ito integrals whose definition is based on Riemann sums of some Wick's products.

[2]This is a stochastic integral defined using the following Riemann sum: $\sum_i \frac{F(t_i)+F(t_{i+1})}{2} \cdot (B(t_{i+1}) - B(t_i))$.

Consider a probability space $(\Omega, \mathcal{F}, P^H)$ for a fixed Hurst parameter $H \in (1/2, 1)$. We can define the space of random variables $F : \Omega \to \mathbb{R}$ by

$$L^p := L^p(\Omega, \mathcal{F}, P^H) = \{F : \Omega \to \mathbb{R} | (E|F|^p)^{1/p} < \infty\}$$

for each fixed $p \geq 1$. Define the exponential functions $\epsilon : L^2_\phi \to L^1(\Omega, \mathcal{F}, P)$ by

$$\epsilon(f) := \exp\left\{\int_0^\infty f_t dB_t^H - \frac{1}{2}\int_0^\infty \int_0^\infty f_s f_t \phi(s,t) ds dt\right\}$$

for any $f \in L^2_\phi$. It can be proved that (see [2]) the linear span $\mathcal{E}$ of these exponentials is a dense set of $L^p(\Omega, \mathcal{F}, P)$ for each $p \geq 1$. This fact is crucial for the development of the Wick-Ito integrals.

After that, [2] borrows the idea of Malliavin derivative to define the ϕ-derivative as follows:

Definition 3.3.1 ([2]) *1. For any $g \in L^2_\phi$, define Φg by*

$$(\Phi g)(t) := \int_0^\infty \phi(t,u) g_u du.$$

2. The ϕ-derivative of $F \in L^p(\Omega, \mathcal{F}, P)$ in the direction of Φg is defined as

$$D_{\Phi g} F(\omega) = \lim_{\delta \to 0} \frac{1}{\delta}\left\{F\left(\omega + \delta \int_0^\infty \int_0^\infty \phi(u,v) g(v) dv du\right) - f(\omega)\right\}$$

if such a limit exists in $L^p(\Omega, \mathcal{F}, P)$. Furthermore, if there is a process f_s such that

$$D_{\Phi g} F = \int_0^\infty f_s g_s ds \qquad a.s., \forall g \in L^2_\phi,$$

then F is said to be ϕ-differentiable, $D^\phi F$ is said to exist and f_s is denoted by $D^\phi_s F$, i.e.

$$D_{\Phi g} F(\omega) = \int_0^\infty D^\phi_s F(\omega) g_s ds.$$

Here comes the definition of Wick product $\diamond$. First, we define Wick product for two arbitrary exponentials:

$$\epsilon(f) \diamond \epsilon(g) := \epsilon(f+g). \tag{3.14}$$

Since exponentials span the linear space $\mathcal{E}$, (3.14) can be easily extended to the definition of $F \diamond G$ for any $F, G \in \mathcal{E}$.

In general, $\int_0^\infty g_s dB_s^H$ does not belong to $\mathcal{E}$. As a result, further extension to (3.14) is required in order to define Wick products on general integrals of the form $\int_0^\infty g_s dB_s^H$ for $g \in L_\phi^2$.

Lemma 3.3.2 ([2]) *If $f, g \in L_\phi^2$, then*

$$\epsilon(f) \diamond \int_0^\infty g_s dB_s^H = \epsilon(f) \int_0^\infty g_s dB_s^H - D_{\Phi g}\epsilon(f). \tag{3.15}$$

Proof The lemma follows by differentiating $\epsilon(f) \diamond \epsilon(\delta g) = \epsilon(f + \delta g)$ with respect to δ and evaluating the equality at $\delta = 0$. Notice that by the definition of ϕ-derivative, we have $D_{\Phi g}\epsilon(f) = \epsilon(f) \int_0^\infty \int_0^\infty \phi(s,t) f_s g_t ds dt$. ■

Theorem 3.3.3 (Proposition 3.4 in [2]) *If $g \in L_\phi^2$, and suppose $F, D_{\Phi g}F \in L^2(\Omega, \mathcal{F}, P)$, then*

$$F \diamond \int_0^\infty g_s dB_s^H = F \int_0^\infty g_s dB_s^H - D_{\Phi g}F. \tag{3.16}$$

Proof Extend the result in Theorem 3.3.2 to any $F \in \mathcal{E}$, then the extension to $F \in L^2(\Omega, \mathcal{F}, P)$ follows by a continuity argument. ■

An extension of Ito's isometry can be obtained for $F \diamond \int_0^\infty g_s dB_s^H$:

Theorem 3.3.4 ([2]) *Assume that $g \in L_\phi^2$ and $F \in \mathcal{E}$. Then*

$$E\left(F \diamond \int_0^\infty g_s dB_s^H\right)^2 = E\left[(D_{\Phi g}F)^2 + F^2 |g|_\phi^2\right]. \tag{3.17}$$

Proof As before, we can derive the equality for the case when $F = \epsilon(f)$, then extend to $F \in \mathcal{E}$. ■

We can now give the definition of $\int_0^T F_s \delta B_s^H$ in the Wick-Ito's sense:

Definition 3.3.5 ([2]) *Let $F \in L^2(\Omega, \mathcal{F}, P)$ and consider an arbitrary partition π of $[0,T]$ with $0 < t_0 < t_1 < \cdots < t_n = T$. Define the Riemann sum*

$$S_\pi = \sum_{i=0}^{n-1} F_{t_i} \diamond (B^H_{t_{i+1}} - B^H_{t_i}).$$

Denote $|\pi| = \max_i(t_{i+1} - t_i)$ and $F^\pi_t := F_{t_i}$ for $t \in [t_i, t_{i+1})$. Suppose that as $|\pi| \to 0$, we have $E|F^\pi - F|^2_\phi \to 0$ and

$$\sum_{i=0}^{n-1} E\left[\int_{t_i}^{t_{i+1}} |D^\phi_s F_{t_i} - D^\phi_s F_s| ds\right]^2 \to 0 \qquad in\ L^2,$$

then the Riemann sum has a limit in $L^2(\Omega, \mathcal{F}, P)$ and is denoted as $\int_0^T F_s \delta B^H_s$, i.e.

$$\int_0^T F_s \delta B^H_s = \lim_{|\pi| \to 0} \sum_{i=0}^{n-1} F^\pi_{t_i} \diamond (B^H_{t_{i+1}} - B^H_{t_i}).$$

Denote $\mathcal{L}(0,T)$ as the set of stochastic processes F on $[0,T]$ such that $\int_0^T F_s \delta B^H_s$ is well-defined.

The Wick-Ito integral as defined above preserves several nice properties in the standard Brownian motion:

$$E\left[\int_0^T F_S \delta B^H_s\right] = 0 \tag{3.18}$$

$$E\left[\int_0^T F_S \delta B^H_s\right]^2 = E\left[\left(\int_0^T D^\phi_s F_s ds\right)^2 + |F|^2_\phi\right] \tag{3.19}$$

By (3.19), if F is deterministic or F satisfies $D^\phi_s F_s ds = 0$ for $s \in [0,T]$, then

$$E\left[\int_0^T F_S \delta B^H_s\right]^2 = |F|^2_\phi,$$

which resembles the Ito's isometry in standard Ito's integral.

The relation between Wick-Ito and Stratonovich integrals is given by Theorem 3.9 in [2], retrieved here:

Theorem 3.3.6 *If $F \in \mathcal{L}(0,T)$, then*

$$\int_0^T F_s dB_s^H = \int_0^T F_s \delta B_s^H + \int_0^T D_s^\phi F_s ds \qquad a.s.$$

where $\int_0^T F_s dB_s^H$ denotes the Stratonovich integral.

Note that when F is deterministic, then the two types of integrals coincide. In next chapter, we will solely deal with integrals of deterministic functions, and hence we will not distinguish these two types of integrals unless ambiguity arises.

Finally, we state without proof the Ito's lemma for a general fBm:

Theorem 3.3.7 ([2]) *Suppose that $F_u, u \in [0,T]$ is a stochastic process in $\mathcal{L}(0,T)$ satisfying the following regularity conditions:*

- *There exists $\alpha > 1 - H$ and $\delta > 0$, such that for all u, v such that $|u - v| \leq \delta$,*

$$E|F_u - F_v|^2 \leq C|u - v|^{2\alpha}.$$

- $\lim\limits_{0 \leq u,v \leq t, |u-v| \to 0} E|D_u^\phi (F_u - F_v)|^2 = 0.$

Also suppose that $E[\sup_{s \in [0,T]} |G_s|] < \infty$ and denote $\eta_t = \xi + \int_0^t G_u du + \int_0^t F_u \delta B_u^H$ with $\xi \in \mathbb{R}$ and $\frac{\partial f}{\partial x}(s, \eta_s) F_s \in \mathcal{L}(0,T)$. Then for all $t \in [0,T]$,

$$\begin{aligned} f(t, \eta_t) = f(0, \xi) &+ \int_0^t \frac{\partial f}{\partial s}(s, \eta_s) ds + \int_0^t \frac{\partial f}{\partial x}(s, \eta_s) G_s ds \\ &\int_0^t \frac{\partial f}{\partial x}(s, \eta_s) F_s \delta B_s^H + \int_0^t \frac{\partial^2 f}{\partial x^2}(s, \eta_s) F_s D_s^\phi \eta_s ds \qquad a.s. \end{aligned} \tag{3.20}$$

The proof of Ito's lemma can be found in [2]. Here, we consider only an application to this lemma to a particular function: $f(t, \eta_t) := e^{kt}\eta_t, k \in \mathbb{R}$, which is relevant to the next Chapter. Since $\dfrac{\partial f}{\partial t} = ke^{kt}\eta_t$, $\dfrac{\partial f}{\partial x} = e^{kt}$ and $\dfrac{\partial^2 f}{\partial x^2} = 0$, a direct application to the Ito's lemma gives

$$f(t, \eta_t) = \eta_0 + \int_0^t ke^{ks}\eta_s ds + \int_0^t e^{ks} G_s ds + \int_0^t e^{ks} F_s \delta B_s^H + 0.$$

In particular, if F_t is a deterministic function, then $\int_0^t e^{ks} F_s \delta B_s^H = \int_0^t e^{ks} F_s dB_s^H$ and hence

$$f(t, \eta_t) = \eta_0 + \int_0^t k e^{ks} \eta_s ds + \int_0^t e^{ks} G_s ds + \int_0^t e^{ks} F_s dB_s^H.$$

It is in this sense that we can, with some abuse of notation, write the above equality in differential form:

$$\begin{aligned} df(t, \eta_t) &= k e^{kt} \eta_t dt + e^{kt}(G_t dt + F_t dB_t^H) = \eta_t k e^{kt} dt + e^{kt} d\eta_t \\ &= \eta_t d(e^{kt}) + e^{kt} d\eta_t, \end{aligned}$$

which retrieves the usual product rule. It should be reminded that such a formulation holds only for some specific cases, such as when F_t (i.e. the coefficient of the volatility term in η_t) is deterministic.

Chapter 4

Bias Estimation for a Fractional Ornstein-Uhlenbeck Process

In this chapter, we derive the second order bias for the OLS estimate of the mean reversion parameter for the fractional Brownian process with $\frac{1}{2} < H < 1$. It turns out that most part of the work rests on computation of covariance of fractional Ito's integrals.

This chapter is divided into several sections. First, the covariance matrix involved in the bias calculation will be derived. Then, the theoretical bias formula is compared against the actual bias obtained from Monte-Carlo simulation. Afterwards, some observations, as well as the implications from the perspective of risk modeling, related to the estimate of mean reversion parameter for a fOU process are given.

4.1 Introduction

Recall the stochastic differential equation for a fractional Ornstein-Uhlenbeck process:

$$dS_t = k(\mu - S_t)dt + \sigma dB_t^H. \tag{4.1}$$

As mentioned in Chapter 3, the solution to the above SDE can be obtained in a similar (formally speaking) as in the standard OU process. First, Ito's product rule states that $d(e^{kt}S_t) = e^{kt}dS_t + ke^{kt}S_tdt$ and hence we can multiply the integrating factor e^{kt} to (4.1) to get

$$dS_t + kS_tdt = k\mu dt + \sigma dB_t^H$$

$$\Rightarrow \qquad d\left(e^{kt}S_t\right) = k\mu e^{kt}dt + \sigma e^{kt}dB_t^H.$$

In practice data are collected at discrete time steps. As a result, we can assume that these data are recorded in evenly spaced time intervals, i.e. $S_i := S(t_i), i = 0, 1, \cdots, n$, with $t_i = i \cdot h$ where $h > 0$ is fixed. Integrating the above SDE over $[t_{i-1}, t_i]$ gives

$$e^{kt_i}S_i - e^{kt_{i-1}}S_{i-1} = \mu\left(e^{kt_i} - e^{kt_{i-1}}\right) + \sigma\int_{t_{i-1}}^{t_i} e^{ks}dB_s^H$$

$$\text{or} \qquad S_i = e^{-kh}S_{i-1} + \mu(1 - e^{-kh}) + \sigma e^{-kh}\int_{t_{i-1}}^{t_i} e^{ks}dB_s^H. \tag{4.2}$$

Without loss of generality, we can from now on assume that $\mu = 0$ and consider the following solution to (4.1):

$$S_i = e^{-kh}S_{i-1} + \sigma e^{-kh}\int_{t_{i-1}}^{t_i} e^{ks}dB_s^H. \tag{4.3}$$

It should be emphasized that the error terms

$$\epsilon_i := \int_{t_{i-1}}^{t_i} e^{ks}dB_s^H$$

are in general serially correlated for $H \neq \frac{1}{2}$. However, as mentioned in Chapter 3, they are still normally distributed. As a result, even in this generalized situation of fOU process, we are still free from the concern of computing co-kurtosis terms. In particular, by Lemma 2.2.1, the quadratic terms involved in the computation of the (second order) bias formula depends only on the covariance matrix. This reduces our calculation to the computations of the covariance of S_i and S_j, where $i, j = 1, \cdots, n$.

4.2 Computation of the Covariance Terms

If we compute the covariance terms directly from (4.3), we can only arrive at an iterative expression defining these covariances because under the general fBm framework, S_{i-1} is correlated with the error term ϵ_i. Correlation occurs because S_{i-1} also contains other error terms (which are ϵ_{i-1}, and other ϵ's implicitly implied in the recursive formula (4.3)) and as mentioned above, all of these error terms are correlated with ϵ_i.

Hence, unlike the treatment in [11], it is more convenient to express S_{t_i} by integrating (4.1) over $[-\infty, t_i]$, i.e.

$$\int_{-\infty}^{t_i} d(e^{ks}S_s) = \sigma \int_{-\infty}^{t_i} e^{ks} dB_s^H$$
$$S_i = \sigma e^{-kih} \int_{-\infty}^{ih} e^{ks} dB_s^H. \tag{4.4}$$

It suffices to compute

$$c_{i,j} := E_{\mu_\phi}\left[\sigma e^{-kih} \int_{-\infty}^{ih} e^{ks} dB_s^H \cdot \sigma e^{-kjh} \int_{-\infty}^{jh} e^{kt} dB_t^H\right]$$

or equivalently, the following expression

$$I(\alpha,\beta) := E_{\mu_\phi}\left[\int_{-\infty}^{\alpha} e^{ks} dB_s^H \int_{-\infty}^{\beta} e^{kt} dB_t^H\right], \qquad \alpha, \beta \geq 0.$$

By Lemma 3.2.2, $I(\alpha,\beta)$ in turn reduces to

$$I(\alpha,\beta) = \int_{-\infty}^{\alpha}\int_{-\infty}^{\beta} e^{k(u+v)}\phi(u,v)dvdu = H(2H-1)\int_{-\infty}^{\alpha}\int_{-\infty}^{\beta} e^{k(u+v)}|u-v|^{2H-2}dvdu.$$

Simplifying $\iota(\alpha,\beta) := \int_{-\infty}^{\alpha}\int_{-\infty}^{\beta} e^{k(u+v)}|u-v|^{2H-2}dvdu$ is straightforward but tedious. First, due to symmetry we can assume $\alpha \geq \beta (> 0)$ without loss of generality. Then, apply the following change of variables:

$$\begin{cases} s = u + v \\ t = u - v \end{cases}$$

so that $dudv = \frac{1}{2}dsdt$ and ι becomes

$$\iota(\alpha,\beta) = \frac{1}{2}\iint_A e^{ks}|t|^{2H-2}dsdt = \frac{1}{2}\left[\iint_{A_+} e^{ks}t^{2H-2}dsdt + \iint_{A_-} e^{ks}|t|^{2H-2}dsdt\right],$$

where $A := A_+ \cup A_-$ and A_+, A_- are 2-dimensional regions defined as in Figure 4.1.

Lemma 4.2.1 *For any fixed $k > 0$, $H \in (\frac{1}{2}, 1)$ and $\alpha \geq \beta \geq 0$, we have*

$$\iint_{A_+} e^{ks}t^{2H-2}dsdt = \frac{1}{2H-1}\left[e^{2k\alpha}\int_{\alpha-\beta}^{\infty} e^{-kt}t^{2H-1}dt - e^{2k\beta}\int_{0}^{\alpha-\beta} e^{kt}t^{2H-1}dt\right]$$
$$\iint_{A_-} e^{ks}|t|^{2H-2}dsdt = \frac{e^{2k\beta}}{2H-1}\int_0^{\infty} e^{-kt}t^{2H-1}dt.$$

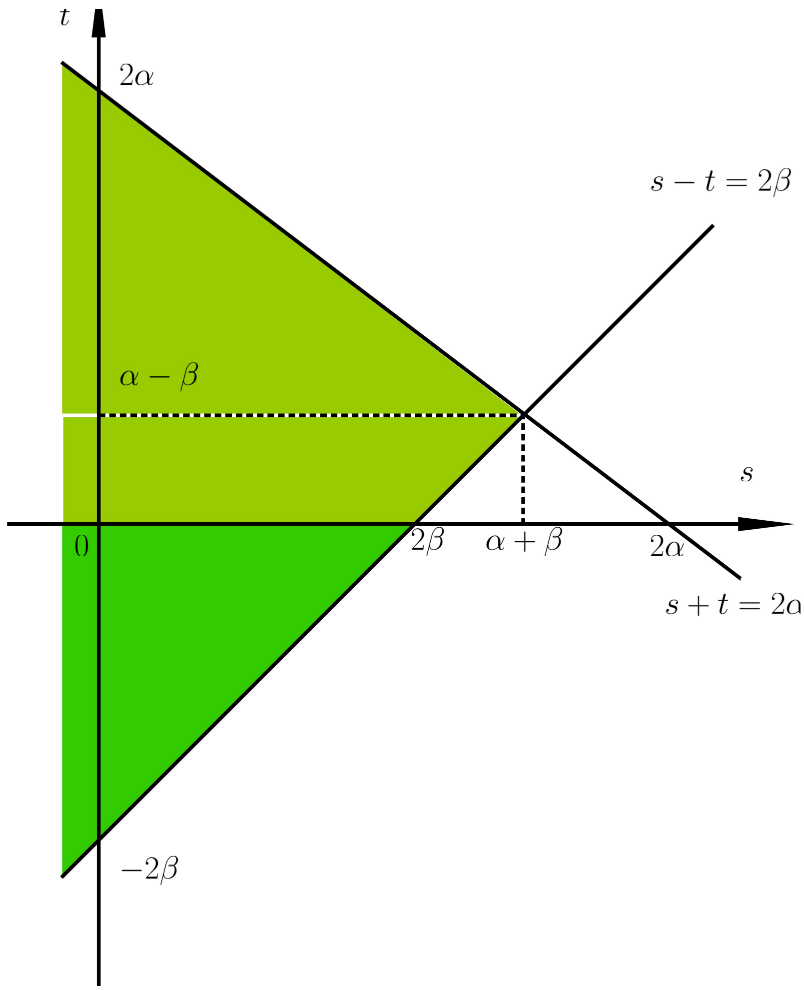

Figure 4.1: Region of integration, with A_+ in pale green and A_- in bright green.

Proof The basic idea is to simplify the innermost integral with respect to s, followed by an integration by part so as to raise the power of t from $2H-2$ to $2H-1$:

$$\begin{aligned}\iint_{A_+} e^{ks}t^{2H-2}dsdt &= \int_{\alpha-\beta}^{\infty}\int_{-\infty}^{2\alpha-t} e^{ks}t^{2H-2}dsdt + \int_{0}^{\alpha-\beta}\int_{-\infty}^{2\beta+t} e^{ks}t^{2H-2}dsdt \\ &= \frac{1}{k}\left[\int_{\alpha-\beta}^{\infty} t^{2H-2}e^{k(2\alpha-t)}dt + \int_{0}^{\alpha-\beta} t^{2H-2}e^{k(2\beta+t)}dt\right]\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{k(2H-1)}\left[\int_{\alpha-\beta}^{\infty} e^{k(2\alpha-t)} d(t^{2H-1}) + \int_{0}^{\alpha-\beta} e^{k(2\beta+t)} d(t^{2H-1})\right] \\
&= \frac{1}{k(2H-1)}\left[e^{k(2\alpha-t)} t^{2H-1}\Big|_{\alpha-\beta}^{\infty} + k\int_{\alpha-\beta}^{\infty} e^{k(2\alpha-t)} t^{2H-1} dt \right.\\
&\qquad \left. + e^{k(2\beta+t)} t^{2H-1}\Big|_{0}^{\alpha-\beta} - k\int_{0}^{\alpha-\beta} e^{k(2\beta+t)} t^{2H-1} dt\right] \\
&= \frac{1}{2H-1}\left[e^{2k\alpha}\int_{\alpha-\beta}^{\infty} e^{-kt} t^{2H-1} dt - e^{2k\beta}\int_{0}^{\alpha-\beta} e^{kt} t^{2H-1} dt\right]
\end{aligned}
$$

$$
\begin{aligned}
\iint_{A_-} e^{ks}|t|^{2H-2} ds dt &= \int_{-\infty}^{0}\int_{-\infty}^{2\beta+t} e^{ks}|t|^{2H-2} ds dt \\
&= \frac{1}{k}\int_{-\infty}^{0} e^{ks}\Big|_{-\infty}^{2\beta+t} |t|^{2H-2} dt = \frac{1}{k}\int_{-\infty}^{0} e^{k(2\beta+t)}|t|^{2H-2} dt.
\end{aligned}
$$

Now to eliminate the absolute sign in the above integral, we introduce a dummy variable $\tau = -t$ so that $|t| = -t = \tau$ and

$$
\begin{aligned}
\iint_{A_-} e^{ks}|t|^{2H-2} ds dt &= \frac{1}{k}\int_{\infty}^{0} e^{k(2\beta-\tau)}\tau^{2H-2}(-d\tau) = \frac{1}{k}\int_{0}^{\infty} e^{k(2\beta-\tau)}\tau^{2H-2} d\tau \\
&= \frac{1}{k(2H-1)}\int_{0}^{\infty} e^{k(2\beta-\tau)} d(\tau^{2H-1}) \\
&= \frac{1}{k(2H-1)}\left[e^{k(2\beta-\tau)}\tau^{2H-1}\Big|_{0}^{\infty} + k\int_{0}^{\infty} e^{k(2\beta-\tau)}\tau^{2H-1} d\tau\right] \\
&= \frac{e^{2k\beta}}{2H-1}\int_{0}^{\infty} e^{-k\tau}\tau^{2H-1} d\tau.
\end{aligned}
$$

It should be noted that in the valuation of upper and lower limits it is necessary to employ the fact that e^{-t} decays at a much faster rate than the rate at which t^{2H-1} increases. ■

Remark: By raising the power of t from $2H-2$ to $2H-1$ by integration by part, it helps avoid the $0 \cdot \infty$ indeterminate form when we consider the behaviour of the covariance terms $I(\alpha, \beta)$ when $H \to \frac{1}{2}^{+}$, and provide some numerical stability when we develop numerical schemes based on the above expressions.[1]

[1] It should be reminded that if integration by part is not done here, it is *incorrect* to directly substitute $H = \frac{1}{2}$ to obtain the covariance terms under the standard Brownian motion case; indeed by so doing we will erroneously get 0 for all covariance terms because they have a $2H-1$ factor, which is zero when $H = 1/2$.

From the above lemma, it becomes clear that the covariance terms are related to incomplete gamma functions:

Definition 4.2.2 *Given any fixed $s, x \in \mathbb{R}$, the upper and lower incomplete gamma functions are defined as*

$$\gamma(s,x) = \int_0^x t^{s-1}e^{-t}dt,$$
$$\Gamma(s,x) = \int_x^\infty t^{s-1}e^{-t}dt.$$

In particular, we always have $\gamma(s,x) + \Gamma(s,x) = \Gamma(s)$, where $\Gamma(s)(= \Gamma(s,0))$ is the gamma function.

Lemma 4.2.3 *For $k > 0$, $H \in (\frac{1}{2}, 1)$, $x \in \mathbb{R}$, we have*

$$\int_x^\infty e^{-kt}t^{2H-1}dt = \frac{1}{k^{2H}}\Gamma(2H, kx)$$
$$\int_0^x e^{kt}t^{2H-1}dt = \frac{1}{(-k)^{2H}}\gamma(2H, -kx).$$

Proof Straightforward exercise. ■

Theorem 4.2.4 *For $\alpha \geq \beta \geq 0$,*

$$I(\alpha,\beta) = \frac{H}{2k^{2H}}\left[e^{2k\alpha}\Gamma(2H, k(\alpha-\beta)) + e^{2k\beta}\left(\Gamma(2H) - \frac{\gamma(2H, -k(\alpha-\beta))}{(-1)^{2H}}\right)\right]$$

Moreover, we have $E_{\mu_\phi}[S_\alpha \cdot S_\beta] = \sigma^2 e^{-k(\alpha+\beta)}I(\alpha,\beta)$, leading to

$$E_{\mu_\phi}[S_\alpha \cdot S_\beta] = \frac{H\sigma^2}{2k^{2H}}\left[e^{k(\alpha-\beta)}\Gamma(2H, k(\alpha-\beta)) + e^{-k(\alpha-\beta)}\left(\Gamma(2H) - \frac{\gamma(2H, -k(\alpha-\beta))}{(-1)^{2H}}\right)\right] \quad (4.5)$$

When $\alpha = \beta$, the variance term is given by

$$E_{\mu_\phi}[S_\alpha^2] = \frac{H\sigma^2\Gamma(2H)}{k^{2H}}. \quad (4.6)$$

The reason why this is incorrect is because, one of the integrals, namely $\int_0^{\alpha-\beta} t^{2H-2}e^{k(2\beta+t)}dt$, will blow up when $H \to \frac{1}{2}^+$, leading to a $0 \cdot \infty$ indeterminate form when it is multiplied by the $2H-1$ factor.

Thus, the correlation term is given by

$$
\begin{aligned}
corr(S_\alpha, S_\beta) &= \frac{E_{\mu_\phi}[S_\alpha \cdot S_\beta]}{\sqrt{E_{\mu_\phi}[S_\alpha^2] E_{\mu_\phi}[S_\beta^2]}} \\
&= \frac{1}{2\Gamma(2H)} \left[e^{k(\alpha-\beta)} \Gamma(2H, k(\alpha-\beta)) + e^{-k(\alpha-\beta)} \left(\Gamma(2H) - \frac{\gamma(2H, -k(\alpha-\beta))}{(-1)^{2H}} \right) \right].
\end{aligned}
$$

Remark:

1. By exchanging α and β, (4.5) implies that the covariance terms is always a function of $|\alpha - \beta|$. We can write

$$E_{\mu_\phi}[S_\alpha \cdot S_\beta] = C(|\alpha - \beta|), \qquad \forall \alpha, \beta \geq 0,$$

   where $C(|\alpha-\beta|)$ is the RHS of (4.5), with $\alpha - \beta$ replaced by $|\alpha - \beta|$.

2. As a check, it is worthwhile to consider the case when $H \to \frac{1}{2}^+$. Since by definition $\Gamma(1, x) = e^{-x}$ and $\gamma(1, x) = 1 - e^{-x}$, (4.5) and (4.6) will be reduced to

$$
\begin{aligned}
E[S_\alpha \cdot S_\beta] &\to \frac{\sigma^2}{4k} \left[e^{k(\alpha-\beta)} e^{-k(\alpha-\beta)} + e^{-k(\alpha-\beta)} \left(1 - \frac{1 - e^{k(\alpha-\beta)}}{-1} \right) \right] = \frac{\sigma^2}{2k} e^{-k(\alpha-\beta)} \\
E[S_\alpha^2] &\to \frac{\sigma^2}{2k},
\end{aligned}
$$

   which matches with the facts regarding the standard Ornstein-Uhlenbeck process.

3. From (4.5) and (4.6), when $k \to 0^+$, all variance and covariance terms will tend to infinity because of the presence of k^{2H} in the denominator of the equations (while the numerator is still bounded).

4. Using L'Hospital'rule, both $e^{k(\alpha-\beta)}\Gamma(2H, k(\alpha-\beta))$ and $e^{-k(\alpha-\beta)}\frac{\gamma(2H, -k(\alpha-\beta))}{(-1)^{2H}}$ will approach $(k(\alpha-\beta))^{2H-1}$ as $k \to \infty$ and $\alpha > \beta$. Hence, $E[S_\alpha \cdot S_\beta] \to \frac{H\sigma^2}{2k^{2H}} e^{-k(\alpha-\beta)}\Gamma(2H)$, i.e. exponentially decaying when $k \to \infty$. In other words, from the perspective of the covariance of S_t, the behaviour of fOU process will look more "alike" to that of the standard OU process when k is large.

5. As to the computational aspect, many programming languages have library support to compute the incomplete gamma functions numerically. For instance, MATLAB has a

`gammainc()` function callto calculate the "normalized" incomplete gamma functions, i.e.

$$\Gamma_n(s,x) := \Gamma(s,x)/\Gamma(s)$$
$$\gamma_n(s,x) := \gamma(s,x)/\Gamma(s)$$

The C++ Boost package also includes a gamma.hpp to calculate these special functions.

6. Recall that we have defined

$$c_{i,j} := E_{\mu_\phi}[S_{ih} \cdot S_{jh}] = E_{\mu_\phi}\left[\sigma e^{-kih}\int_{-\infty}^{ih} e^{ks} dB_s^H \cdot \sigma e^{-kjh}\int_{-\infty}^{jh} e^{kt} dB_t^H\right]$$

By (4.5), we have $c_{i,j} = C(h|i-j|)$, where

$$C(x) := \frac{H\sigma^2}{2k^{2H}}\left[e^{kx}\Gamma(2H,kx) + e^{-kx}\left(\Gamma(2H) - \frac{\gamma(2H,-kx)}{(-1)^{2H}}\right)\right]. \tag{4.7}$$

4.3 Expectation of Stochastic Quadratic Forms

From Chapter 2, we know that in order to arrive at the estimation bias formula for the mean reverting parameter of the fOu process, we need to compute $E(U_n), E(V_n), E(U_n^2)$ and $E(V_n^2)$ where

$$U_n = \frac{1}{n}\sum_{i=1}^n S_{i-1}S_i, \qquad V_n = \frac{1}{n}\sum_{i=1}^n S_{i-1}^2.$$

Using (4.5) and (4.6), the following results are immediate:

Theorem 4.3.1 *Define $C(x)$ as in (4.7), then*

$$E_{\mu_\phi}[U_n] = \frac{1}{n}\sum_{i=1}^n c_{i,i-1} = C(h)$$
$$E_{\mu_\phi}[V_n] = \frac{1}{n}\sum_{i=1}^n c_{i-1,i-1} = C(0)\left(= \frac{H\sigma^2\Gamma(2H)}{k^{2H}}\right).$$

The above theorem immediately implies that for general $H \neq \frac{1}{2}$,

$$E[U_n] - e^{-kh}E[V_n] \neq 0,$$

since $C(h)/C(0) \neq e^{-kh}$. In other words, $E[a_{-\frac{1}{2}}]$ is never zero for a general fOU process. Nevertheless, for $k \approx 0$, we can still have the following asymptotic result:

Lemma 4.3.2 *When $k \to 0$, we have*

$$\frac{E[U_n]}{E[V_n]} = \frac{C(h)}{C(0)} = \frac{1}{2}\left[e^{kh}\Gamma_n(2H, kh) + e^{-kh}\left(1 - \frac{\gamma_n(2H, -kh)}{(-1)^{2H}}\right)\right] \to e^{-kh}.$$

for $H > \frac{1}{2}$.

Proof The result is immediate when we go back to the definition of the incomplete gamma functions. First,

$$e^{kh}\Gamma(2H, kh) = e^{kh}\int_{kh}^{\infty} t^{2H-1}e^{-t}dt = e^{kh}\int_0^{\infty}(y + kh)^{2H-1}e^{-y-kh}dy,$$

by a change of variable $y := t - kh$. As a result,

$$e^{kh}\Gamma(2H, kh) = \int_0^{\infty}(y + kh)^{2H-1}e^{-y}dy \approx \int^{\infty} y^{2H-1}e^{-y}dy = \Gamma(2H),$$

as $kh \to 0$. Thus, $e^{kh}\Gamma_n(2H, kh) \to 1$.

For $\gamma_n(2H, -kh)$, observe that

$$\frac{\gamma(2H, -kh)}{(-1)^{2H}} = \frac{\int_0^{-kh} t^{2H-1}e^{-t}dt}{(-1)^{2H}} = \int^{kh} \tau^{2H-1}e^{\tau}d\tau,$$

by a change of variable $\tau := -t$. Now, when $k \approx 0$, $\tau^{2H-1}e^{\tau} \approx e^{\tau}$ for all $\tau \in [0, kh]$ and $H > \frac{1}{2}$, hence

$$\frac{\gamma(2H, -kh)}{(-1)^{2H}} \to \int_0^{kh} e^{\tau}d\tau = e^{kh} - 1$$

$$\Rightarrow \quad \frac{C(h)}{C(0)} \to \frac{1}{2}\left[1 + e^{-kh}\left(1 - (e^{kh} - 1)\right)\right] = e^{-kh}. \qquad \blacksquare$$

Computation of the quadratic forms $E(U_n^2)$ and $E(V_n^2)$ is more involved but still straightforward. We start with the characteristic function of fBm integrals with deterministic integrands, as discussed in Chapter 3:

$$E_{\mu_\phi}\left[e^{i\int_{\mathbb{R}} F(s)dB_s^H}\right] = e^{-\frac{1}{2}|F|_\phi^2} \tag{4.8}$$

Theorem 4.3.3 *For any deterministic $f, g, p, q \in L_\phi^2(\mathbb{R})$, we have*

1. $E_{\mu_\phi}\left[\left(\int_{\mathbb{R}} f(s)dB_s^H\right)^4\right] = 3|f|_\phi^4$.
2. $E_{\mu_\phi}\left[\left(\int_{\mathbb{R}} f(s)dB_s^H\right)^2\left(\int_{\mathbb{R}} g(s)dB_s^H\right)^2\right] = |f|_\phi^2|g|_\phi^2 + 2(f,g)_\phi^2$.
3. $E_{\mu_\phi}\left[\int_{\mathbb{R}} f(s)dB_s^H \int_{\mathbb{R}} g(s)dB_s^H \int_{\mathbb{R}} p(s)dB_s^H \int_{\mathbb{R}} q(s)dB_s^H\right] = (f,g)_\phi(p,q)_\phi + (g,p)_\phi(f,q)_\phi + (f,p)_\phi(g,q)_\phi$.

Proof

1. Substitute $F(s) = tf(s)$ for some fixed $t \in \mathbb{R}$ in (3.12). Then the t^4-term of the Taylor series expansion of both sides of (4.8) gives $\frac{1}{24}E\left[\left(\int_{\mathbb{R}} f(s)dB_s^H\right)^4\right] = \frac{1}{2}\left(-\frac{1}{2}|f|_\phi^2\right)^2$, hence the results.

2. Based on the result in the 1st bullet, for any fixed $t \in \mathbb{R}$, we have

$$E\left[\left(\int_{\mathbb{R}}(f(s)+tg(s))dB_s^H\right)^4\right] = 3|f+tg|_\phi^4.$$

   Considering the t^2-terms of both sides of the equation will give the results.

3. The 2nd bullet implies that

$$E\left[\left(\int(f+sg)dB\right)^2\left(\int(p+tq)dB\right)^2\right] = |f+sg|^2|p+tq|^2 + 2(f+sg,p+tq)^2,$$

   for any fixed $s, t \in \mathbb{R}$, and subscripts/superscripts/arguments are omitted whenever understood without causing any confusion. Then comparison of the st-terms of both sides will give the results. Notice that by definition and linearity $(f+sg,p+tq)^2 = ((f,p)+s(g,p)+t(f,q)+st(g,q))^2$. ■

Remark: The above theorem essentially states that due to the Gaussian nature of fBm, any higher order moments can always be expressed in terms of the second order moment. If this Gaussian nature was not present (e.g. in CEV process), the above computation of the quadratic forms would become much more tedious.

Theorem 4.3.4 *Define $C(x)$ as in (4.7), then*

$$E[U_n^2] = C(h)^2 + \frac{1}{n^2}\left(nC(0)^2 + nC(h)^2 + 2\sum_{i=1}^{n-1}(n-i)\Big[C((i+1)h)C((i-1)h) + C(ih)^2\Big]\right)$$

$$E[V_n^2] = C(0)^2 + \frac{2}{n^2}\left(nC(0)^2 + 2\sum_{i=1}^{n-1}(n-i)C(ih)^2\right).$$

Proof We first calculate $E[V_n^2]$. Based on the 2nd bullet of Theorem 4.3.3, we have

$$\begin{aligned}
E[V_n^2] &= \frac{1}{n^2}E\left[\left(\sum_{i=1}^{n} S_{i-1}^2\right)^2\right] \\
&= \frac{1}{n^2}\sum_{i,j=1}^{n} E\left[S_{i-1}^2 S_{j-1}^2\right] = \frac{1}{n^2}\sum_{i,j=1}^{n}\left[E\left[S_{i-1}^2\right]E\left[S_{j-1}^2\right] + 2E\left[S_{i-1}\cdot S_{j-1}\right]\right] \\
&= \frac{1}{n^2}\sum_{i,j=1}^{n}\left[c_{i-1,i-1}c_{j-1,j-1} + 2c_{i-1,j-1}^2\right] = \frac{1}{n^2}\sum_{i,j=1}^{n}\left[C(0)^2 + 2C(h|i-j|)^2\right] \\
&= C(0)^2 + \frac{2}{n^2}\sum_{i,j=1}^{n} C(h|i-j|)^2.
\end{aligned}$$

By counting the number of (i,j), such that $|i-j| = 0, 1, 2, 3, \cdots$, the last summation is equal to $n \cdot C(0)^2 + 2(n-1)C(h)^2 + 2(n-2)C(2h)^2 + \cdots + 2C((n-1)h)^2$, and hence the result.

Now we apply the 3rd bullet to compute $E[U_n^2]$:

$$\begin{aligned}
E[U_n^2] &= \frac{1}{n^2}E\left[\left(\sum_{i=1}^{n} S_i S_{i-1}\right)^2\right] = \frac{1}{n^2}\sum_{i,j=1}^{n} E\left[S_i S_{i-1} S_j S_{j-1}\right] \\
&= \frac{1}{n^2}\sum_{i,j=1}^{n}\left[c_{i,i-1}c_{j,j-1} + c_{i-1,j}c_{i,j-1} + c_{i,j}c_{i-1,j-1}\right]
\end{aligned}$$

$$= \frac{1}{n^2} \sum_{i,j=1}^{n} \left[C(h)^2 + C(h|i-j-1|)C(h|i-j+1|) + C(h|i-j|)^2 \right]$$

$$= C(h)^2 + \frac{1}{n^2} \Big[\underbrace{\sum_{i,j=1}^{n} C(h|i-j-1|)C(h|i-j+1|)}_{\text{denoted as } Expr_1} + \underbrace{\sum_{i,j=1}^{n} C(h|i-j|)^2}_{\text{denoted as } Expr_2} \Big]$$

From above, we know that $Expr_2 = nC(0)^2 + 2\sum_{i=1}^{n-1}(n-i)C(ih)^2$ while a similar counting argument for $Expr_1$ will give $Expr_1 = nC(1)^2 + 2\sum_{i=1}^{n-1}(n-i)C((i-1)h)C((i+1)h)$. ■

Now, based on a similar calculation as described in Chapter 2, we can present the bias formula for the fractional Ornstein-Uhlenbeck process:

Theorem 4.3.5 *Given a time series $\{S_i\}_{0 \le i \le n}$ (equally spaced by $h > 0$) whose dynamics is governed by a fractional Ornstein-Uhlenbeck process $dS_t = -kS_t dt + \sigma dB_t^H$ $(k > 0)$, the second-order bias formula of using OLS estimate for k is given by*

$$Bias(k) = E[\hat{k}] - k \approx a_{-1/2} + a_{-1},$$

where $a_{-1/2}$ and a_{-1} are defined by

$$a_{-1/2} = -\frac{E[U_n] - e^{-kh}E[V_n]}{he^{-kh}E[V_n]},$$
$$a_{-1} = \frac{E[U_n^2] - e^{-2kh}E[V_n^2]}{2he^{-kh}(E[V_n])^2} + a_{-1/2},$$

with the expectations $E[U_n], E[V_n], E[U_n^2], E[V_n^2]$ being calculated using Theorem 4.3.1 and 4.3.4.

4.4 Monte Carlo Simulation

To confirm our theoretical results, we compare the bias formula as described in Theorem 4.3.5 against the empirical bias we would get from the OLS estimate from some simulated fOU paths. In particular, our work here is an extension of [11], and includes it as a special case (by setting Hurst parameter to be $H = \frac{1}{2}$).

We adopt the same simulation scheme as described in [11]. In particular, for each fixed $H \in (\frac{1}{2}, 1)$ and true mean reversion parameter $k > 0$, we simulate 10000 paths based on the solutions as shown in (4.4) for the fractional Ornstein-Uhlenbeck process, and compute for each path the difference between the OLS estimate and k. The (empirical) estimation bias is then obtained by averaging these differences over each path. This bias is also compared against the "theoretical" bias calculated by using the formulas shown in Theorem 4.3.5.

The comparison is shown graphically in Figures 4.2-4.5. In each of these figures the horizontal axis is the true mean reversion parameter k while the vertical axis is the estimation bias. The empirical biases obtained by Monte-Carlo simulation are shown in red circles while the theoretical biases are shown in blue lines. In each of the four figures a confidence interval of 2 standard deviation is indicated with green dash lines for each fixed k that are tested.

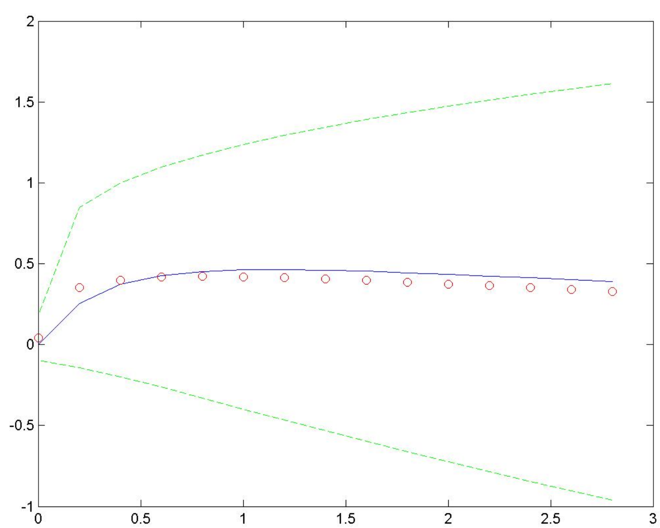

Figure 4.2: Theoretical and Empirical Bias when $T = 3, h = 1/252, H = 0.51$.

Several observations can be drawn by comparing the biases shown in these figures against those in the standard OU case, i.e. Figure 2.1-2.3.

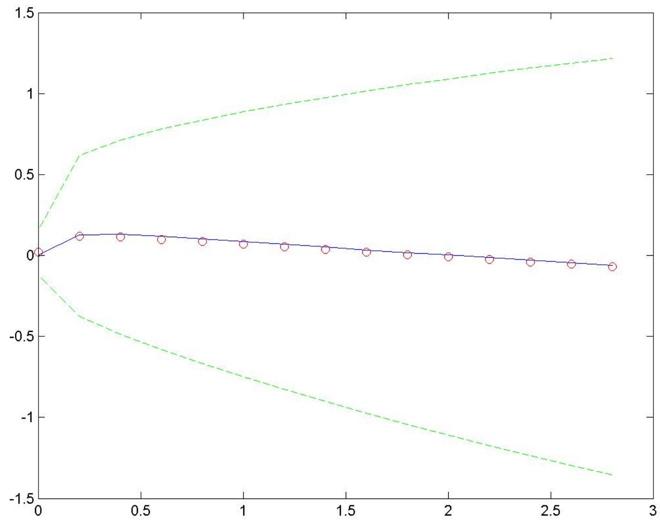

Figure 4.3: Theoretical and Empirical Bias when $T = 10, h = 1/252, H = 0.51$.

When H approaches $\frac{1}{2}$, e.g. when $H = 0.51$ as in Figure 4.2, the behavior of the bias is similar to that in the case of standard OU process, i.e. the bias approaches 0 when k approaches 0, and is positively biased for most values of k. However, when bias tends to be positively sloped in the standard OU case, the bias under the fOU process can be decreasing with increasing k, for k larger than 1.

When H is further away from $\frac{1}{2}$, the bias can decrease into negative values as k increases (Figures 4.4 and 4.5). This is contrary to the case of the standard OU process where the bias is always positive. The bias when $H = 0.6$ as shown in Figure 4.5 tends to be more negative compared to the corresponding case in Figure 4.4, when $H = 0.53$. Indeed, similar simulations also point to the fact the higher the value of H, the more negative the estimation bias can be.

The negative biasedness of the OLS estimate can be explained by taking a closer look at the stochastic differential equation governing the fOU process. In particular, as H increases, the stochastic process will become more persistent, i.e. a shock at time t will have an impact for

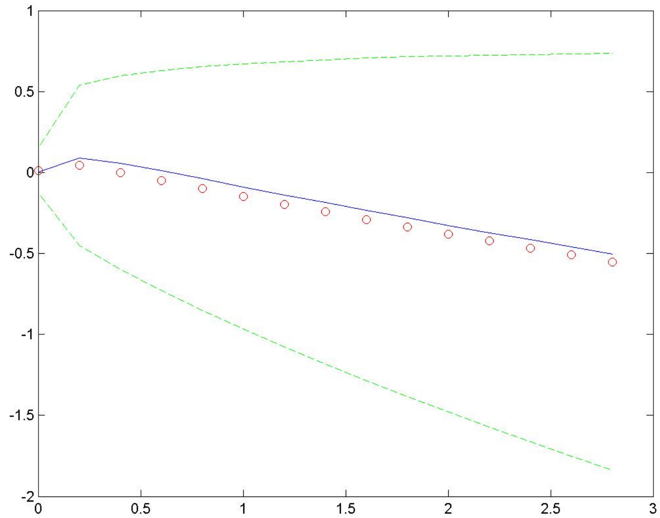

Figure 4.4: Theoretical and Empirical Bias when $T = 10, h = 1/252, H = 0.53$.

a longer range of future time. As a result, given the same shock at that initial time, a fOU with $H > \frac{1}{2}$ will tend to propagate this shock for a longer time than a usual OU process does, and heuristically speaking this implies that more time is required in order to revert to the long term mean, and hence, the mean reversion speed will appear to be smaller if we look at the fOU process through the lens as if it were still standard OU. Recall that the OLS estimate is usually positively biased in the standard OU case. The negative biasedness for some of the fOU examples we present here means that the "drag" due to the persistence of the fractional noise can sometime be so large that it outweighs the intrinsic over-estimation of the OLS estimator.

The negative biasedness of the OLS estimate under a fOU process raises some concern from the perspective of risk management. Suppose that we have different bias curves for various Hurst parameters H, such as those in Figure 4.6 showing how the estimated mean reversion changes with the actual mean reversion. Suppose also that there exists a time series of financial data which is known to follow a fOU process with $H = 0.6$ and the mean reversion is calibrated to be 1.5 using OLS on 3-year data. Then, according to the bias relation in Figure 4.6, its true mean reversion should be approximately 2. However, if initially we did not know

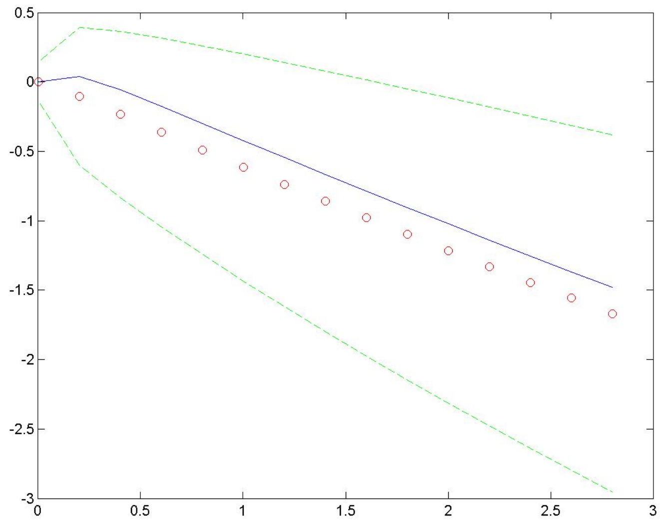

Figure 4.5: Theoretical and Empirical Bias when $T = 10, h = 1/252, H = 0.6$.

that the data are driven by fOU but instead assume the bias formula under the standard OU framework, then we would reach a conclusion that the true mean reversion should be around 1, a 50% reduction from the true value of k. In this sense, the OLS estimate without any bias adjustment appears to be a better estimate compared to the adjusted value assuming a standard OU process.

In reality, risk models tend not to capture persistence to avoid unnecessary computation effort. Instead, risk factors are assumed to follow standard Brownian processes. The above discussion reveals that under such a simplification the speed of a mean-reverting factor will be greatly under-estimated. In other words, while historical data tend to support that many time series have small mean reversion, it might be the case that these mean reversion speeds are small just because we apply the wrong model.

Moreover, since the calibration of the mean reversion parameter by OLS is sensitive to the persistence (or equivalently, the auto-correlation) of the time series in question, it is advisable to investigate the persistence property of the time series to be calibrated before applying any bias formula.

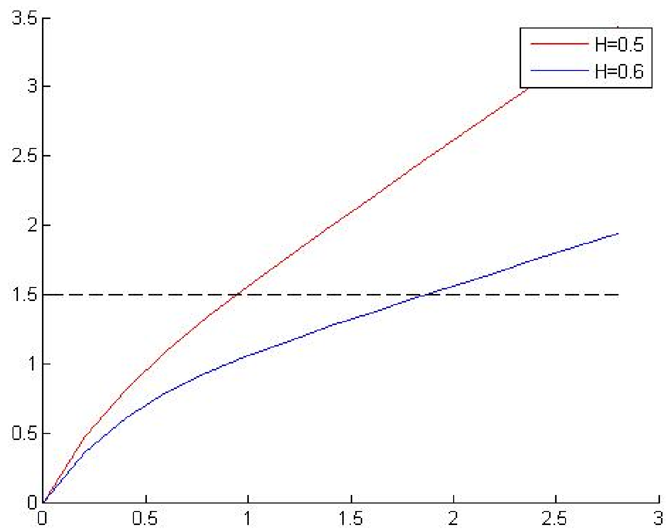

Figure 4.6: Plot of Estimated versus Actual Mean Reversion under Different Hurst Parameters.

Chapter 5

Conclusion

In this thesis, we have extended the previous work of [11] to investigate the behaviour of the bias when applying the OLS to estimate the mean reversion parameter under the fractional Brownian motion framework. The fractional Brownian motion model is chosen as an example to study the effect of persistence in the time series on the bias of the estimate of the mean reversion parameter.

It turns out that unlike the situation where the stochastic process is driven by standard Brownian noises, the OLS estimate for the mean reversion parameter can be negatively biased when the Hurst parameter H and/or the true mean reversion parameter is high. The autocorrelation present in the time series drags the underlying from reverting to its long term mean, and hence if we measure the mean reversion as if there were no persistence behaviour, the mean reversion speed would be under-estimated.

This result highlights an important model risk when one tries to calibrate mean reversion by the usual OLS method. Very often the model developer applies the OLS estimate without taking the persistence of the time series to be calibrated into consideration. The resulting estimate will almost certainly rendered to be biased. If one further naively applies the bias formula developed in [11] to this time series, the "adjusted" estimate can under-estimate the true mean reversion parameter considerably.

One may argue that one can resort to a generalized least square approach, which transformw the original question into bias estimation of the standard Ornstein-Uhlenbeck process. However, to achieve this, one still needs the information regarding the persistence of the time series in question, as we need the covariance matrix of the error terms in order to transform these error terms into approximately uncorrelated ones. One can define some estimates for the covariance matrix (a natural candidate is the empirical covariance matrix based on the available historical data), but how the estimation bias on the covariance matrix impacts the final bias of estimating mean reversion will require a further study in the future.

References

[1] Y. Bao and A. Ullah. The second-order bias and mean squared error of estimators in time-series models. *Journal of Econometrics*, 140:650–669, 2007.

[2] T.E. Duncan, Y. Hu, and B.Pasik-Duncan. Stochastic calculus for fractional Brownian motion - I. theory. *SIAM J. Control Optim.*, 38:582–612, 2000.

[3] Y. Hu and D. Nualart. Parameter estimation for fractional Ornstein-Uhlenbeck processes. *Statistics and Probability Letters*, 80(12):1030–1038, 2010.

[4] Y. Hu and B. Øksendal. Fractional white noise calculus and applications to finance. *Infin. Dimens. Anal. Quantum. Probab. Relat. Top.*, 6:1–32, 2003.

[5] Y. Hu and J. Song. Parameter estimation for fractional Ornstein-Uhlenbeck processes with discrete observations. In *Malliavin calculus and stochastic analysis*, pages 427–442. Springer, 2013.

[6] T.R. Johansen. The Bochner-Minlos theorem for nuclear spaces and an abstract white noise space. `http://www3.math.uni-paderborn.de/~johansen/seminars/minlos.pdf`. Technical report, 2003. Accessed: 14 Nov 2016.

[7] P. Rilstone, V.K. Srivastava, and A. Ullah. The second-order bias and mean squared error of nonlinear estimators. *Journal of Econometrics*, 75:369–395, 1996.

[8] L.C.G. Rogers. Arbitrage with fractional Brownian motion. *Mathematical Finance*, 7:95–105, 1997.

[9] D.W. Stroock and S.R. Srinivasa Varadhan. *Multidimensional diffusion processes, volume 233 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, 1979.

[10] A. Ullah. *Finite sample econometrics.* Oxford University Press, Oxford, UK, 2004.

[11] J. Yu. Bias in the estimation of the mean reversion parameter in continuous time models. *Journal of Econometrics*, 169:114–122, 2012.