Department of Mathematics and Statistics

Modeling and Pricing Financial Assets Under Long Memory Processes

Masnita Misiran

This thesis is presented for the Degree of

Doctor of Philosophy

of

Curtin University of Technology

May 2010

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university. To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgement has been made.

Acknowledgement

I would like to express my gratitude to my supervisors, Dr. Zudi Lu and Prof. K.L. Teo for their priceless shared knowledge, guidance, patience and supports during my study. They patiently reading many drafts of this thesis and papers and listened to constant problems that arise during the completion of this thesis. I really appreciate their kindness and precious help. Thank you. Your wisdom teaches me to live humble and sincere.

I would also like to thank Dr. Changzi Wu for his valuable discussions on optimal control, and also his enthusiasm in doing research. I am very fortunate to having opportunity studying from him, even for a short period of time. I am also grateful to Department of Mathematics and Statistics. The constant seminars throughout the years, together with friendly lecturers and supporting staff helped motivate me, and renewed my curiosity in many diverse disciplines of knowledge.

I would also like to extend my gratitude to Malaysian Ministry of Higher Education and Universiti Utara Malaysia for sponsoring my study and stay in Australia. Their financial aids are important for the completion of this thesis.

Thank you to all my collective friends I met here in Australia, for your friendship. To Rui Li and Jimmy, for being so kind and helped me get through my chaotic and self-cautious beginning of my new life here. To Rahman, Kak Ita, Kak Jem and family, it is always a good feeling to know I can count on you. To Mick, Renato, Cris and Waseem, for such a good times. We will see each other again in the future. To Siti, Nize, Ibrahim and Marlizah, thanks for joining me hitting a tennis ball most Saturday evenings. It helps hit all the stress away.

Last but by no means least, I would like to thank my mother and my father for their unconditional love. They are the reasons of my being who I am today. My sister, my brother in law and their three little princesses for allowing me being part of their beautiful family. To Harda, thank you. You always make my heart smiles.

Abstract

An important research area in financial mathematics is the study of long memory phenomenon in financial data. Long memory had been known long before suitable stochastic models were developed. Fractional Brownian motion (FBM) can be used to characterize this phenomenon. This thesis examines the use of FBM and its long memory parameter H, from the view point of estimation method, approximation, and numerical performance.

How to estimate the long memory parameter H is important in financial pricing. This thesis starts by reviewing the performance of some existing preliminary methods for estimating H. It is then applied to some Malaysia financial data. Although these methods are easy to use, their performance are in doubts, in particular these methods can only get an estimator of H, without providing the dynamic, long-memory behaviour of financial price process.

This thesis is therefore concerned with the estimation of the dynamic, long-memory behaviour of financial processes. We propose estimation methods based on models of two stochastic differential equations (SDEs) perturbed by FBM, that play important role in option pricing and interest rate modelling. These models are the geometric fractional Brownian motion (GFBM) and the fractional Ornstein-Uhlenbeck (FOU) model, respectively. These methods are able to obtain H and other parameters involved in the models. The efficiency of these methods are investigated through simulation study. We applied the new methods to some financial problems.

We also extend this study to filtering the SDE driven by FBM in multi-dimensional case. We propose a novel approximation scheme to this problem. The convergence property is also established. The performance of this method is evaluated through solving some numerical examples. Results demonstrate that methods developed in this thesis are applicable and have advantages when compared with other existing approaches.

Publications arising from the thesis

Misiran, M., Wu, C., Lu, Z. and Teo, K.L. (2010). Optimal filtering of linear system driven by fractional Brownian motion. To appear in Dynamic Systems and Applications.

Misiran, M., Lu, Z. and Teo, K.L. (2009). Estimation of geometric fractional Brownian motion with application to option pricing. Submitted to Quantitative Finance.

Misiran, M., Lu, Z. and Teo, K.L. (2010). Fractional Black-Scholes model: Complete MLE with application to fractional option pricing. To appear in Proceeding of the International Conference on Optimization and Control, Guiyang, China, July 2010.

Misiran, M., Lu, Z. and Teo, K.L. (2008). Investigating Malaysia financial markets by Hurst exponent. Proceeding of the National Management Conference, Terengganu, Malaysia.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Market is terrifying, yet it is equally captivating. Whether you are a lone trader, a member of a well known corporate, or a researcher, the market brings with itself the awe in silence. We give the honor to the first responsible scholar, Louis Bachelier [10], a French mathematician, who laid the foundation for financial theory from the view point of probabilistic analysis in his thesis, entitled *Theorie de la Speculation*. Unfortunately, Bachelier's work was ignored for years after his time. It was until 1956 that the importance of his ideas was finally recognised by Paul E. Samuelson, an economic student from MIT. He studied option pricing in his thesis, where he encountered the phenomenon that was called "fair game" by Bachelier. This phenomenon is now recognised as a random process.

Many outstanding works have appeared since then. These included the works by Sharpe, Black, Scholes and Merton ([126], [18], [106]). In particular, Black and Scholes provided the financial world with the famous Black-Scholes Market Model in 1973. Merton extended this study in continuous time. They were awarded Nobel prizes for their contributions. These are celebrated achievements as we can now understand the market behaviour better, and we are also provided with tools to better price the market.

However, the event that took place on the Black Monday in October 1987, the largest one-day decline in stock market history, had given a blow to all the traders. Some regarded this phenomenon as the Black Swan event. This very phenomenon was believed to be an extremely rare event, almost impossible to occur. For detailed discussions, see the book [132] by Nassim Nicholas Taleb. This event seems quite similar to earlier events, such as the Black Monday of Wall Street in 1929 and the crash of the market in 1914. In fact, this so-called rare event doesn't seem so rare anymore as it keeps recurring in recent time, such as the Asian financial crisis in 1997, the event on September 17, 2001, and most recently, the market downturn in October 2008. This latest event has caused economic slowdown and recession worldwide. All these events have confirmed that the market behaviour is far from being able to be predicted by previous models.

Motivated by these disturbing events, researchers begin to have a closer look in the underlying market models. Some started to question the assumption of the Brownian motion, or the Wiener process, that is being used extensively in many financial models. This assumption seems too naive to model the chaotic real-world market behaviour. This leads to the suggestion of the long memory properties, such as jump-diffusions, subordinated processes, Levy processes, and processes with stochastic volatility driven by fractional Brownian motion [44]. These properties can capture the memory dependency in past data. The idea appears to be more reasonable in comparison to the Brownian motion process. For financial processes driven by fractional Brownian motion, the works ([97], [98], [103], [99], [93], [90]) by Mandelbrot should be mentioned for recognizing the long memory phenomenon in financial data. His early interest was at the index of self similarity, which is found in power law problems. This important index is known as Hurst index, first introduced by H.E. Hurst in [63]. On this basis, Mandelbrot introduced the fractional Brownian motion and its underlying statistics. He believed that the main contribution to the fragility of a market model is caused by the use of Brownian motion. His enthusiasm was shared by his students and colleagues which is clearly evident from many of their works in the area.

Since the work by Mandelbrot, many econometricians started to investigate the properties of these processes. Earlier works were more focused on finding the best estimator for Hurst parameter. Subsequently, several methods have been developed, which include the R/S analysis method, the modified R/S method by Lo, the periodogram based method and the Whittle estimator. However, all these methods come with their own strengths and drawbacks. The perfect method has yet to be proposed to cater for this problem. As the financial world advances, so does the mathematical theory. Researchers started to incorporate the long memory properties in the economic models. We consider, in this thesis, some of the important models in the area.

We will investigate the extensions of the Black-Scholes model and also the Ornstein-Uhlenbeck model. Black-Scholes model has been used extensively as a model to value the option, while Ornstein-Uhlenbeck model is well known for being a tool to model interest rate. However, both these two models used Brownian motion as the governing noise. In this thesis, these problems are considered with their long memory properties being taken into account. That is, Brownian process is replaced by fractional Brownian motion (FBM) for modelling noise process. How to estimate the unknown parameters in these models are extremely important in applications. Our aim is to find the parameters involved in these models.

For filtering, its aim is to extract the best information on the state process based on the measured data. However, besides [4], there do not appear to have other papers available in the literature, whose main concern is on the filtering problems with noises being characterized by FBMs. This is especially so in the multi-dimensional cases, where both the state and observation are governed by respective stochastic differential equations (SDE). We understand that SDE is very vital in finance, as most of the financial problems are described by appropriate SDEs. By incorporating the FBM process in SDE, we are able to capture the memory dependency in the model, a notable property which is absent in the standard Brownian motion. Our aim is to investigate properties which are inherited in this model. Furthermore, we will provide computational schemes to calculate the crucial parameters involved in this model.

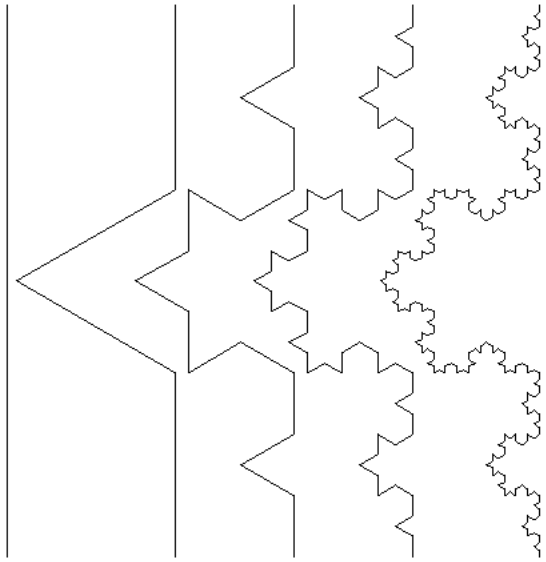

Figure 1.1: First few iterations for a fractal geometry [20]

1.2 Fractional Brownian Motion and Long Memory in Finance Modeling

We begin this section by introducing a fractal theory, taken from Latin word 'fractus', meaning broken or fractured. It was introduced by Mandelbrot as early as 1982. Fractal can be understood as a rough or fragmented geometric shape that can be split into parts, each of which is (at least approximately) a reduced-size copy of the whole [95]. This idea is inspired from nature. For example, fern, cloud and cauliflower have all exhibited such fractal behaviours.

This idea is then adopted in financial world. The structure of price changes can be seen as similar if we look closely at the successive monthly fluctuations or daily price changes. This approach allows us to compare these prices in different scales and look at how they behave. We can then analyse the global market through the understanding of the local market as they are statistically identical. They differ only by a scale factor or a self-similar parameter. The Brownian motion is an example of this fractal theory, as the Brownian motion has a property that $B(\alpha t)$ has the same law as $\alpha^H B(t)$, where $H = 1/2$. For fractional Brownian motion, H can take a value from $0 < H < 1$. However, it still enjoys the self-similar property.

The self-similar property is known to have many real world applications in diverse disciplines. This is because some persistent order rather than a mere chance of random occurrence is often observed in their behaviours, for example in financial data. Thus, the financial market should not be viewed as being an efficient market. In fact, the phenomena with long memory have been noticed well before the development of stochastic models. Scientists have observed empirically that correlations between observations that are far apart decay to zero at slower rate as compare to independent data or data generated by classical ARMA or Markov-type models [14]. In the area related to long memory, the pioneer work was done by Hurst. He has done an extensive study on the history of Nile River's discharge, and afterward, many other great rivers world wide. He has found similarity in all his findings. He believed that the range is widened not by a square-root law as in coin tossing, but as a three-fourths-power (at about 0.73). He named the method as Rescaled Range analysis ([63], [64]). The published work on this finding has aroused the interest of many researchers and practitioners. Some agree with this new method of measuring the range, while some despised the idea as miscalculation and statistically improper.

Intrigued by these intensed discussions, Mandelbrot has decided to engage in an in-depth study on the idea and compared it with his power law problems in the context of cotton prices. He discovered that Hurst finding is, in fact, describing long dependency behaviour, and it is a pillar of fractal geometry. Excited with this finding, Mandelbrot ([91], [99], [100], [102], [101], [103], [98], [92], [96]) has refurbished the analysis and applied it to problems in financial world. He has also introduced the fractional Brownian motion, an extension of the Brownian motion, where a self similarity index, called Hurst parameter, named after Hurst, is involved. Since then, the study of long memory, particularly for finding the Hurst parameter, has attracted an intense interest in the financial world. Mandelbrot [93], who introduced the R/S method to this field, investigated the financial returns. These problems have been further investigated by Green and Fielitz [53], Lo [85] and many others (see, for example, [29], [51], [108], [119]). There are also estimators, mostly heuristic, for the estimation of the Hurst parameter. Each of which is having its advantages and disadvantages.

We present here some important definitions that will be used in later chapters.

Definition 1.2.1. *A stochastic process $B(t)$ is a Brownian motion if it satisfies the following properties:*
1. $B(t)$ is a continuous function of time with $B(0) = 0$.
2. $B(t)$ has independent increments, i.e., for all $t > s$, $v > u$ and $u > t$, $B(t) - B(s)$ and $B(v) - B(u)$ are independent.
3. $B(t)$ has normal increments, i.e., for all $t > s$, $B(t) - B(s) \sim N(0, t - s)$.

Definition 1.2.2. *Let H be a constant belonging to $(0, 1)$. A fractional Brownian motion (FBM), $(B_H(t))_{t\geq 0}$, with the Hurst index H is a continuous and centered Gaussian process with covariance function*

$$E[B_H(t)B_H(s)] = \frac{1}{2}(t^{2H} + s^{2H} - |t - s|^{2H}).$$

For $H = \frac{1}{2}$, the FBM is a standard Brownian motion.

By Definition 1.2.2, the following properties hold:

(i) $B_H(0) = 0$ and $E[B_H(t)] = 0$ for all $t \geq 0$.

(ii) B_H has homogeneous increments, i.e., for $s, t \geq 0$, $B_H(t + s) - B_H(s)$ has the same law as that of $B_H(t)$.

(iii) B_H is a Gaussian process and $E[B_H(t)]^2 = t^{2H}, t \geq 0$, for all $H \in (0, 1)$.

(iv) B_H has continuous trajectories.

The FBM can be divided into three different families according to $0 < H < \frac{1}{2}$, $H = \frac{1}{2}$ and $\frac{1}{2} < H < 1$ [15]. For FBM with $0 < H < \frac{1}{2}$, it is known as having short memory dependency, and for FBM with $\frac{1}{2} < H < 1$, it corresponds to having long memory dependency. For FBM with $H = \frac{1}{2}$, it is a standard Brownian motion.

We can also define FBM by using an integral representation on a compact interval.

Definition 1.2.3. *Let $\{B_t, t \in [0,1]\}$ be a standard Brownian motion. Then*

$$B_{H,t} = \int_0^t K_H(t,s)dB_s, \quad t \in [0,1]$$

is the FBM on $[0,1]$ and $K_H(t,s) = \frac{(t-s)^{H-\frac{1}{2}}}{\Gamma(H+\frac{1}{2})}F(\frac{1}{2}-H, H-\frac{1}{2}, H+\frac{1}{2}, 1-\frac{t}{s})\mathbf{1}_{(0,t)}(s)$, where Γ is a gamma function and F is the Gauss hypergeometric function [62].

Definition 1.2.4. *An $\mathbb{R}^d$-valued random process $\{X(t), t \geq 0\}$ is self similar or satisfies the property of self-similarity if for every $a > 0$ there exists a constant $b > 0$ such that*

$$\{X(at)\} =_d \{bX(t)\}, \tag{1.1}$$

where $=_d$ denotes equality of all joint distributions for $\mathbb{R}^d$-valued stochastic processes defined on some probability space $(\Omega, \mathcal{F}, P)$.

$\{X(t), t \geq 0\}$ is stochastically continuous at t if for any $\varepsilon > 0$, $\lim_{h\to 0} P\{|X(t+h) - X(t)| > \varepsilon\} = 0$.

Definition 1.2.5. *If $b = a^{-H}$, then we say that $X = (X_t)_{t\geq 0}$ is a self-similar process with the Hurst index H (i.e., it satisfies the property of (statistical) self-similarity with the Hurst index H). The quantity $D = \frac{1}{H}$ is called the statistical fractal dimension of X.*

Remark 1.2.1. *Let $X(t)$ be a geometric fractional Brownian motion (GFBM) defined by*

$$dX(t) = \mu X(t)dt + \sigma X(t)dB_H(t), \qquad t \geq 0, \qquad X(0) = x > 0,$$

with x being a constant, where $B_H(t)$ is a one-dimensional FBM with $H \in (0,1)$. By using the Wick calculus developed by Hu and Oksendal [60], the solution of this equation is found to be expressible as:

$$X(t) = x\exp(\sigma B_H(t) + \mu t - \frac{1}{2}\sigma^2 t^{2H})$$

Remark 1.2.2. *The fractional Ornstein-Uhlenbeck process can be written in the form of*

$$\begin{aligned} dx(t) &= -ax(t)dt + \sigma dB_H(t) \\ x(0) &= x_0, \end{aligned} \tag{1.2}$$

where $a > 0$ and σ_1 are constant parameters, known as the drift and volatility parameters, respectively. $(B_H(t), t \geq 0)$ is called a fractional Brownian motion with the Hurst index H.

Here, we summarize the relationship of some important stochastic processes in Figure 1.2.

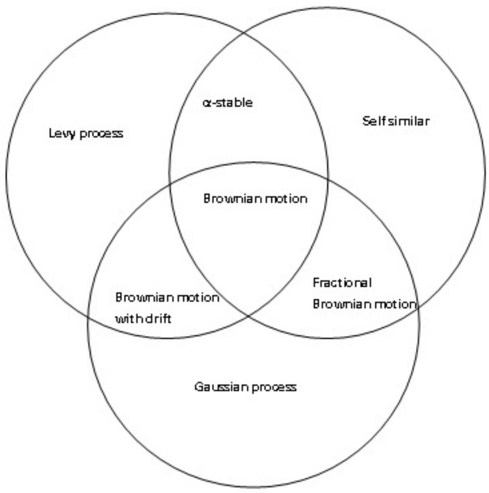

Figure 1.2: Relationship between different groups of stochastic processes [123]

1.3 Literature Review

For a financial market, it is important to understand the random fluctuation process representing the error term. In normal circumstances, it is governed by Brownian motion, i.e. Gaussian process. A good understanding on this term will enable us to construct a reliable model for a financial market. On this basis, good methods for the estimation of parameters can be developed, providing a good prediction of the market behavior. The idea of Brownian motion was first introduced by a Botanist, Robert Brown, who studied pollen grains in suspension in the early nineteenth century. While studying the movements of these pollens that were moving erratically. He first thought that this behavior was due to the fact that the pollen grains were alive. He later discovered that this phenomenon was the result of collisions among grains.

Interestingly, this concept was later introduced independently by Bachelier [10] in his dissertation. Bachelier, who was a mathematics student with a sound background on mathematical physics, studied the financial market behavior in France, where he tried to model the price movements of stocks and commodities.

Note that Brownian motion does not have memory for the past. Thus, the present occurring could not have been forecasted using the historical data. This model has two major drawbacks when used to model stock or commodity prices [122]. Firstly, since the price of a stock is a normal random variable, it can theoretically become negative. Secondly, for Brownian motion, it is assumed that the difference in price over an interval of fixed length has the same normal distribution irrespective of the initial price for those time intervals. This assumption does not appear to reflect the real situation in financial market.

Motivated by these drawbacks, fractional Brownian motion (FBM) was introduced. The most important feature in FBM is that it has memory. This feature makes more sense in understanding the behaviour of financial time series.

Fractional Brownian motion (FBM) was first studied by Kolmogorov [76] in 1940. FBM has been used extensively in many fields, such as medical, environment and internet. We refer readers to [117], [87], [145], [141] and the relevant papers cited therein for further details. However, it was Mandelbrot and Van Ness [99] who brought FBM to finance world and subsequently motivated other researchers in this field. Beran, Hu, Oksendal, and Mishura ([14], [60], [15], [107]) are amongst those who have made important contributions in the field.

The main parameter in FBM is the Hurst index, H. Intensive research has been conducted, aiming to find the best estimator for this index. The Rescaled Range (R/S) analysis is a widely used estimator in the empirical study to measure this index. This is due to its simplicity. The modified R/S by Lo is introduced later in [85]. However, it was subject to criticism in [137]. Some other estimators are Higuchi method [55], detrended fluctuation analysis [142], Whittle estimators [134], to name just a few. Over the years, many empirical works have been conducted by using these methods.

A more ambitious investigation started to emerge, focusing on the implementation of long memory in financial models. New methods for parameter estimation have been developed for FBM, and later, for models with long memory. Examples include those reported in [48], [37] and [119]. Beran [14] summarizes estimation related problem with long memory in discrete time. Later, Comte [33] and Comte and Renault [34] initiated the study of estimation related problems with long memory in continuous time. These works have aroused the interest of many researchers, leading to the publication of many papers. See, for example, [66], [27], [8] and [81].

There was some criticism reported in [121] against FBM as a financial model because this model admits an arbitrage opportunity. However, recent developments ([60], [46]) have helped overcome this problem. Consequently, the stochastic calculus for FBM was being developed ([15], [107]) extensively to accommodate the needs for a wide spread of applications. Some works on the application of FBM in financial market were reported in ([99], [127], [123]). There are also recent developments on some finance models with long memory. Here, we would like to mention some of the works with focus on the geometric fractional Brownian motion model in [78], [23], [60] and [15]. Active research on the fractional Ornstein-Uhlenbeck model can also be found in ([75], [62], [56], [30] and [59]). There are also some important works on optimal filtering of a model driven by FBM. See, for example, [4],

[36], [74] and [80].

1.4 Thesis Layout

Chapter 2 is devoted to preliminary analysis of existing methods for estimating the index of self similarity, H. Most of these methods are heuristic in nature. Background on some of these existing methods will be revealed. Then, we will move on to consider two most favorable methods, the R/S analysis and Whittle method, and their application to some Malaysia financial data sets. This preliminary investigation will help provide us with insight to the behavior of Malaysia financial market. Our aim is to find whether or not Malaysia market exhibits the existence of a long memory behavior in its financial data series.

Malaysia financial data is of interest, because Malaysia is a fast growing developing country. This is especially so during and after the Asian financial crisis in mid 1997. In this chapter, we will consider three major components in Malaysian economy, the Kuala Lumpur Composite Index (KLCI), the currency, and the interest rate. We will investigate if there is a sign of memory in these time series. The reason for choosing the KLCI is that it is the aggregation of stock prices, the main indicator for the performance of Malaysian economy. The currency will be with reference to the USD.

Chapter 3 will focus on method for estimating the parameters which are used for pricing the fractional option pricing for the fractional Black-Scholes model. This is an extension of the famous Black-Scholes model. This model can capture the memory that exists in some financial data, which is clearly not following a Brownian motion. We will propose a complete maximum likelihood estimation method for estimating the parameters of the geometric fractional Brownian motion (GFBM) model. GFBM is an extended model of the traditional geometric Brownian motion that is widely used for Black-Scholes option pricing. By considering GFBM, we are now able to capture the memory dependency. This method will enable us to derive the estimators of the drift, μ, volatility, σ^2, and also the index of self similarity, H, simultaneously. This will enable us to use the fractional Black-Scholes model with all the needed parameters. We will also carry out simulation study to illustrate the effectiveness and reliability of the method proposed. In this chapter, we will also carry out an empirical application to stock exchange index with option pricing under the GFBM.

Chapter 4 is concerned with estimating parameters in the fractional Ornstein-Uhlenbeck (FOU) model. This model is an extension of the famous Ornstein-Uhlenbeck model that is widely applied to model interest rates, currency exchange rates and commodity prices stochastically. By considering FOU model, we are now able to incorporate the memory dependency in this process. It will give more reasonable explanation to the behavior of financial time series. By using the dynamic measurement error, we will make use of the innovation process to derive the likelihood function for this process. A simulated annealing method is used to get the parameters involved in the system simultaneously. Further, we will present a simulation investigation to look at the efficiency of our method. An empirical work to the interest rate data will also be presented.

Chapter 5 is concerned with a continuous time filtering of a multi-dimensional stochastic differential system driven by a fractional Brownian motion process. We will show that this filtering problem is equivalent to an optimal control problem involving convolutional integrals in its dynamical system. In this chapter, we will develop a novel approximation scheme and applied it to this optimal control problem. This yields a sequence of standard optimal control problems. We will then establish the convergence of the approximate standard optimal control problem to the optimal control problem involving convolutional integrals in its system dynamics. Further, some numerical examples are solved by using the method proposed.

We will conclude the thesis with discussions in Chapter 6.

Chapter 2

Preliminary estimation of long memory index: Evidence from Malaysia financial markets

2.1 Introduction

Hurst index, H, has been used in a vast literature to detect the existence of a long range dependency in a time series data. Since its discovery in 1951 [63], this parameter has found applications in many fields of study. These include environment ([130], [88]), network and telecommunication traffics ([2], [138]) and economics and finance ([23], [127]). For $H \in (\frac{1}{2}, 1)$, it indicates the existence of a long range dependency in the time series data under consideration. Methods based on heuristic approaches are often used for finding the index H. Further details on these methods are discussed in the next section. These methods are easy to implement. However, they are also subject to criticism mostly for lack of statistical properties.

This chapter is dedicated to explore some of these methods. We focus on two most popular methods, i.e., the R/S analysis method and Whittle estimator, for testing the existence of long range dependency within some simulated data. Further in this chapter, we will look at the Hurst index in Malaysia financial data.

2.2 Overview of Hurst Index Estimators

In most existing literature, heuristic approaches are used to get the value of the Hurst index, H. The main reason is that these approaches are easy to follow, especially for practitioners who are not statisticians. As we have discussed in the previous chapter, this index is very important in many applications such as hydrology and internet traffics.

One of the estimators worth mentioning is the celebrated Rescaled Range analysis (R/S analysis). It is very popular among researchers until today ([90], [53], [51], [143], [105], [24], [35]). However, this method is examined critically by Lo [85] and, subsequently, Lo's modified R/S analysis is being introduced to compensate the drawback of the R/S analysis. However, this work is also being criticized few years later in [137].

There are also many other methods available in the literation for estimating the Hurst exponent, such as Higuchi method [55], the modified R/S analysis [85], the multiaffine analysis [116], the detrended analysis [9], periodogram regression [49], the moving average analysis technique [43], the ARFIMA estimation by exact maximum likelihood [129], the Whittle estimator [50], the variance method [26], bivariate multi fractal model [83] and the wavelet based estimators [50]. However, the R/S analysis is still the most widely used one in various applications. As it is mentioned earlier, this is due to the fact that it is developed based on simple statistics without the need of any assumption on the underlying process of the time series. Furthermore, it works well in many empirical cases, especially those with large sample size [111].

As far as the estimation is concerned, there is yet a perfect method that is agreed by all researchers. Each method has its own drawbacks and cannot be used as a sole estimator in all cases. For more detailed discussions, see the work by Robinson and Beran ([120], [14]).

The R/S analysis will be used in the study of Malaysia financial time series data. For comparison, we will also discuss briefly the efficiency of Whittle estimator in its application to a simulated data. Some other available estimators will also be briefly reported.

2.2.1 Rescaled Range (R/S) Analysis

The idea behind the R/S analysis is to find whether or not the amount of data varying from maximum to minimum over varying periods of time is greater or smaller than what is expected if each data point is independent of the last one. It is expressed as the equation below.

$$R/S = \frac{\max_{1\leq k\leq n}\sum_{j=1}^{k}(r_j - \overline{r}_n) - \min_{1\leq k\leq n}\sum_{j=1}^{k}(r_j - \overline{r}_n)}{\frac{1}{n}[\sum_j (r_j - \overline{r}_n)^2]^{\frac{1}{2}}}. \tag{2.1}$$

Let r denote the return series over different time periods, say one day, two days, and so on up to the full period n of the time series. The average return of n, $\bar{r}_n$, is calculated. We then calculate the differences for each period of time, namely one day, two days, until n days. We further identify the maximum and the minimum for these obtained values. By subtracting these two values, we get the estimation of the range from peak to trough in the accumulated deviation. This estimation is then denominated by the conventional standard deviation in the data series. In the work ([63], [91]) by Hurst and Mandelbrot, it is found that there exists a scaling behaviour with exponent $H > \frac{1}{2}$ in many geophysical time series. This indicates a tendency of deviation from the mean. This behavior also appears in time series models driven by fractional Gaussian noises, and fractional integrated ARMA models. This type of process behaves like a slow decaying auto correlation function with long term dependency. It follows the asymptotic scaling relationship as

$$(R/S)_t \sim at^H. \tag{2.2}$$

The Hurst index, H, can be estimated by a simple linear regression over a sample of increasing time horizons ($s = t_1, t_2, ..., T$):

$$\ln(R/S)_s = \ln(a) + H\ln(s). \tag{2.3}$$

Any value other than $H = \frac{1}{2}$ indicates the presence of a memory dependency. If $H > \frac{1}{2}$, then the data are "persistent", and there are long runs in the data. If the data is up/down in the last period, then most likely the data will continue going up/down in the next period. On the other hand,

$H < \frac{1}{2}$ implies "anti persistent" in the data under consideration. Meaning, if the data is up/down in the last period, then most likely the data will go down/up in the next period. We refer the reader to ([13], [24], [31], [35], [53], [51], [58], [65], [105], [108], [110], [113], [114], [143]), where the R/S analysis method is used.

2.2.2 Whittle Estimators

Whittle method has been widely used in the literature to estimate the parameter of a long range dependency ([7], [50], [77], [109], [134]). This method is based on the parametric method under the assumption that the spectral density of the time series is known with an exception of few parameters, which are to be estimated. Let $X_j, j = 1, ..., N$, be a mean zero time series with spectral density $f(\upsilon;\eta), -\pi < \upsilon < \pi$. The function f is assumed known, except for the parameter vector η,

$$f(\upsilon) = C_H(2\sin\frac{\upsilon}{2})^2 \sum_{k=-\infty}^{\infty} \frac{1}{|\upsilon + 2\pi k|^{2H+1}} \sim C_H|\upsilon|^{1-2H}, \tag{2.4}$$

where C_H is a constant. Whittle estimator is based on the periodogram of the time series,

$$I(\upsilon) = \frac{1}{2\pi N}|\sum_{j=1}^{N} X_j e^{ij\upsilon}|^2. \tag{2.5}$$

This estimator is the vector η which minimizes the following function

$$Q(\eta) := \int_{-\pi}^{\pi} \frac{I(\upsilon)}{f(\upsilon,\eta)} d\upsilon + \int_{-\pi}^{\pi} \log f(\upsilon,\eta). \tag{2.6}$$

In a real life application, the sum over the Fourier frequencies $\upsilon_j = \frac{2\pi j}{N}$, where $j = 1, 2, ..., [(N-1)/2]$ and N is the length of the series, is computed instead of taking the integration. The actual function to be minimized is

$$Q^*(\eta) = \sum_{j=1}^{[(N-1)/2]} \frac{I(\upsilon_j)}{f^*(\upsilon_j;\eta)}, \tag{2.7}$$

where $f^* = \beta f$, and $\int_{-\pi}^{\pi} \log f^*(\upsilon;\eta) d\upsilon = 0$. However, the evaluations of Whittle estimators do not involve the parameter β.

When dealing with a fractional Gaussian noise, η represents the parameter H. $\hat{H}$ converges to its true value H at the rate of $N^{\frac{1}{2}}$ and the asymptotic distribution of $\sqrt{N}(\hat{H} - H)$ is Gaussian [134].

2.2.3 Aggregated Variance Method

From [14], the aggregated variance method can be obtained by first dividing the time series $X = \{X_i, i \geq 1\}$ into several blocks, where m_k is the size of the k-th block. Then, we take the average within each block,

$$\bar{X}(k) = \frac{1}{m_k} \sum_{j=1}^{m_k} \bar{X}_j(k), \tag{2.8}$$

for successive values of m_k, where k denotes the k-th block. The sample variance of the sample, for each k, is

$$s^2(k) = \frac{1}{m_k - 1} \sum_{k=1}^{m_k} (\bar{X}_j(k) - \bar{X}(k))^2. \tag{2.9}$$

The graph of $\log s^2(k)$ against $\log k$ is then plotted. This sample variance should be asymptotically proportional to m^{2H-2} for large N/m and m [32].

2.2.4 Higuchi Method

This method is introduced by Higuchi [55]. By considering the time series $X(i), i = 1, ..., N$, the main idea behind this method is to find the normalized length of the curve,

$$L_m(k) = \frac{1}{k}\{(\sum_{i=1}^{[\frac{N-m}{k}]} |X(m+ik) - X(m+(i-1)k)|)\frac{N-1}{k[\frac{N-m}{k}]}\} \tag{2.10}$$

for $m = 1, ..., k$. Note that this method divides time series to m blocks. $N - 1/[(N-m)/k] \cdot k$ represents the normalization factor for the curve length for each block. It follows that the average of $L_m(k)$ is

$$L(k) = \frac{1}{k} \sum_{m=1}^{k} L_m(k). \tag{2.11}$$

For a statistically self-similar curve, we have

$$L(k) \propto k^{-D}, \tag{2.12}$$

where $D = 2 - H$. Thus, a log-log plot of $L(k)$ versus k should produce a straight line with the slope D.

2.2.5 Periodogram Method

This method is first described in [49] by using the frequency domain of a time series data. For simplicity, we use the definition given in [133]. Let X_j be a given time series data. Then, calculate

$$I(\lambda) = \frac{1}{2\pi N} |\sum_{j=1}^{N} X_j e^{ij\lambda}|^2, \tag{2.13}$$

where λ is a frequency and N is the length of the time series. Since $I(\lambda)$ is an estimator of the spectral density, a time series with a long-range dependency should have a periodogram, which is proportional to $|\lambda|^{1-2H}$, close to the origin. Then, we can plot the logarithm of the spectral density of the time series versus the logarithm of the frequencies. Its slope will provide us with the value of H.

2.3 Simulation Results

Here, we use models with a long memory to test the performances of some of the existing methods. We give a brief description for each of these models.

2.3.1 Fractional Integrated Model and Fractional Integrated Generalized Auto Regressive Conditional Heteroskedasticity (FIGARCH) model

The easiest way to generate a long memory series is by using the difference operator $(1 - L)^d$ for a fractional values of d, so a basic long memory series gets generated as

$$(1 - L)^d X(t) = \varepsilon_t, \tag{2.14}$$

where L is the lag operator and $d > 0$ in $\mathbb{R}$ [128]. This process can be regarded as an infinite-order moving average process. This is because

$$X(t) = (1-L)^{-d}\varepsilon(t) \tag{2.15}$$

and equivalently,

$$X(t) = \sum_{\tau=0}^{\infty} b(\tau)L^{\tau}\varepsilon(t), \qquad \varepsilon \sim \text{iid}(0, \sigma_{\varepsilon}^{2}), \tag{2.16}$$

which is a weighted summation of the white noise $\varepsilon(t)$, where the moving average coefficient $b(\tau)$ can be expressed in terms of the gamma function as follows:

$$b(\tau) = (-1)^{\tau}\begin{pmatrix} -d \\ \tau \end{pmatrix} = \frac{\Gamma(\tau+d)}{\Gamma(d)\Gamma(\tau+1)}. \tag{2.17}$$

It has been shown [73] that a long term dependent irregular process is a process with auto covariance function $\gamma(\tau)$ given by

$$\gamma(t) = E\{x(t)x(t-\tau)\} = \int_{-\infty}^{\infty} x(t)x(t-\tau) = \left\{ \begin{array}{cc} \tau^{\lambda}L(\tau) & \text{for}\lambda \in [-1,0] \\ -\tau^{\lambda}L(\tau) & \text{for}\lambda \in [-2,-1] \end{array} \right\} \tag{2.18}$$

and as $\tau \to \infty$, $L(\tau)$ is a slowly varying function at infinity, for example, a constant.

The auto covariance function (ACF) of a fractional differenced time series is given by

$$\begin{aligned} \gamma(t) &= \frac{\sigma_{\varepsilon}^{2}(-1)^{\tau}(-2d)!}{(\tau-d)!(-\tau-d)!} \\ &= \frac{\sigma_{\varepsilon}^{2}\Gamma(1-2d)\Gamma(\tau+d)}{\Gamma(d)\Gamma(1-d)\Gamma(\tau+1-d)} \sim \sigma_{\varepsilon}^{2}\tau^{2d-1} \end{aligned} \tag{2.19}$$

where $\varepsilon(t) \sim \text{iid}(0, \sigma_{\varepsilon}^{2})$ and $d \in (-0.5, 0.5)$. If we have the fractional differenced process, then the dependency exponent $\lambda = 2d - 1$ and the slowly varying function is a constant white noise with variance σ_{ε}^{2} . This auto covariance is slowly decaying.

As $d \downarrow 0.5$, the white noise, $\varepsilon(t)$, is fractionally differentiable, and the

ACF decays hyperbolically, i.e., $\gamma(\tau) \to \sigma_\varepsilon^2 \tau^{-2}$. These properties indicate that the increment of the series is anti persistent. When $d = 0$, $\varepsilon(t)$ is a white noise and its ACF decays hyperbolically, i.e., $\gamma(\tau) \to \sigma_\varepsilon^2 \tau^{-1}$. As $d \uparrow 0.5$, the white noise is fractionally integrable, and the ACF will decay slower than hyperbolically, i.e., $\gamma(\tau) \to \sigma_\varepsilon^2$. These indicate that the increment of the series is persistent.

These models include the ARCH model [47] and the GARCH model [19]. They have been widely applied. There are also several extensions after they are introduced in [29]. These models are particularly suitable for problems that arise in risk management, portfolio analysis and derivative pricing.

The fractional integrated GARCH (FIGARCH) is proposed by Baillie *et. al.* [11]. Its primary purpose is to develop a more flexible class of processes with conditional variance more capable to explain and represent the observed temporal dependency in the financial market volatility. The FIGARCH(p,d,q) process for ε_t is defined by

$$\phi(L)(1-L)^d \varepsilon_1^2 = w + [1 - \beta(L)]\upsilon_t, \tag{2.20}$$

and equivalently,

$$[1-\beta(L)]\sigma_t^2 = w + [1 - \beta(L) - \phi(L)(1-L)^d]\varepsilon_t^2, \tag{2.21}$$

where $0 < d < 1$, and all the roots of $\phi(L)$ and $[1-\beta(L)]$ lie outside the unit cycle. Note that L is a lag operator, $\beta(L) = \beta_1 L + B_2 L^2 + ... + \beta_p L^p$ is the backshift operator, $\phi(L) = \sum_{i=1}^{q-1} \phi_i L^i$ is of order $q-1$ and $\upsilon_t = \varepsilon_t^2 - \sigma_t^2$. This model observes a slow hyperbolic rate of decay for its conditional variance. It is known from [29] that financial data can be described by fractional integrated or fractional integrated GARCH models rather satisfactorily.

We use the fractional integrated model as well as the fractional integrated GARCH model to look at the performances of H in the R/S analysis and Whittle estimator. We use the package, *fSeries*, which is available in *R*. The simulation is carried out 100 times. We present the findings in the following tables.

Table 2.1: Simulation using fractional integrated model

	H	Mean	Variance	$y = \frac{1}{n}\sum_{t=1}^{n} \lvert\frac{H-H_t}{H}\rvert$	$z = \frac{1}{n}\sum_{t=1}^{n} \lvert H_t - H\rvert$
R/S	0.6	0.6088241	0.001408671	0.04911205	0.002946723
	0.7	0.6816487	0.001275037	0.04585656	0.03209959
	0.8	0.744181	0.001912977	0.07468207	0.05974565
	0.9	0.7902214	0.001536024	0.1219762	0.1097786
	1.0	0.8416431	0.001375958	0.1583569	0.1583569
Whittle	0.6	0.5755584	0.000421959	0.04582390	0.02749434
	0.7	0.6593072	0.0005052445	0.05853925	0.04097748
	0.8	0.7427115	0.000509533	0.07161339	0.05729071
	0.9	0.830812	0.0004483971	0.07687548	0.06918793
	1.0	0.921394	0.0005135379	0.07860597	0.07860597

Table 2.2: Simulation using fractional integrated generalized autoregressive conditional heteroskedasticity model

	H	Mean	Variance	$y = \frac{1}{n}\sum_{t=1}^{n} \lvert\frac{H-H_t}{H}\rvert$	$z = \frac{1}{n}\sum_{t=1}^{n} \lvert H_t - H\rvert$
R/S	0.6	0.6048012	0.001304686	0.04963525	0.02978115
	0.7	0.6744801	0.001040458	0.0455274	0.03186918
	0.8	0.734113	0.001259956	0.08493953	0.06795162
	0.9	0.792369	0.001190705	0.1195900	0.1076310
	1.0	0.8373854	0.001551047	0.1626146	0.1626146
Whittle	0.6	0.583954	0.000463034	0.03740914	0.02244549
	0.7	0.6614076	0.000505010	0.05580307	0.03906215
	0.8	0.7403529	0.0004606759	0.07455885	0.05964708
	0.9	0.8327198	0.0005328518	0.07475576	0.06728019
	1.0	0.9219956	0.0005956759	0.0780043	0.07800443

From the findings, we can see that the R/S analysis performs better than Whittle estimator for small value of $H(H = 0.6, 0.7)$. Both the fractional integrated model and fractional integrated GARCH model give 0.6088 and 0.6048 for $H = 0.6$ by using R/S analysis (the actual value should be 0.6). This represents a very small difference from the actual expected value. When $H = 0.7$, the R/S analysis gives rise to 0.6816 and 0.67 (actual value is 0.7), leaving behind the poor performance of Whittle estimator. The values for Whittle estimator are 0.6593 and 0.6614. Only for large $H(H = 0.9, 1.0)$, Whittle estimator has outperformed the R/S analysis. From the literature, we note that most empirical studies show that the value of the Hurst exponent should not be more than 0.9. Thus, we will use the R/S analysis in our investigation of the financial data in the later section.

2.4 Empirical Results

2.4.1 Malaysia Financial Data

We give a brief history on Malaysia financial performance in this section. Malaysia is a fast growing developing country which has experienced a high rate of increase in its gross domestic product (GDP) every year, from 24.489 billion US dollars in 1980 to 138.270 billion US dollars in 2006. The performance of the financial market in Malaysia has received an increasing attention throughout the world. Research conducted attributes the good performance of Malaysia financial market to the impact of the financial liberalization ([69], [52], [131]). It is caused by ASEAN's economic growth started in the late 1980's when ASEAN began its free trade market policy in conjunction with liberal policies for foreign investment [84].

Malaysia is a member of ASEAN countries, and has been one of the main contributors in ASEAN economy. It has the fourth largest GDP after Indonesia, Thailand and Singapore. Its contribution towards the growth of ASEAN economy is clearly noticeable. Comparing with other ASEAN members, Malaysia has a low inflation rate, just 3.1% in 2005, and a low unemployment rate of 3.8% in 2006, which is the second lowest after Thailand. With an estimated population of 26,686,000 in 2006, Malaysia is a relatively small country, yet it has been contributing significantly in the ASEAN economy.

Asian currency crisis erupted in July 1997, starting in Thailand, and not long after, it hit other countries in the region. For Malaysia, it could maintain the value of its Ringgit (RM2.50 to 1 USD) during the first half of the year. However, Ringgit started to depreciate by 35% against the USD. The situation became worst during the first six months of 1998. Ringgit dropped to RM4.88 to one USD on 9 January 1998, the lowest ever recorded since the start of the free trade of Ringgit [111]. In order to stop the free fall of Ringgit, the Government decided to impose the currency control by pegging Ringgit to USD at a fixed exchange rate (RM3.80 to one USD). Edison and Reinhart [42] and Kaplan and Rodrik [72], after careful examination of this decision of pegging Ringgit to USD at a fixed exchange rate, confirmed that this policy has, indeed, stabilized the exchange rate, avoiding the bankruptcy of Malaysia financial institution. This policy has stopped the attack by big overseas funds against Ringgit.

There have been several research projects conducted on Malaysia financial data, such as a test [57] of whether or not Malaysia stock data exhibits a random walk behaviour, and the linkage between exchange rate and stock prices [115], the comparison of Malaysia and Singapore equity market in price discovery [40], and the relationship between the two stock returns [79]. We now turn our attention to investigate whether or not there exists a long memory. The presence of a long memory in returns series of the stock markets in 4 countries of ASEAN is investigated by Navarro et. al in [112] based on ARFIMA model. However, no economic implication is given in their findings.

There are also research projects devoted to the study of the behaviour of Malaysia financial time series in regards to the existence of a long memory. However, these research projects mainly focus on the overall performance of Malaysia economy in comparison with economies of other ASEAN countries ([25], [24], [31], [73], [111]).

We make use of daily observation of closed price Kuala Lumpur Composite Index (KLCI) stock indices, foreign exchange (FX) rates of Malaysia Ringgit to US dollar for the period from 1994 to 2006, and the interest rates from 1998 to 2006. We assume that the returns series of the data is stationary. We understand that the three main indicators for most countries are the stock market indices, the exchange rates, and the interest rate. Unfortunately, we can only get the interest rate data starting from the year of 1998. However, it still meets our purpose - making a preliminary investigation to the estimators of the H index.

2.4.2 Statistical Data

Figure 2.1 reveals some trends about the return series of these three financial indicators. Statistics factors (namely, mean, variance, skewness and kurtosis) are listed in Table 2.3 - Table 2.5 for the three financial indicators of the return series data obtainable from *http://www.econstats.com*.

Figure 2.1 shows the graphs of the plotted data together with their respective returns. From the graphs, we can see that there is quite a big jump in every graph (KLCI, FX and Interest rates). This is due to the financial instability during the Asian financial crisis caused by the attack to many Asian currencies in the middle of 1997. Malaysia is one of the countries severely affected by this crisis. The economy in Malaysia relies heavily on some imported electronic components from overseas. We present some analysis on the data.

Table 2.3: Statistic data of mean, variance, skewness and kurtosis for KLCI

		KLCI		
Year	Mean	Variance	Skewness	Kurtosis
1994	-0.1098426	2.969164	0.2093	8.5198
1995	0.0100	1.4541	0.7335	5.7005
1996	0.0880	0.6641	-0.3895	5.6070
1997	-0.2946	5.5978	0.3748	9.9141
1998	-0.0057	14.6089	0.5370	16.3290
1999	0.1311	2.9886	0.1119	4.4715
2000	-0.0743	1.9390	-0.1903	4.7543
2001	0.0098	1.6343	-0.9729	8.5232
2002	-0.0299	0.6661	0.1462	4.1406
2003	0.0836	0.5313	0.3871	4.1878
2004	0.0541	0.5259	-0.0311	4.1927
2005	-0.0034	0.2380	0.2499	4.0130
2006	0.08	0.2760	-0.5018	5.0728

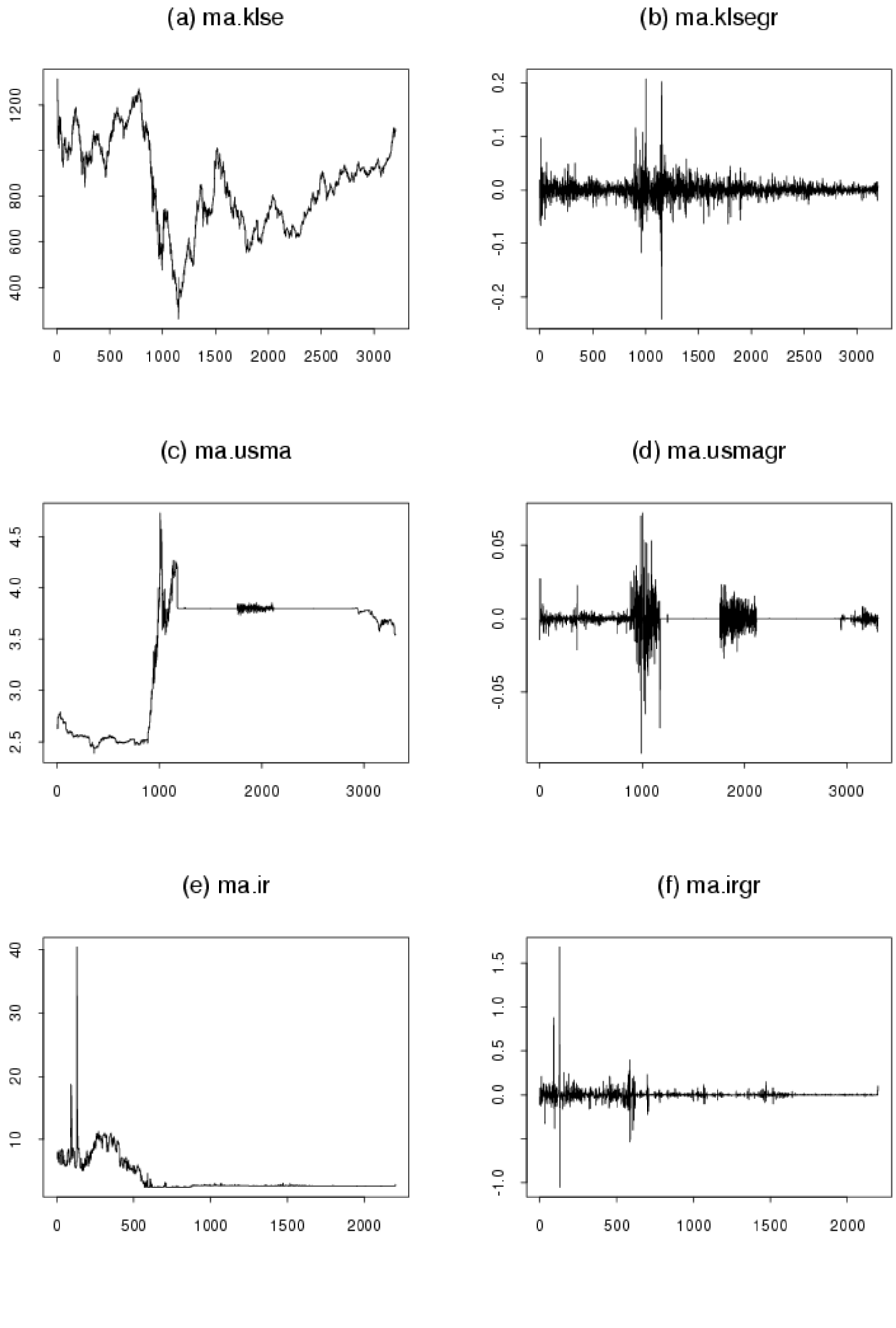

Figure 2.1: The KLCI, foreign exchange and interest rate respective to their return series

Table 2.4: Statistic data of mean, variance, skewness and kurtosis for FX

		FX		
Year	Mean	Variance	Skewness	Kurtosis
1994	-0.0178	0.1033	2.1045	27.4589
1995	-0.0023	0.0799	0.2027	31.2118
1996	-0.0022	0.0196	-0.056	5.3126
1997	0.1715	1.4193	-1.1585	21.9994
1998	-0.0088	2.5916	-0.1193	8.3181
1999	0	5.5185e-7	0	103.3209
2000	0	1.3793e-6	0	126
2001	-1.544e-11	0.6933	0.0307	3.1802
2002	7.722e-12	0.1033	0.2774	8.571
2003	-8.3722e-21	6.7105e-6	-0.1221	2.4781
2004	-3.336e-21	7.647e-6	-0.06	2.3967
2005	-0.0021	0.0054	-5.4963	59.1465
2006	-0.0263	0.0601	0.1453	5.3136

Table 2.3 - Table 2.5 represent the statistical properties of the return series for KLCI, FX and Interest Rates. We can see from the KLCI table that there exists a heavy tail in the 1998 data. A similar condition exists also in the FX data, except for the years of 2003 and 2004. This is probably due to the currency control imposed by the Government by pegging Ringgit to USD at a fixed rate of MYR3.80 to USD 1. For interest rate, a heavy tail is also consistently observed in all the data.

By taking into account the simulation results reported in previous section, we will now consider the R/S analysis for our data. We present the findings in Table 2.6 - Table 2.8.

In Table 2.6 - Table 2.8, a long memory is observed in every financial indicator. We consider these indicators based on the returns, absolute returns and square return series. The findings have suggested the existence of long memory behaviour for KLSE data. On the other hand, there are varieties

Table 2.5: Statistic data of mean, variance, skewness and kurtosis for interest rate

		Interest		
Year	Mean	Variance	Skewness	Kurtosis
1998	-0.219	22.5899	0.0439	5.8124
1999	-0.2883	67.9891	-0.9584	15.0159
2000	0.0402	2.2759	-0.4021	10.2379
2001	-2.0492e-11	4.7535	-0.2238	13.16
2002	-0.0074	3.3638	1.4103	22.7771
2003	0.0015	1.864	-1.3009	22.92
2004	-0.0044	0.0686	0.2175	4.8545
2005	0.0442	0.4991	10.9506	155.1376
2006	0.0617	0.5217	8.6651	89.5470

of outcomes being observed for the FX and interest rates. We summarize the findings in Figure 2.2.

Here, we find that for the case of KLCI, all the yearly data for returns, absolute returns and square returns exhibit a long term dependency behaviour. We use the returns, square returns and absolute returns to analyze the data. This is because the square and the absolute returns can give more information on the volatility in the financial series. The returns series shows that the Hurst exponent takes its value from the minimum of 0.52 to the maximum of 0.63. They show mild long memory behaviors. However, if we use the absolute returns, we can see a clear long memory characteristic during 1997 to 1998, where the Hurst index is 0.78 and 0.75, respectively. This is consistent with the findings of Ding *et. al.* in [41], where it is found that the long memory property is most noticeable in absolute returns.

Our investigation of the financial activities in Malaysia over these particular years shows that these are the turbulence years in Malaysia economy, the worst ever since its independence in 1957. Malaysia suffers the financial crisis basically due to the attack of its currency. The exchange rate of MYR2.50 to USD1 before the crisis jumped up quickly to MYR4.8 to USD 1 after the attack. The Hurst exponent for the square return during 1997 is large. There was also political instability during these years. The political disagreement between the then Prime Minister and his Deputy ended

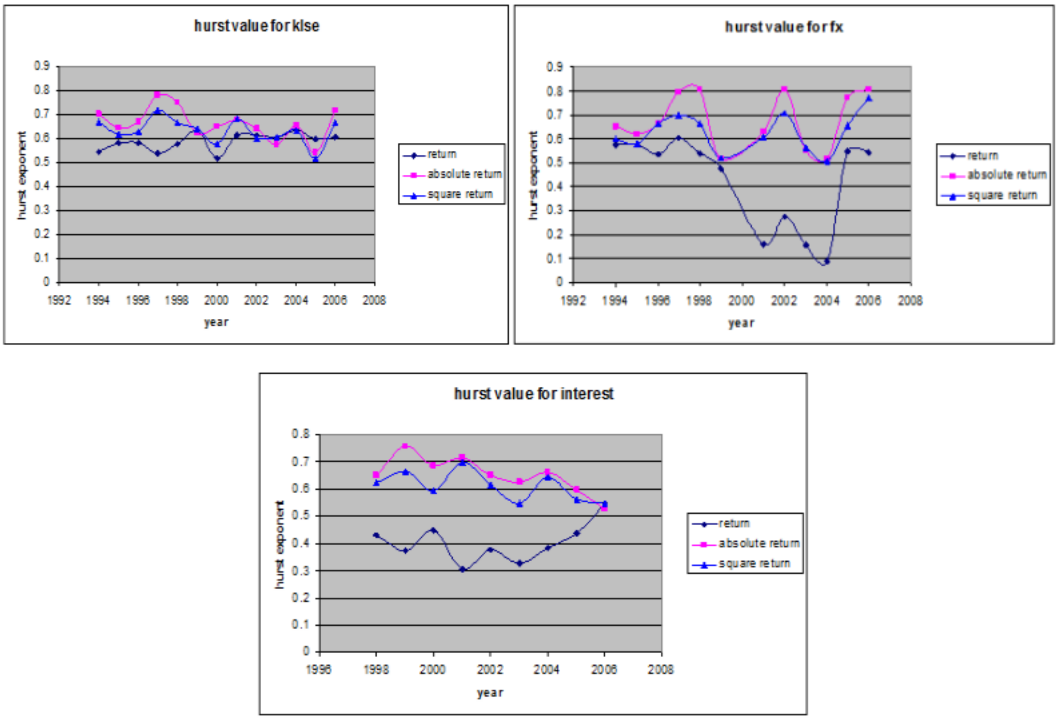

Figure 2.2: Hurst values for returns, absolute and square returns

Table 2.6: The value of Hurst exponent for KLCI

Year	Returns	Absolute returns	Square returns
1994	0.5444799	0.6982398	0.6664263
1995	0.5818081	0.6420021	0.6144591
1996	0.5787778	0.6679039	0.6262075
1997	0.5342115	0.7816373	0.7142305
1998	0.5743141	0.7499512	0.6644766
1999	0.6321186	0.6214217	0.6380659
2000	0.5155318	0.6480894	0.5743059
2001	0.6128591	0.6761504	0.6783971
2002	0.6139401	0.6395184	0.6005312
2003	0.6050676	0.5729191	0.6039314
2004	0.6321451	0.6526607	0.6305329
2005	0.5970594	0.540061	0.5141837
2006	0.6048106	0.7170437	0.663979

up with the sacking of Anwar Ibrahim in 1998. He was arrested on some charges not long after. This led to several protests led by Anwar's followers. The economy and political instability are believed to be the two causes for the observed behaviour in the market during these years.

The FX also exhibits long memory behaviour during the period from 1994 to 1998 for the return series. The values obtained for the period from 1999 to 2004 do not show long memory behaviour. This may be due to the currency control imposed by the Government during these years. It was imposed to prevent the free fall of Malaysian currency by pegging Ringgit to USD at a fixed rate of MYR3.8 to USD 1. The value of the Hurst exponent is significantly larger for 2005 and 2006 soon after the pegging was lifted by the Government, indicating long memory behaviour.

If we look at the absolute return, we can see almost similar responses in comparison to the KLCI situation. There are strong long memory behaviours during 1997 and 1998. We also notice a large Hurst exponent in 2002. We learn that during this year, Malaysia economy was slowing down due to the global crisis caused by the attack to the World Trade Centre on 11 September, 2001. We also find that the resignation of the former Prime Minister, Dr. Mahathir Mohammad, was occurred in this year, after 21

Table 2.7: The value of Hurst exponent for FX

Year	Returns	Absolute returns	Square returns
1994	0.5759926	0.6528175	0.600992
1995	0.5748038	0.6205861	0.5785149
1996	0.5340298	0.662195	0.6645818
1997	0.60505235	0.7965435	0.7009597
1998	0.5382436	0.8096447	0.6628409
1999	0.4755805	0.5164264	0.5212018
2001	0.1588013	0.6302862	0.6098399
2002	0.2749463	0.8089852	0.7090517
2003	0.1550481	0.5500234	0.5634557
2004	0.08788315	0.5162005	0.5098352
2005	0.5495628	0.7706903	0.6568638
2006	0.5457544	0.8065743	0.7713298

Table 2.8: The value of Hurst exponent for interest rate

Year	Returns	Absolute returns	Square returns
1998	0.4301006	0.6518281	0.6210953
1999	0.372699	0.7575185	0.6637135
2000	0.4492246	0.6853204	0.5922305
2001	0.3068688	0.7161545	0.6957458
2002	0.3763427	0.6505131	0.6160992
2003	0.3257871	0.6253157	0.5491196
2004	0.3845908	0.6630796	0.6427257
2005	0.4387428	0.5980355	0.5632757
2006	0.5480132	0.5263451	0.5466853

years of serving as the Prime Minister - the longest in Malaysian history. The square returns of FX also show similar behaviours as those of KLCI square returns.

We also run the Hurst exponent test to interest rates and observe that there exist long memory in absolute returns and square returns, even though there is a short memory dependency in the return series of the interest rates. We can also see from all these three graphs that the absolute values produce the largest Hurst exponents in most of the data. This finding is consistent with the result reported by Cajeiro in [23].

2.5 Discussion

In this chapter, we made a preliminary investigation to the Hurst parameter in Malaysia financial data. We used the fractional integrated model and fractional integrated GARCH model in our simulation study, looking at the performance corresponding to each of the estimators. Then, we used the R/S analysis to compare the performances of the returns, absolute returns and square returns series.

The R/S analysis is practical and straightforward to use. However, there are many deficiencies in association with these estimators in real applications. For example, there is no asymptotic distribution that could be derived [89]. Hence, the significance of the findings could be subjective in nature. In view of this drawback, it requires a large data set so as to be effective [142]. However, in real world applications, a large date set could come with a high price. Normally, we only look at financial data of reasonable length. This is more useful and beneficial for our understanding of the behaviour of the data.

The R/S analysis also assumes that the data is stationary. This means that the underlying process remains the same throughout the sample. This may not be true for most of the financial data. There is also a criticism on its sensitivity to short range correlation [85]. In fact, we also found that most of the estimators available in the literature produce different results when they are applied to the same set of data. This gives room for doubts in the reliability of these estimators.

However, these methods are also having many advantages. In addition to the fact that they are easy to use, the approaches used in most of these methods are nonparametric. This means that we do not need to impose prior assumption on the structure of the underlying model. This is very helpful and is much cheaper when compared with those based on the parametric approach. Furthermore, the R/S analysis gives consistent results when there exists a long range dependency in the data under consideration [111].

Chapter 3

Fractional Black-Scholes models: Complete maximum likelihood estimation with application to fractional option pricing

3.1 Introduction

In this chapter, we will look at one of the model based estimators of long memory indexes. One of the most important models in financial world is geometric Brownian motion (GBM) introduced by Samuelson in [125]. This model is widely used as the underlying process of a risky market. The values of the parameters obtained from this model are used in the famous Black-Scholes model for option pricing.

Geometric Brownian motion (GBM) is a continuous time stochastic process describing the dynamics of stock price in Black-Scholes option pricing model. With this in mind, our aim in this chapter is to incorporate long memory properties to the standard GBM. This is achieved by replacing Brownian motion (BM) with fractional Brownian motion (FBM) in the model, leading to Geometric Fractional Brownian motion (GFBM) model.

The values of the parameters involved in GFBM are important, as they will be used in the fractional Black-Scholes model, an extension of Black-Scholes model with long memory properties. Since the work by Hurst in early 1950s, it is still an active research area for the investigation of the best suited index of self similarity, H. This problem becomes particularly apparent in the study of financial market. In fact, in addition to the Hurst parameter, we should also estimate other important parameters, such as the volatility parameter and drift parameter. They have influences on the financial model. There are two main objectives in the financial world - minimizing the risks while maximizing the profit.

Works on estimating parameters in problems involving fractional Brownian motion in model based approaches are recently becoming very popular among researchers. Hult approximates some Volterra type stochastic integrals in [61] and also uses the approach of the maximum likelihood estimator to get the values of the parameters in Ornstein-Uhlenbeck process in [62]. Further in this chapter, we will consider GFBM as a model to estimate the long memory index, as well as other parameters involved in this model. For practical applications in the fractional Black-Scholes market, it is essential to know the values of the parameters in GFBM. In particular, there are two key parameters, the volatility σ and the long memory parameter H, that are crucially important in valuing, say, European option [107].

However, it appears that there are few works addressing the estimators of these parameters in the literature. An exception is the paper by Kukush *et al.* [78], where an incomplete maximum likelihood estimation (IMLE) method for the volatility σ is developed. However, the estimation of the long memory parameter H is carried out separately by some specially designed estimation methods, such as the R/S analysis or the variation analysis. As discussed in Chapter 2, no single method is most suited to all situations for the estimation of H. Consequently, no matter what method is used, it may still lead to poor results in some situations.

In this chapter, we study the problem of estimating the unknown parameters, which include the drift μ, volatility σ and the Hurst index H, involved in GFBM. The estimation is carried out based on discrete observations in the range of $0 < H < 1$. Unlike [78], we propose a complete maximum likelihood estimation (CMLE) approach, which enables us not only to derive the estimators of μ and σ^2, but also the estimator of the long memory parameter, H, simultaneously, for the risky assets in the fractional Black-Scholes market governed by GFBM with large sample distributions. The simulation outcomes will illustrate that the CMLE approach is effective and reliable for GFBM model. The method for estimating σ^2 and H separately by IMLE and the widely used R/S analysis may lead to poor results.

3.2 Overview on Some Estimators for Stochastic Differential Equations

We start this subsection by first introducing the definition of Geometric Brownian motion (GBM).

Definition 3.2.1. *A stochastic process S_t is said to follow a GBM if it satisfies the following stochastic differential equation:*

$$dS(t) = \mu S(t)dt + \sigma S(t)dB(t) \tag{3.1}$$

where $B(t)$ is a Brownian motion, and μ, the drift, and σ, the volatility, are constants.

For any arbitrary initial value s_0, the equation has the analytic solution

$$S(t) = s_0 \exp(\sigma B(t) + \mu t - \frac{1}{2}\sigma^2). \tag{3.2}$$

Both Brownian motion and geometric Brownian motion are important tools in finance. They are the two core ingredients for many concepts in finance. The efficient market hypothesis (EMH), for example, is being developed by using either Brownian motion or geometric Brownian motion, so is the valuation of options. However, researchers have begun to pursuit for other alternatives due to the occurring of crashes in financial market - a clear evidence showing the deficiency of using EMH and Brownian motion. This is the motivation for the introduction of fractional Brownian motion (FBM) in finance. The main idea is rather simple. We just introduce the long memory parameter, H, to Brownian motion. By doing this, we can capture the long memory phenomenon in the financial market behaviour so that we can get a better understanding of a real financial market.

Remark 3.2.1. *Let $X(t)$ be a geometric fractional Brownian motion defined by*

$$dX(t) = \mu X(t)dt + \sigma X(t)dB_H(t), \qquad t \geq 0, \qquad X(0) = x > 0, \quad (3.3)$$

with x a constant, where $B_H(t)$ is a one-dimensional FBM with $H \in (0,1)$. By using the Wick calculus developed by Hu and Øksendal in [60], the solution of this equation is found to be

$$X(t) = x \exp(\sigma B_H(t) + \mu t - \frac{1}{2}\sigma^2 t^{2H}). \quad (3.4)$$

However, works, which make use of fractional Brownian motion (FBM) in finance, do not arouse too much interest in the early years. The use of pathwise integration theory shows that mathematical markets based on FBM could have arbitrage opportunity, and hence the usefulness of this theory in financial modelling may be in doubt [121]. Nonexistence of arbitrage in a market is viewed as the basic equilibrium condition for the market [15]. This problem has discouraged further investigation in this field for quite a number of years. Only until recently, there are researchers who work on FBM, where the ordinary product pathwise is used as an alternative approach. In this way, promising results have emerged, producing no arbitrage condition.

Consequently, much works have been actively carried out by researchers using FBM as an underlying process in mathematical market models. Hu and Oksendal [60] proved that the white noise calculus based on FBM with $1/2 < H < 1$, corresponding to Ito type fractional Black-Sholes market, has no arbitrage and the market is complete. There are, however, critical views in regard of this approach. Bjork and Hult [17] question the meaning of self-financing used in this framework and they point out that the self financing defined in the study does not give any meaningful explanation in economic perspective. They, however, agree that the method used does not admit arbitrage, and therefore, further investigation into the geometric version of fractional Brownian motion should be carried out. It is called geometric fractional Brownian motion (GFBM). Since the product pathwise representation has been developed for FBM, works related to financial markets using GFBM have started to blossom again in finance literature.

Under this new framework, the option pricing under geometric fractional Brownian motion is being developed. It covers the famous Black-Scholes option pricing as a special case, when $H = 0.5$. For latest works, see, for example, [44], [15], [107] and [123].

For comparison, we show two sample paths, one generated by Brownian motion and one generated by fractional Brownian motion. We can see that the sample path generated by FBM with $H = 0.7$ seems mimic better the financial environment when compared to that generated by the random walk process.

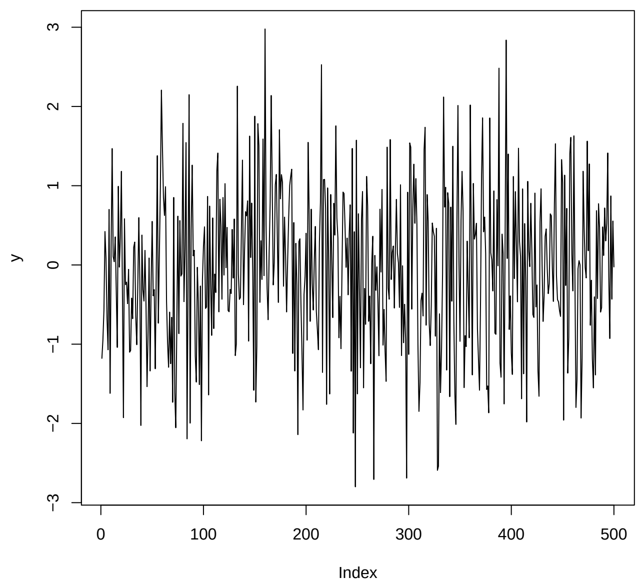

Figure 3.1: Sample path generated by Brownian motion.

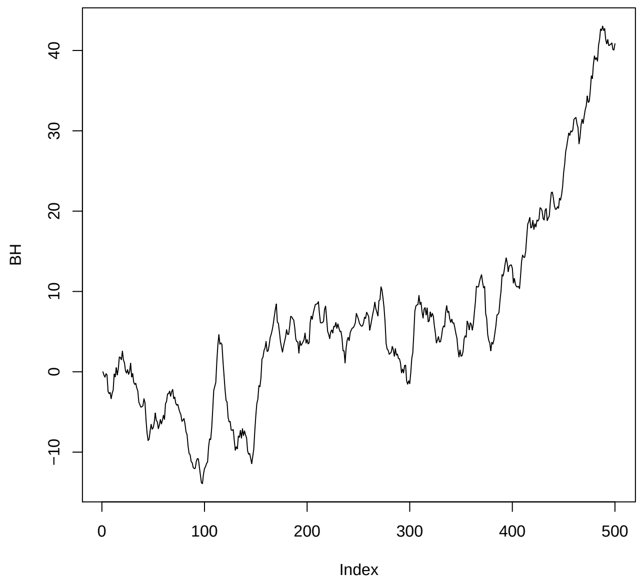

Figure 3.2: Sample path generated by fractional Brownian motion.

For geometric fractional Brownian motion (GFBM), by using the Wick-Ito-Skorohod (WIS) integration in [15], the solution for the stochastic differential equation

$$dS(t) = \mu S(t)dt + \sigma S(t)dB_H(t), \tag{3.5}$$

is given by

$$S(t) = s \ \exp\{\sigma B_H(t) + \mu t - \frac{1}{2}\sigma^2 t^{2H}\}, \tag{3.6}$$

where S is the risky assets, $\mu, \sigma \neq 0$ and $s > 0$ are constants. As mentioned earlier, there is no "the best method" available to estimate the parameters involved in this model.

3.3 Overview on Option Pricing

3.3.1 Option Valuation

Traders in financial world started to venture on the option trading in the early 1960s. However, the first option exchange, called Chicago Board Option Exchange (CBOE) was formally founded in US in 1973 where 16 securities of call option were being traded. By the end of the year, over one million contracts had been traded. Few years later, they introduced the put options and subsequently the index options started to appear ten years later. The trading of options has become increasing popular - as many as 18 million option contracts were traded in 1975 and the number was increased to 60 million by 1978. This promising trend continued until the Great Stock Market Crash in 1987.

For option, investors can decide whether to make profit to an investment or to protect the investment by hedging the risk involved. It is still unknown exactly when the first option contract ever took place. In earlier history, it appeared that Romans and Phoenicians used a similar contract in shipping. In 1600s, people in Holland also used the idea of option when trading tulips. In [45], an option on a stock is defined by Elliott and Kopp as a contract that gives the owner the right, but not the obligation, to trade a given number of shares of a common stock for a fixed price at a future date (the expiry date T). A call option gives the owner the right to buy stocks, whereas a put option gives the right to sell at the fixed expiry date T. The option is called European option when it can only be exercised at the fixed expiry date T. It is called American option if the owner can exercise her right to trade at any time up to the expiry date.

3.3.2 Classical Black Scholes Model

As the first organized option trading took place in 1973, another important discovery was made by Fischer Black and Myron Scholes [18] to cater for option trading. They construct the model describing the market value of a call option. It is now known as Black-Scholes option pricing model, where the price of a call option to buy stock at a specific price and at a specific time is formulated as:

$$C_0 = S_0 N(d_1) - Ke^{-rT} N(d_2). \tag{3.7}$$

Here, C_0 is the price of a call option, S_0 is the current stock price, K is the exercise price, r is the risk-free interest rate, and T is the time to maturity. $N(d_1)$ and $N(d_2)$ are probabilities of the random numbers, d_1 and d_2, respectively, with bell-curve distributions given below.

$$d_1 = \frac{\ln(S_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}} \tag{3.8}$$

$$d_2 = d_1 - \sigma\sqrt{T}, \tag{3.9}$$

where σ is the standard deviation of the stock price. In this model, the price is assumed to follow a geometric Brownian motion with a constant drift and volatility. A stochastic process $S(t)$ is said to follow a GBM if it satisfies the following stochastic differential equation:

$$dS(t) = \mu S(t)dt + \sigma S(t)dB(t) \tag{3.10}$$

where $B(t)$ is a Brownian motion and μ (i.e., the drift) and σ (i.e., the volatility) are constants. The analytic solution of this stochastic differential equation corresponding to an arbitrary initial value S_0 is:

$$S(t) = S_0 \exp((\mu - \frac{1}{2}\sigma^2)t + \sigma B(t)) \tag{3.11}$$

3.3.3 Fractional Black-Scholes Model

Applications of FBM and GFBM to financial option pricing are a natural extensions of the famous Black-Scholes option theory based on BM and GBM. However, works related to FBM in the early years where the pathwise integration theory is used, are subject to criticism, as the mathematical markets based on FBM could have arbitrage opportunity. Thus its usefulness in financial modelling is in doubt [121]. This problem had discouraged further investigation in this field for many years. It is only recently that researchers working on FBM have used ordinary product pathwise as an alternative approach. This gives rise to promising results, producing no arbitrage situation. Consequently, it has stimulated active research with FBM taken as an underlying process in mathematical market models.

Hu and Oksendal [60] proved that white noise calculus based on FBM with $1/2 < H < 1$ for Ito type fractional Black-Scholes market, has no arbitrage and the market is complete. Elliot and van der Hoek [46] extend H into the range of $[0, 1]$. They work on option pricing and consider FBM as the driving noise process. Though there is some criticism regarding this approach, the option pricing under GFBM has been well developed based on this new approach, with Black-Scholes option pricing as a special case (when $H = 0.5$).

Mishura [107] shows that the price at time $t_0 \in [0, T_0]$ of a European call option with the strike price K and maturity T_0 is given by

$$\begin{aligned} C(t_0, S) = {} & S\Phi\left(\frac{\ln\frac{S}{K} + r(T_0 - t_0) + (T_0^{2H} - t_0^{2H})\frac{\sigma^2}{2}}{\sigma\sqrt{T_0^{2H} - t_0^{2H}}}\right) \\ & - Ke^{-r(T_0 - t_0)}\Phi\left(\frac{\ln\frac{S}{K} + r(T_0 - t_0) - (T_0^{2H} - t_0^{2H})\frac{\sigma^2}{2}}{\sigma\sqrt{T_0^{2H} - t_0^{2H}}}\right), \end{aligned} \quad (3.12)$$

where S is the underlying stock price at time t_0, r is the risk free interest rate, and $\Phi(\cdot)$ is the cumulative function of a standard normal distribution. Note that it coincides with the solution of the usual Black-Scholes option pricing if $H = \frac{1}{2}$.

3.4 Model Simplification

We are concerned with fractional Black-Scholes markets, in which the risky asset price process, $S(t)$, driven by FBM is modeled by GFBM, in the form

$$dS(t) = \mu S(t)dt + \sigma S(t)dB_H(t), \tag{3.13}$$

where $S(0) = s > 0$ and μ and $\sigma > 0$ are the drift and volatility, respectively. The solution to this fractional differential equation [60] is given by

$$S(t) = s\ \exp\{\sigma B_H(t) + \mu t - \frac{1}{2}\sigma^2 t^{2H}\}. \tag{3.14}$$

The estimations of the Hurst index, H, and the volatility in this model is particularly important in financial asset pricing. Later, we will show how to estimating them.

We begin with a review of the incomplete likelihood estimation proposed by Kukush *et al.* in [78]. Let

$$X(t) = \ln(\frac{S(t)}{s}). \tag{3.15}$$

Then, it follows from (3.14) that

$$X(t) = \sigma B_H(t) + \mu t - (\frac{\sigma^2}{2})t^{2H}, \qquad t \geq 0. \tag{3.16}$$

As in [78], we assume that the historical data are observed at discrete times $t_k = \frac{kT}{n}$, $k = 0, 1, ..., n$, over the time interval $[0, T]$. By setting $X_k = X(t_k)$ and $B_{Hk} = B_H(t_k)$ and considering $k = 1, ..., n$, we have

$$\Delta X_k = \sigma \Delta B_{Hk} + \mu \Delta t_k - \frac{\sigma^2}{2}\Delta(t^{2H})_k, \tag{3.17}$$

where $\Delta X_k = X_k - X_{k-1}$, ΔB_{Hk} and Δt_k are defined similarly, and $\Delta(t^{2H})_k = t_k^{2H} - t_{k-1}^{2H}$.

Let us briefly describe the incomplete maximum likelihood estimation (IMLE) procedure developed by Kukush *et al.* [78]. First, it is assumed that H can be estimated in advance by some well designed estimation methods for H in the literature, such as the R/S analysis and the variation analysis, Then, they set

$$Y_k = \frac{n^H \Delta X_k}{T^H}, \tag{3.18}$$

and write (3.17) as

$$Y_k = \sigma\varepsilon_k + \frac{n^H \mu \Delta t_k}{T^H} - \frac{1}{2}\sigma^2 T^H n^H \Delta\tau_k^{2H} \tag{3.19}$$

for $k = 1, ..., n$, where $\Delta\tau_k^{2H} = (\frac{k}{n})^{2H} - (\frac{k-1}{n})^{2H}$, and $\varepsilon_k = \frac{n^H \Delta B_{Hk}}{T^H}$.

Simple calculation shows that ε_k is normally distributed with $E\{\varepsilon_k\} = 0$, $E\{\varepsilon_k^2\} = 1$ and the covariance of ε_k is the same as in (3.24). Using (3.19), Kukush *et al.* [78] then suggest an IMLE of the volatility σ, based on Y_k, with

$$\hat{\sigma}^2_{\text{IMLE}} = \frac{1}{n}\sum_{k=1}^{n}(Y_k - \bar{Y})^2, \tag{3.20}$$

where $\bar{Y} = \frac{1}{n}\sum_{k=1}^{n} Y_k$. This estimation method is applied to option pricing by Cajueiro and Barbachan in [23]. Note that $\hat{\sigma}^2_{\text{IMLE}}$ is just the usual sample variance of Y_k. It is essentially the same as being assumed that the Y_k in the model (3.19) is stationary. This, however, may not be true in general for the model (3.19), because $\Delta\tau_k^{2H}$ will depend on k if $H \neq 0.5$. Under some appropriate conditions imposed on $n = n(T)$, such as $\frac{n}{T^{2H}} \to \infty$ as $T \to \infty$ for $0.5 < H < 0.75$, it is shown that $\hat{\sigma}^2_{\text{IMLE}}$ is consistent (see Kukush *et. al.* ([78], page 88)). However, it may not be statistically justifiable, especially when the sample size n does not converge to ∞ faster than T. Furthermore, a bad estimate of H that is obtained in advance may lead to a poor estimate of σ^2.

In this work, we study the problem of estimating the unknown parameters, which include the drift, μ, volatility, σ, and Hurst index, H, involved in the GFBM based on the discrete observations in the setting of $0 < H < 1$. Unlike Kukush et al. [78], we propose a complete maximum likelihood estimation (CMLE) approach, which enables us to estimate μ, σ^2 and H simultaneously. We will follow an alternative approach and consider the return series $Z_k = \Delta X_k$, rather than Y_k, as follows (following from (3.19)):

$$\begin{aligned}
Z_k = \Delta X_k &= (\frac{T}{n})^H Y_k \\
&= (\frac{T}{n})^H \sigma \varepsilon_k + \mu \frac{T}{n} - \frac{1}{2}\left\{(\frac{T}{n})^H \sigma\right\}^2 n^{2H} \Delta\tau_k^{2H} \\
&\equiv \sigma_1 \varepsilon_k + \mu_1 - \frac{1}{2}\sigma_1^2 n^{2H} \Delta\tau_k^{2H},
\end{aligned} \tag{3.21}$$

where $\sigma_1 = (\frac{T}{n})^H \sigma$ and $\mu_1 = \frac{\mu T}{n}$. We construct our complete maximum likelihood estimation based on (3.21).

3.5 Complete Maximum Likelihood Estimation (CMLE)

In this subsection, we are concerned with the estimation of $\theta = (\sigma_1^2, \mu_1, H)'$ by using the method of CMLE. Here, A' stands for the transpose of a vector or a matrix A.

3.5.1 Likelihood function of $\theta = (\sigma_1^2, \mu_1, H)'$

Based on (3.21), our observations are $Z = (Z_1, ..., Z_n)'$, and for notational convenience, set $\mathbf{x}_H = n^{2H}(\Delta\tau_1^{2H}, ..., \Delta\tau_n^{2H})'$ and $\varepsilon = (\varepsilon_1, ..., \varepsilon_n)'$.

Then, the vector form of (3.21) is as follows:

$$Z = \sigma_1 \varepsilon + \mu_1 \mathbf{1} - \frac{1}{2}\sigma_1^2 x_H, \tag{3.22}$$

where $\mathbf{1}$ is the n-dimensional vector of components 1's.
Set

$$\Sigma = var(Z) = \sigma_1^2 (E\varepsilon\varepsilon')_{n\times n} = \sigma_1^2 \Sigma_0, \tag{3.23}$$

with $\Sigma_0 = \Sigma_0(H) = (\gamma_{ij})_{n\times n}$ given by

$$\gamma_{ij} = E\varepsilon_i\varepsilon_j = \frac{1}{2}(|i-j+1|^{2H} - 2|i-j|^{2H} + |i-j-1|^{2H}). \tag{3.24}$$

Since the process is Gaussian, the log likelihood for Z is given below.

$$\begin{aligned}\ell_n(\theta) &= -\frac{1}{2}\log|\Sigma| - \frac{1}{2}(Z - \mu_1\mathbf{1} + \frac{1}{2}\sigma_1^2\mathbf{x}_H)'\Sigma^{-1}(Z - \mu_1\mathbf{1} + \frac{1}{2}\sigma_1^2\mathbf{x}_H) \\ &= -\frac{1}{2}(n\log\sigma_1^2 + \log|\Sigma_0|) - \frac{1}{2\sigma_1^2}(Z - \mu_1\mathbf{1} + \frac{1}{2}\sigma_1^2\mathbf{x}_H)'\Sigma_0^{-1}(Z - \mu_1\mathbf{1} + \frac{1}{2}\sigma_1^2\mathbf{x}_H).\end{aligned} \tag{3.25}$$

Therefore, the CMLE of $\theta = (\sigma_1^2, \mu_1, H)'$ is

$$\hat{\theta} = (\hat{\sigma_1^2}, \hat{\mu_1}, \hat{H})' = \arg\max_{\theta\in\Theta} \ell_n(\theta), \tag{3.26}$$

where Θ is a compact subset of $\mathbb{R}^+ \times \mathbb{R} \times (0, 1)$, which contains the actual parameter vector $\theta_0 = (\sigma_{10}^2, \mu_{10}, H_0)'$.

We finally obtain the estimators of σ^2 and μ as follows:

$$\widehat{\sigma}^2 = \left(\frac{n}{T}\right)^{\hat{H}} \widehat{\sigma}_1^2, \tag{3.27}$$

and

$$\hat{\mu} = \frac{n}{T}\hat{\mu}_1. \tag{3.28}$$

3.5.2 Algorithm

Our aim now is to calculate $\hat{\theta}$ in (3.26). Maximizing (3.25) directly is quite involved. We suggest a profile method to simplify the calculation. For a given H, we can derive the maximum likelihood estimators for σ_1^2 and μ_1, by maximizing (3.25) with respect to σ_1^2 and μ_1. They are achieved by setting the first order partial derivatives of $\ell_n(\theta)$ with respect to σ_1^2 and μ_1 equal to 0 (see Appendix for the details of the derivation), giving

$$\widehat{\sigma}_1^2 = \frac{2Z'\Sigma_1 Z}{\sqrt{n^2 + \mathbf{x}_H'\Sigma_1\mathbf{x}_H Z'\Sigma_1 Z} + n} \tag{3.29}$$

and

$$\hat{\mu}_1 = \frac{1}{\mathbf{1}'\Sigma_0^{-1}\mathbf{1}}(\mathbf{1}'\Sigma_0^{-1}Z + \frac{1}{2}\widehat{\sigma}_1^2\mathbf{1}'\Sigma_0^{-1}\mathbf{x}_H), \tag{3.30}$$

where

$$\Sigma_1 = \Sigma_0^{-1}\left(\mathbf{I} - \frac{\mathbf{1}\mathbf{1}'\Sigma_0^{-1}}{\mathbf{1}'\Sigma_0^{-1}\mathbf{1}}\right), \tag{3.31}$$

with $\mathbf{I}$ being the $n \times n$ identity matrix.

Now, in order to estimate H, we replace σ_1^2 and μ_1 in (3.25) by (3.29) and (3.30). Consequently, we obtain

$$\begin{aligned}\ell_{1n}(H) &= \ell(\widehat{\sigma}_1^2, \hat{\mu}_1, H)\\ &= -\frac{1}{2}(n \log \widehat{\sigma}_1^2 + \log |\Sigma_0|)\\ &- \frac{1}{2\widehat{\sigma}_1^2}(Z - \hat{\mu}_1\mathbf{1} + \frac{1}{2}\widehat{\sigma}_1^2\mathbf{x}_H)'\Sigma_0^{-1}(Z - \hat{\mu}_1\mathbf{1} + \frac{1}{2}\widehat{\sigma}_1^2\mathbf{x}_H). \qquad (3.32)\end{aligned}$$

This is a function of H. Taking the differentiation of $\ell_{1n}(H)$ with respect to H is difficult. However, note that $\ell_{1n}(H)$ is a univariate profile likelihood function of H. There are many available numerical methods that can be used to maximize $\ell_{1n}(H)$ without using the information of differentiation, for example, Golden Section Search method. Thus, we can get the estimator $\hat{H}$ of H.

With this in mind, we propose an algorithm given below.

- Maximize (3.32) numerically to get the estimator, $\hat{H}$, of H.
- Calculate the estimators, $\hat{\sigma}_1^2$ and $\hat{\mu}_1$, by replacing H with $\hat{H}$ in (3.29) and (3.30), respectively.
- Compute the estimators of σ^2 and μ by (3.27).

3.5.3 Large sample property

It is well known that the complete maximum likelihood estimate is statistically efficient. For $\hat{\theta} = (\hat{\sigma_1^2}, \hat{\mu_1}, \hat{H})'$, we have

$$I_{F,n}^{1/2}(\hat{\theta} - \theta) \rightsquigarrow N(0, I_3), \qquad (3.33)$$

as $n \to \infty$, where

$$I_{F,n} = -E[\frac{\partial^2 \ell_n(\theta)}{\partial\theta\partial\theta'}] \qquad (3.34)$$

is the Fisher information matrix, where the convergence in distribution is used. I_3 is the 3×3 identity matrix. Therefore, for $\hat{\vartheta} = (\hat{\sigma^2}, \hat{\mu}, \hat{H})'$ if $\frac{n}{T} \to c_0$,

$0 < c_0 < \infty$, we have

$$V_n^{-1/2}(\hat{\vartheta} - \vartheta) \rightsquigarrow N(0, I_3), \tag{3.35}$$

where $\vartheta = (\sigma^2, \mu, H)'$, $V_n = CI_{F,n}^{-1}C'$, and

$$C = \begin{pmatrix} c_0^H & 0 & \sigma^2 \log c_0 \\ 0 & c_0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

In particular, if $c_0 = 1$, then $C = diag(1, 1, 1)$ is an identity matrix. For this asymptotic property, the requirements on the sample size n and T are much simpler and milder than that in Kukush *et al.* [78], where the consistency property of $\hat{\sigma}^2_{IMLE}$ is obtained by using IMLE under some restrictive sample size conditions.

The proof of the large sample property is very involved and lengthy. Thus, it is omitted and we refer the interest reader to [37] for the proof. The derivation of $I_{F,n}$ is highly complicated. An alternative method of estimating the asymptotic variance is by using the Bootstrap simulation, which we will do in the next section.

3.6 Simulation Study

In order to examine the performance of the proposed estimators, we have carried out some simulation experiments. Let us describe how the data is generated. We first consider the model (3.14). As in the last section, we take $t_k = \frac{kT}{n}$. Note that $B_H(t_k)$ has Gaussian distribution with $EB_H(t_k) = 0$ and $E(B_H^2(t_k)) = t_k^{2H} = (\frac{kT}{n})^{2H}$. With this property, equation (3.14) becomes

$$S_k = S(t_k) = s \exp[\sigma(\frac{T}{n})^H B_H(k) + \mu(\frac{kT}{n}) - \frac{1}{2}\sigma^2(\frac{kT}{n})^{2H}]. \tag{3.36}$$

We take the parameters $\mu = 0.2752908$, $\sigma^2 = 0.2554078$, $H = 0.549$, and the initial value of $s = 903.84$. We simulate the time series from this discrete time model and apply our methodology to estimate the parameters $\vartheta = (\sigma^2, \mu, H)'$ using the simulated data set. The simulation is repeated one hundred times.

To have an idea on the performance of the estimators suggested by Kukush *et al.* [78], we also consider the estimation method by Kukush *et al.* as a comparison. No doubt, the R/S analysis of Hurst [63] and Mandelbrot ([93], [94]) is the most widely used method for the estimation of the Hurst index in the literature. See also Mandelbrot and Taqqu [98] and Mandelbrot and Wallis ([100], [101], [103], [102]). We use the Hurst value obtained from the R/S analysis method. The simulated outcomes of the average value of estimates based on 100 replications, with bias and variance, are reported in Tables 3.1–3.4, for $T = 15$, $T = 30$, $T = 40$ and $T = 50$, respectively. The 5 cases of sample sizes $n = 100, 200, 300, 400, 500$ are considered in each table, where

- $\hat{H}_{CMLE}$ = Hurst index obtained by using the method proposed in this chapter;
- $\hat{\mu}_{CMLE} = \mu$ obtained by the method proposed in this chapter;
- $\hat{\sigma}^2_{CMLE} = \sigma^2$ obtained by the method proposed in this chapter;
- $\hat{H}_{RS}$ = Hurst index obtained by using the method of R/S analysis;
- $\hat{\sigma}^2_{IMLE} = \sigma^2$ obtained by the method of Kukush et. al (2005) with $\hat{H}_{RS}$.

Table 3.1: Outcome of simulation with $T = 15$: average value of estimates based on 100 replications, with bias in ( ) and variance in [ ]

n	100	200	300	400	500
$\hat{H}_{CMLE}$	0.5378 (−0.0112) [0.0019]	0.5395 (−0.0095) [0.0011]	0.5446 (−0.0044) [0.0012]	0.5438 (−0.0052) [0.0008]	0.5409 (−0.0081) [0.0009]
$\hat{\mu}_{CMLE}$	0.2590 (−0.0163) [0.0321]	0.2593 (−0.0160) [0.1022]	0.3199 (0.0446) [0.0270]	0.2512 (−0.0240) [0.0254]	0.2697 (−0.0056) [0.0159]
$\hat{\sigma}^2_{CMLE}$	0.2439 (−0.0115) [0.0043]	0.2424 (−0.0131) [0.0025]	0.2520 (−0.0034) [0.0037]	0.2500 (−0.0054) [0.0029]	0.2448 (−0.0106) [0.0034]
$\hat{H}_{RS}$	0.6575 (0.1085) [0.0318]	0.6275 (0.0785) [0.0200]	0.6099 (0.0609) [0.0141]	0.6326 (0.0836) [0.0113]	0.5969 (0.0479) [0.0057]
$\hat{\sigma}^2_{IMLE}$	0.4739 (0.2185) [0.1158]	0.5127 (0.2573) [0.2747]	0.4748 (0.2194) [0.1925]	0.5717 (0.3163) [0.2226]	0.4060 (0.1506) [0.0524]

Table 3.2: Outcome of simulation with $T = 30$: average value of estimates based on 100 replications, with bias in ( ) and variance in [ ]

n	100	200	300	400	500
$\hat{H}_{CMLE}$	0.5392 (−0.0098) [0.0017]	0.5374 (−0.0116) [0.0010]	0.5448 (−0.0042) [0.0012]	0.5425 (−0.0065) [0.0009]	0.5454 (−0.0036) [0.0008]
$\hat{\mu}_{CMLE}$	0.2398 (−0.0355) [0.0194]	0.2788 (0.0035) [0.0172]	0.2475 (−0.0278) [0.0152]	0.2691 (−0.0062) [0.0163]	0.2841 (0.0088) [0.0184]
$\hat{\sigma}^2_{CMLE}$	0.2457 (−0.0097) [0.0019]	0.2457 (−0.0097) [0.0018]	0.2520 (−0.0034) [0.0024]	0.2527 (−0.0027) [0.0019]	0.2538 (−0.0016) [0.0021]
$\hat{H}_{RS}$	0.6165 (0.0675) [0.0291]	0.6302 (0.0812) [0.0165]	0.6175 (0.0685) [0.0104]	0.5950 (0.0460) [0.0113]	0.6155 (0.0665) [0.0073]
$\hat{\sigma}^2_{IMLE}$	0.3230 (0.0676) [0.0232]	0.3902 (0.1348) [0.0389]	0.3937 (0.1383) [0.0573]	0.3806 (0.1252) [0.0481]	0.4177 (0.1623) [0.0498]

Table 3.3: Outcome of simulation with $T = 40$: average value of estimates based on 100 replications, with bias in ( ) and variance in [ ]

n	100	200	300	400	500
$\hat{H}_{MLE}$	0.539 (−0.0100) [0.0019]	0.5363 (−0.0127) [0.0011]	0.5409 (−0.0081) [0.0012]	0.5455 (−0.0035) [0.0007]	0.5433 (−0.0057) [0.0007]
$\hat{\mu}_{CMLE}$	0.2853 (0.0100) [0.0869]	0.2662 (−0.0091) [0.0130]	0.2841 (0.0089) [0.0153]	0.3083 (0.0330) [0.0155]	0.2799 (0.0046) [0.0133]
$\hat{\sigma}^2_{CMLE}$	0.2504 (−0.0050) [0.0021]	0.2512 (−0.0042) [0.0016]	0.2477 (−0.0077) [0.0016]	0.2510 (−0.0044) [0.0015]	0.2487 (−0.0067) [0.0016]
$\hat{H}_{RS}$	0.6113 (0.0623) [0.0289]	0.6258 (0.0768) [0.0226]	0.6275 (0.0785) [0.0139]	0.6282 (0.0792) [0.0091]	0.6095 (0.0605) [0.0095]
$\hat{\sigma}^2_{IMLE}$	0.3011 (0.0457) [0.0115]	0.3752 (0.1198) [0.0409]	0.3883 (0.1328) [0.0394]	0.4052 (0.1498) [0.0420]	0.3945 (0.1390) [0.0628]

Table 3.4: Outcome of simulation with $T = 50$: average value of estimates based on 100 replications, with bias in ( ) and variance in [ ]

n	100	200	300	400	500
$\hat{H}_{MLE}$	0.541 (−0.0080) [0.0018]	0.54 (−0.0090) [0.0014]	0.5423 (−0.0067) [0.0012]	0.5428 (−0.0062) [0.0008]	0.5438 (−0.0052) [0.0008]
$\hat{\mu}_{CMLE}$	0.2907 (0.0154) [0.0714]	0.2675 (−0.0078) [0.0150]	0.2867 (0.0114) [0.0141]	0.2690 (−0.0062) [0.0101]	0.2644 (−0.0109) [0.0113]
$\hat{\sigma}^2_{CMLE}$	0.2469 (−0.0085) [0.0013]	0.2478 (−0.0076) [0.0014]	0.2508 (−0.0046) [0.0015]	0.2520 (−0.0034) [0.0012]	0.2526 (−0.0028) [0.0018]
$\hat{H}_{RS}$	0.6249 (0.0759) [0.0327]	0.6167 (0.0677) [0.0127]	0.6134 (0.0644) [0.0100]	0.6212 (0.0722) [0.0095]	0.6008 (0.0518) [0.0101]
$\hat{\sigma}^2_{IMLE}$	0.2881 (0.0327) [0.0072]	0.3210 (0.0656) [0.0116]	0.3417 (0.0863) [0.0148]	0.3784 (0.1230) [0.0278]	0.3614 (0.1060) [0.0283]

It is obvious from the results obtained in Tables 3.1–3.4 that our methodology performs considerably better. Most of the biases and variances obtained by using our method are within an acceptable tolerance. All of our estimates for H are obviously quite stable and less biased. The performance on the estimation of σ^2 is also fairly satisfactory. We can also see that the larger the sample size n, the better the estimation performs. Further, overall, with a larger T, the outcome become better for any n, which is consistent with our asymptotic result as shown in (3.35).

We find that the bias by the incomplete likelihood method of Kukush *et al.* is quite large in comparison to ours. We are able to give estimates not only for σ^2, but also for μ and the Hurst index, H, as well. The simulation outcomes indicate that our method is more promising in obtaining statistically efficient estimators for GFBM.

3.7 Empirical Results

3.7.1 Data

Kuala Lumpur Composite Index (KLCI) is introduced in 1986, as a guideline of the actual performance indicator, i.e., the local stock market barometer, for the overall Malaysia stock market as well as the economy in general. It comprises of 100 listed multi-sector companies from the Main Board in Bursa Malaysia, previously known as Kuala Lumpur Stock Exchange (KLSE).

KLCI is used as the main index in Malaysia, and is now one of the three primary indices for Malaysia stock market alongside with FTSE Bursa Malaysia Large 30 Index (FMB30) and FTSE Bursa Malaysia Emas (FMBEMAS). We are particularly interested in the index since this is a composite of different stock prices. Our aim is to achieve a better understanding of the overall performance of the state of economy in Malaysia. The fluctuations observed in price series are normally depending on many factors, such as political events, global economic performance as well as speculations by investors.

We used a data set from KLCI available online at *http : //www.econstats.com*. The daily close price data set of KLCI from 3 January, 2005 to 29 December, 2006 is examined, with 494 observations. The return series is then calculated in logarithm. The return is considered to prevent the high volatility in the data. The changes in the price seem to be more practical as these changes are stationary. The figures of the price and return series are presented in Figures 3.3 and 3.4. A summary of the return series can be found in Table 3.5, where the mean of this series is 0.0003915 and the variance is 0.00002584.

Table 3.5: Summary of the return series of KLCI

Min.	1st Qu.	Median	Mean	Var	3rd Qu.	Max.
-0.0202000	-0.0023850	0.0005159	0.0003915	0.00002584	0.0029860	0.0190700

3.7.2 Estimation based on CMLE method

We present in this subsection the results of our study of modeling the data of KLCI by GFBM. We try to estimate the parameters of the risky asset model by using the proposed complete maximum likelihood estimation (CMLE) based on daily return series. The estimates are summarized in Ta-

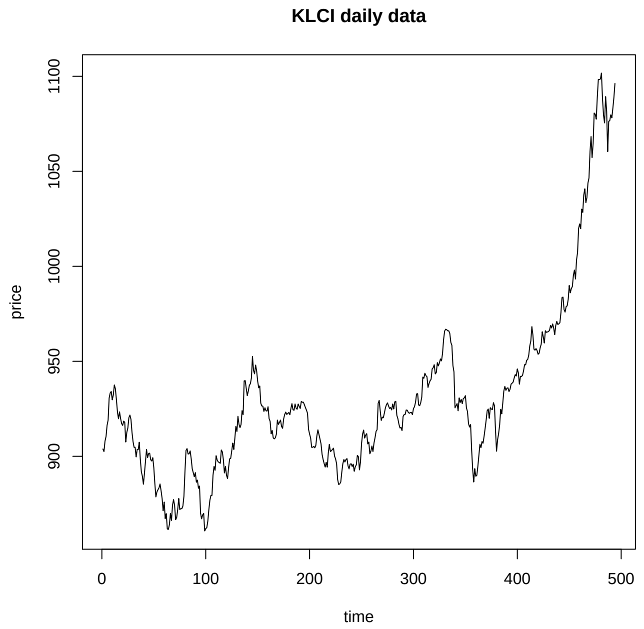

Figure 3.3: Daily close price series of KLCI from 3rd January 2005 to 29 December 2006

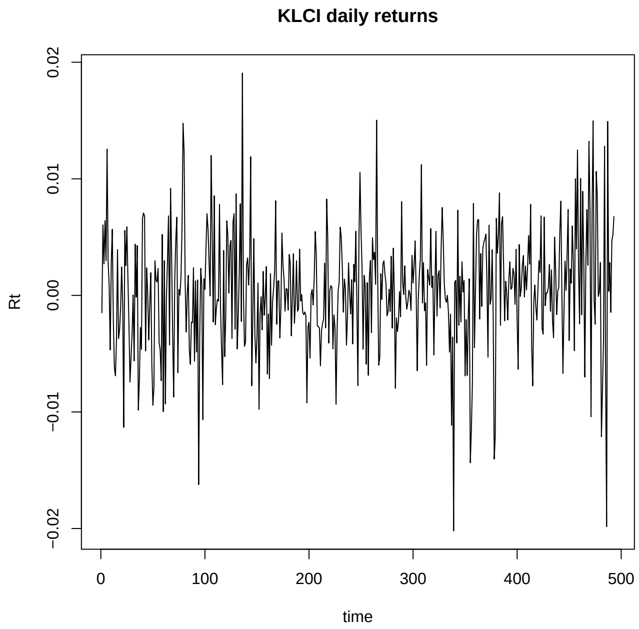

Figure 3.4: Daily return series of KLCI from 3rd January 2005 to 29 December 2006

ble 3.6.

We can clearly see from Table 3.6 that the suggested estimates are $H = 0.575$, $\sigma^2 = 0.00002576$ and $\mu = 0.0004510$. This finding agrees with the work by Sadique and Silvapulle in [124], where the presence of a weak long memory in Malaysia financial data is suggested.

Table 3.6: Likelihood value with respects to the model parameters

H value	$\widehat{\sigma^2}$	$\widehat{\mu}$	**likelihood value**
0.500	0.00002573	0.0004035	2357.961
0.570	0.00002571	0.0004470	2361.453
0.573	0.00002574	0.0004494	2361.464
0.574	0.00002575	0.0004502	2361.465
0.575	**0.00002576**	**0.0004510**	**2361.465**
0.576	0.00002577	0.0004518	2361.464
0.600	0.00002613	0.0004740	2361.115

3.7.3 Application to European Option Pricing

To calculate the appropriate value of European call option, we consider several maturity times (days) for an already traded option. The risk-free interest rate is fixed at 3.5% per annum in accordance with the actual Malaysia conventional interest rate on 29 December, 2006, and we are interested in the daily interest rate in this work. We select the underlying price at time as MYR1096.24, following the price on 29 December, 2006. The volatility and Hurst exponent are estimated based on our method for the historical daily return data of KLCI, with estimates listed in Table 3.6. For comparison, we also calculated the value of European call option using the estimates based on the method of Kukush *et al.* [78], the R/S analysis is used to obtain an estimator of H in advance, as well as the traditional Black-Scholes European option price. The outcomes are listed in Table 3.7.

From Table 3.7, we see that all cases exhibit somewhat differently in their call prices. Call prices valued by the traditional Black-Scholes provide

Table 3.7: Comparison of European call option prices using different methods: C_{CMLE} (this work), C_{IMLE} (Kukush *et al.* with R/S analysis) and C_{BS} (traditional Black Scholes)

$T_0 - t_0$	**K**	C_{CMLE} ($H = 0.575$) [$\sigma^2 = 0.00002576$]	C_{IMLE} ($H = 0.6551$) [$\sigma^2 = 0.00002590$]	C_{BS} ($H = 0.5$) [$\sigma^2 = 0.00002589$]
	1070	30.8566	35.2810	28.7439
	1080	23.2219	28.4503	20.2328
15	1090	16.6880	22.4382	13.0493
	1100	11.3930	17.2847	7.5809
	1110	7.3561	12.9897	3.9079
	1070	35.9385	43.3136	31.9585
	1080	28.9344	36.9983	24.1350
30	1090	22.7585	31.2702	17.3923
	1100	17.4615	26.1410	11.8932
	1110	13.0511	21.6084	7.6796
	1070	38.9955	47.9854	34.0120
	1080	32.2057	41.8453	26.4335
40	1090	26.1415	36.2119	19.8239
	1100	20.8361	31.0917	14.2966
	1110	16.2947	26.4823	9.8847
	1070	41.8534	52.3096	35.9756
	1080	35.2107	46.2922	28.5660
50	1090	29.2202	40.7253	22.0404
	1100	23.9057	35.6126	16.4850
	1110	19.2709	30.9517	11.9273

us with the least values, where the long memory is not taken into account. Method proposed in this work prices the call in an intermediate value between those obtained by the traditional Black-Scholes and the method by Kukush *et al.* with the R/S analysis. Call prices valued by Kukush *et al.* with the R/S analysis are the highest. Our method is based on rigorous theoretical reasoning (see results in the previous sections). It provides practically acceptable results, where the long memory is taken into account. It is seen that the longer the time to expiry, the higher the value of call price becomes. In the case of "in the money", the call price reveals a higher value when compared with the case of "out of the money", as expected.

3.8 Discussion

In this chapter, we studied the problem of estimating the unknown parameters in the GFBM. We proposed a complete MLE approach, which enables us not only to derive the estimators of μ and σ^2, but also the estimate of long memory parameter, H, simultaneously, for risky asset in the fractional Black-Scholes model. This approach helps us to value option pricing by taking into account long memory properties in a market governed model.

Based on the simulation study, we can conclude that our methodology is statistically efficient and reliable. We also carried out empirical study of the stock exchange index with option pricing under fractional Black-Scholes model. Based on the encouraging findings in this chapter, we will further investigate another important model in finance, i.e., the fractional Ornstein-Uhlenbeck model, in the next chapter.

Chapter 4

Fractional Ornstein-Uhlenbeck models: ML estimation with dynamic measurement errors and application to interest rate modeling

4.1 Introduction

Ornstein-Uhlenbeck (OU) model is important in finance. It is the continuous-time analogue of the famous autoregressive moving average (ARMA) process in discrete time. It is developed by Uhlenbeck and Ornstein in [139]. This model is famous for its mean-reverting property. It allows the drift parameter to take a negative value. The drift parameter is positive if the current value is less than the long term mean, and is negative otherwise. This parameter refers to the rate where the shocks in the system dissipate and the variable reverts towards the mean. In this model, the mean always acts as an equilibrium level for the process.

There are vast literature on OU and OU-extension-type models in finance. It is often applied to model interest rate, exchange rates, and stochastic volatility ([16], [28], [38], [39], [12] and [118]). Barndorff-Nielson and Shephard [12] have generalized the OU model to a non-Gaussian process. Recent work by Valdivieso *et al.* [140] has proposed a new method to estimate parameters in OU model by using the maximum likelihood estimator (MLE). Some other interesting works include [104], [1] and [82].

In the current chapter we study an extension of OU model where Brownian motion (BM) is replaced by fractional Brownian motion (FBM). Some recent works on this problem include those reported in [30], [75] and [56]. These results are mainly of theoretical nature. An exemption is the work by Hult [62]. He estimates the parameters involved in fractional Ornstein-Uhlenbeck (FOU) model based on discrete observation. He manipulates the spectral representation of the likelihood function. Hu [59] also studies this process in a financial environment. However, this exploratory study is also concerned only on the theoretical investigation.

The aim of this chapter is to estimate the parameters involved in FOU model. Unlike Hult [62], we use the innovation algorithm to derive a maximum likelihood estimator. Then, a simulated annealing method is applied to find the optimal parameters in FOU model simultaneously. The simulation results are carried out so as to illustrate the performance of our method. This chapter ends with a discussion section.

4.2 Fractional Ornstein-Uhlenbeck process

We begin this section by introducing the definition of Ornstein-Uhlenbeck (OU) model quoted from [136].

Definition 4.2.1. *The OU model is the solution to the linear stochastic differential equation (SDE)*

$$dX(t) = -\theta X(t)dt + \sigma dB(t), \qquad X(0) = U \tag{4.1}$$

where $B(t)$ is a Brownian motion (BM), $\theta \in \mathbb{R}$, and $\sigma > 0$. The initial value U is independent of the BM. The process X can be expressed explicitly *in terms of the $B(t)$ as:*

$$X(t) = e^{-\theta t}U + \sigma \int_0^t e^{-\theta(t-s)} dB(s). \tag{4.2}$$

Note that the OU model with $\mu \in \mathbb{R}$ is given by $Y(t) = X(t) + \mu$, where $Y(t)$ is a solution of the SDE given below.

$$dY(t) = -\theta(Y(t) - \mu)dt + \sigma dB(t). \tag{4.3}$$

In mathematical finance, this model is also known as the Vasiček model for the short-term interest rate. In the following section, we take $\mu = 0$.

In this chapter, we replace BM with FBM. The definition of fractional Ornstein-Uhlenbeck (FOU) model [62] is as follows.

Definition 4.2.2. *The FOU model is defined as the stationary solution to*

$$dX(t) = -\theta X(t)dt + \sigma dB_H(t), \tag{4.4}$$

where $\sigma > 0$, $\theta > 0$, $H \in (0, 1)$, and $\{B_H(t) : t \in \mathbb{R}\}$ is the FBM with the Hurst index H.

The increments of the FBM have long memory if $H > \frac{1}{2}$. This property is also enjoyed by the FOU process.

4.2.1 Model Simplification

We now consider the FOU model in the form of

$$\begin{aligned} dx(t) &= -a_1 x(t)dt + \sigma_1 dB_{H_1}(t) \\ x(0) &= x_0, \end{aligned} \tag{4.5}$$

where $a_1 > 0$ and $\sigma_1 \in \mathbb{R}$ are drift and diffusion parameters, respectively. $\{B_H(t), t \geq 0\}$ is a FBM on some probability space $(\Omega, \mathcal{F}, P)$ with the Hurst parameter, H.

The measurement dynamics is

$$\begin{aligned} dy(t) &= a_2 x(t)dt + \sigma_2 dB_{H_2}(t) \\ y(0) &= 0, \end{aligned} \tag{4.6}$$

where $a_2 \in \mathbb{R}$ and $\sigma_2 \in \mathbb{R}$ are drift and diffusion parameters, respectively. We simplify this problem by considering systems in discrete time. We have

$$x(t + \Delta t) - x(t) = -a_1 x(t)\Delta t + \sigma_1(B_{H_1}(t + \Delta t) - B_{H_1}(t)) \tag{4.7}$$

with the measurement dynamics

$$y(t + \Delta t) - y(t) = a_2 x(t)\Delta t + \sigma_2(B_{H_2}(t + \Delta t) - B_{H_2}(t)). \tag{4.8}$$

By setting $t = k\Delta t$, (4.7) and (4.8) can be written as

$$x((k + 1)\Delta t) - x(k\Delta t) = -a_1 x(k\Delta t)\Delta t + \sigma_1(B_{H_1}((k + 1)\Delta t) - B_{H_1}(k\Delta t)) \tag{4.9}$$

and

$$y((k + 1)\Delta t) - y(k\Delta t) = a_2 x(k\Delta t)\Delta t + \sigma_2(B_{H_2}((k + 1)\Delta t) - B_{H_2}(k\Delta t)). \tag{4.10}$$

Denoting $\tilde{X}_{k+1} = x((k+1)\Delta t)$, $\tilde{Y}_k = y((k+1)\Delta t) - y(k\Delta t)$, $\xi_{k+1} = B_{H_1}((k+1)\Delta t) - B_{H_1}(k\Delta t)$, and $\eta_k = B_{H_2}((k + 1)\Delta t) - B_{H_2}(k\Delta t)$, we obtain

$$\tilde{X}_{k+1} = (1 - a_1\Delta t)\tilde{X}_k + \sigma_1\xi_{k+1} \tag{4.11}$$

and

$$\tilde{Y}_k = a_2\Delta t\tilde{X}_k + \sigma_2\eta_k, \tag{4.12}$$

where $\eta_k \sim N(0, (\Delta t)^{2H_2})$ and $\xi_k \sim N(0, (\Delta t)^{2H_1})$. Note that (4.11) is an autoregressive (AR) process. Following the iteration process and the Cauchy criterion in [21], (4.11) can be re-stated as

$$\tilde{X}_{k+1} = \sum_{j=0}^{\infty}(1 - a_1\Delta t)^j \sigma_1\xi_{k+1-j}, \tag{4.13}$$

where $|1 - a_1\Delta t| < 1$.

The covariances of η and ξ can be expressed as:

$$\gamma_\eta(m) = cov(\eta_k, \eta_{k-m}) = \frac{1}{2}[|(m+1)\Delta t|^{2H_2} + |(m-1)\Delta t|^{2H_2} - 2|m\Delta t|^{2H_2}] \tag{4.14}$$

and

$$\gamma_\xi(m) = cov(\xi_k, \xi_{k-m}) = \frac{1}{2}[|(m+1)\Delta t|^{2H_1} + |(m-1)\Delta t|^{2H_1} - 2|m\Delta t|^{2H_1}], \tag{4.15}$$

respectively. When $H < \frac{1}{2}$, the increments (i.e., $\{B_H(t+\Delta t) - B_H(t)\}$) are negatively correlated. On the other hand, when $H > \frac{1}{2}$, the increments are positive correlated. This process is stationary, and it is known as a fractional Gaussian noise (FGN).

From (4.11), the covariance for $\tilde{X}$ can be expressed as:

$$\begin{aligned} \gamma_{\tilde{X}}(m) &= cov(\tilde{X}_k, \tilde{X}_{k-m}) \\ &= cov(\sum_{j=0}^{\infty}(1-a_1\Delta t)^j \sigma_1 \xi_{k-j}, \sum_{i=0}^{\infty}(1-a_1\Delta t)^i \sigma_1 \xi_{k-m-i}) \\ &= \sum_{j=0}^{\infty}\sum_{i=0}^{\infty}(1-a_1\Delta t)^{j+i}\sigma_1^2 cov(\xi_{k-j}, \xi_{k-m-i}) \\ &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(1-a_1\Delta t)^{i+j}\sigma_1^2 \gamma_\xi(m+i-j). \end{aligned} \tag{4.16}$$

By choosing a sufficiently large L, (4.16) can be written as:

$$\gamma_{\tilde{X}}(m) \approx \sum_{i=0}^{L}\sum_{j=0}^{L}(1-a_1\Delta t)^{i+j}\sigma_1^2 \gamma_\xi(m+i-j), \tag{4.17}$$

where $|a_1\Delta t| < 1$.

Denote $(\tilde{Y}_1, ..., \tilde{Y}_T) \sim N((\mu, ..., \mu), \Sigma)$, with $\Sigma = (\gamma_{\tilde{Y}}(i-j))^T_{\{i,j=1\}}$. The covariance for $\tilde{Y}$ can now be represented as:

$$\begin{aligned} \gamma_{\tilde{Y}}(m) &= cov(\tilde{Y}_k, \tilde{Y}_{k-m}) \\ &= (a_2\Delta t)^2 cov(\tilde{X}_k, \tilde{X}_{k-m}) + (\sigma_2)^2 cov(\eta_k, \eta_{k-m}) \\ &= (a_2\Delta t)^2 \gamma_{\tilde{X}}(m) + \sigma_2^2 \gamma_\eta(m). \end{aligned} \tag{4.18}$$

4.3 Methodology of Estimation

In this section, we estimate the parameters $a_1, a_2, \sigma_1, \sigma_2, H_1$ and H_2 involved in the dynamic and measurement systems given in the previous section.

The likelihood function for $(\tilde{Y}_1, ..., \tilde{Y}_T)$ is given by

$$L(a_1, a_2, \sigma_1, \sigma_2, H_1, H_2) = \frac{1}{(2\pi)^{T/2}(\det(\Sigma_T))^{1/2}} \exp\{-\frac{1}{2}(\tilde{Y}-\mu)'\Sigma_T^{-1}(\tilde{Y}-\mu)\}. \tag{4.19}$$

By maximizing the likelihood function, we obtain efficient estimators of all parameters. However, in (4.19), it is very difficult to analytically maximizing the log likelihood function, as the expression for the covariance function and its inversion are very complicated. As an alternative, the innovation algorithm will be implemented.

Suppose that $\{\tilde{Y}_1, ..., \tilde{Y}_T\}$ is stationary and that

$$\tilde{Y}_{n|n-1} = \phi_{n1}(\tilde{Y}_n - \mu) + \phi_{n2}(\tilde{Y}_{n-1} - \mu) + ... + \phi_{nn}(\tilde{Y}_1 - \mu) + \mu. \tag{4.20}$$

We add μ to the last term of (4.20) to ensure $E[\tilde{Y}_n - \tilde{Y}_{n|n-1}] = 0$.

We need definition for the best linear prediction for stationary processes of [128], which is quoted below.

Definition 4.3.1. *Given data $\tilde{Y}_1, ..., \tilde{Y}_n$, the best linear predictor, $\tilde{Y}_{n|n+m} = \sum_{k=1}^{n} \phi_{nk}\tilde{Y}_k$ of $\tilde{Y}_{n+m}$, for $m \geq 1$, is found by solving*

$$E[(\tilde{Y}_{n+m} - \tilde{Y}_{n|n+m})\tilde{Y}_k] = 0, \qquad k = 0, 1, ..., n$$

where $\tilde{Y}_0 = 1$.

By using Definition 4.3.1, the coefficients $\{\phi_{n1}, \phi_{n2}, ..., \phi_{nn}\}$ satisfy

$$E[(\tilde{Y}_{n+1} - \sum_{k=1}^{n} \phi_{nk}\tilde{Y}_{n+1-k})\tilde{Y}_{n+1-k}] = 0, \quad k = 1, ..., n,$$

and, equivalently,

$$\sum_{k=1}^{n} \phi_{nk}\gamma_{\tilde{Y}}(k-j) = \gamma_{\tilde{Y}}(k), \quad k = 1, ..., n. \tag{4.21}$$

(4.21) can be rewritten in matrix form as follows:

$$\mathbf{\Gamma}_n \boldsymbol{\phi}_n = \boldsymbol{\gamma}_n, \tag{4.22}$$

where $\mathbf{\Gamma}_n = \{\gamma_{\tilde{Y}}(k-j)\}_{j,k=1}^{n}$ is an $n \times n$ covariance matrix, $\boldsymbol{\phi}_n = (\phi_{n1}, ..., \phi_{nn})^T$ is an $n \times 1$ vector, and

$$\boldsymbol{\gamma}_n = (\gamma_{\tilde{Y}}(1), ..., \gamma_{\tilde{Y}}(n))^T \tag{4.23}$$

is an $n \times 1$ vector. Equation (4.20) and (equivalently, equation (4.22)) are called the one-step prediction equation.

By (4.22), we have

$$\boldsymbol{\phi}_n = \mathbf{\Gamma}_n^{-1}\boldsymbol{\gamma}_n. \tag{4.24}$$

Let $\tilde{\boldsymbol{\gamma}}_{n-1} = (\gamma_{\tilde{Y}}(n-1), ..., \gamma_{\tilde{Y}}(1))^T$. Then,

$$\mathbf{\Gamma}_n = \begin{bmatrix} \Gamma_{n-1} & \tilde{\gamma}_{n-1} \\ \tilde{\gamma}_{n-1}^T & \gamma_{\tilde{Y}}(0) \end{bmatrix}. \tag{4.25}$$

Thus,

$$\mathbf{\Gamma}_n^{-1} = \begin{bmatrix} I & -\Gamma_{n-1}^{-1}\tilde{\gamma}_{n-1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Gamma_{n-1}^{-1} & 0 \\ 0 & (\gamma_{\tilde{Y}}(0) - \tilde{\gamma}_{n-1}^T\Gamma_{n-1}^{-1}\tilde{\gamma}_{n-1})^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -\tilde{\gamma}_{n-1}^T\Gamma_{n-1}^{-1} & 1 \end{bmatrix}. \tag{4.26}$$

Let

$$\begin{aligned} \varepsilon_n &= \tilde{Y}_n - \tilde{Y}_{n|n-1} \\ &= \tilde{Y}_n - \mu - \sum_{k=1}^{n} \phi_{nk}(\tilde{Y}_k - \mu), \end{aligned} \tag{4.27}$$

where ε_n is a sequence of random variables, and these random variables are independent and identically distributed $\sim N(0, \upsilon_n^2)$, ϕ_{nk} and υ_n are the autoregressive and standard deviation parameters, respectively.

We can write (4.27) as:

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_T \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & ... & 0 \\ -\phi_{11} & 1 & 0 & ... & 0 \\ \vdots & ... & ... & ... & ... \\ -\phi_{(T-1)1} & -\phi_{(T-1)2} & ... & -\phi_{(T-1)(T-1)} & 1 \end{pmatrix} \begin{pmatrix} \tilde{Y}_1 - \mu \\ \tilde{Y}_2 - \mu \\ \vdots \\ \tilde{Y}_T - \mu \end{pmatrix} \tag{4.28}$$

Now, set $\mathbf{Y} = [\tilde{Y}_1 - \mu, ..., \tilde{Y}_T - \mu]^T$ and $\boldsymbol{\varepsilon} = [\varepsilon_1, ..., \varepsilon_T]$. Note that $\varepsilon_T \sim N(0, \upsilon_T^2)$, where υ_T^2 is the mean square prediction error, given by

$$\upsilon_T^2 = \gamma(0) - \boldsymbol{\gamma}_T\boldsymbol{\prime}\boldsymbol{\Gamma}_T^{-1}\boldsymbol{\gamma}_T. \tag{4.29}$$

Simplifying (4.28), we obtain

$$\boldsymbol{\varepsilon} = \mathbf{AY}. \tag{4.30}$$

Then, from (4.30), we have

$$\mathbf{Y} = \mathbf{A^{-1}}\boldsymbol{\varepsilon}. \tag{4.31}$$

Hence, the autocovariance function is

$$\begin{aligned} \boldsymbol{\Sigma}_T = cov(\mathbf{Y}, \mathbf{Y}) = E(\mathbf{YY^T}) &= \mathbf{A^{-1}}E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T](\mathbf{A^{-1}})^{\mathbf{T}} \\ &= \mathbf{A^{-1}} \begin{bmatrix} E(\varepsilon_1)^2 & 0 & ... & 0 \\ 0 & E(\varepsilon_2)^2 & ... & 0 \\ ... & & & \\ 0 & ... & & E(\varepsilon_T)^2 \end{bmatrix} (\mathbf{A^{-1}})^{\mathbf{T}}. \end{aligned} \tag{4.32}$$

The inversion of (4.32), $\boldsymbol{\Sigma}_T^{\mathbf{-1}}$, is given by

$$\boldsymbol{\Sigma}_T^{\mathbf{-1}} = \mathbf{A^T} \begin{bmatrix} \frac{1}{\upsilon_1^2} & 0 & ... & 0 \\ 0 & \frac{1}{\upsilon_2^2} & ... & 0 \\ ... & & & \\ 0 & ... & & \frac{1}{\upsilon_T^2} \end{bmatrix} \mathbf{A}, \tag{4.33}$$

and its determinant is

$$\det(\mathbf{\Sigma}_T) = \prod_{i=1}^{T} E(\varepsilon_i)^2 = \prod_{i=1}^{T} \upsilon_i^2. \tag{4.34}$$

By using (4.33) and (4.34), the likelihood function (4.19) can now be calculated.

Now, the likelihood function is transformed into the following optimization problem.

Problem P
Maximizes the cost function

$$L(\theta) \tag{4.35}$$

subject to

$$\begin{aligned} E(\tilde{Y} - \mu)^2 &\geq 0 \\ \upsilon^2 &\geq 0. \end{aligned} \tag{4.36}$$

where $\theta = (a_1, a_2, \sigma_1, \sigma_2, H_1, H_2)$.

This optimization problem is very difficult to solve. The constraints are too involved with covariance functions. To simplify this problem, we use the constraint transcription method reported in [68].

Maximizes the cost function:

$$L(\theta) \tag{4.37}$$

subject to

$$g_i(\theta) \leq 0, \quad i = 1, 2, \tag{4.38}$$

where g_i are the constraints in the original problem. Let this problem be referred to as Problem P. For each $i = 1, 2$, we approximate g_i with $G_{i,\varepsilon}(\theta)$, where

$$G_{i,\varepsilon}(\theta) = \begin{cases} g_i, & g_i > \varepsilon \\ \frac{(g_i+\varepsilon)^2}{4\varepsilon} & -\varepsilon < g_i < \varepsilon \\ 0 & g_i < -\varepsilon, \end{cases} \tag{4.39}$$

where ε some small number. We now append the approximate functions

$G_{i,\varepsilon}$ into the cost function $L(\theta)$ to an appended cost function given below.

Problem $P_{\varepsilon,\gamma}$

$$\hat{L}(\theta) = -L(\theta) - \gamma \sum_{j=1}^{m} G_{j,\varepsilon}(\theta), \tag{4.40}$$

where $\gamma > 0$ is a penalty parameter. This is an unconstraint optimization problem, which is referred to as Problem $P_{\varepsilon,\gamma}$. It is known (see [68]) that for any given $\varepsilon > 0$, there exists a $\gamma(\varepsilon)$ such that, for $\gamma > \gamma(\varepsilon)$, the solution of Problem $P_{\varepsilon,\gamma}$ will satisfy the constraints of Problem P. Let $\hat{\gamma}(\varepsilon)$ be such a γ for each $\varepsilon > 0$. Furthermore, the solution of Problem $P_{\varepsilon,\hat{\gamma}(\varepsilon)}$ converges to the solution of Problem P.

We propose an algorithm to solve Problem $P_{\varepsilon,\gamma}$.

4.3.1 Algorithm

Step 0. Set the initial value of parameters, $\{a_1, a_2, \sigma_1, \sigma_2, H_1, H_2\}$. Initialize $\gamma_\xi(0) = 0$, $\gamma_\eta = 0$, $\gamma_{\tilde{X}}(0) = 0$ and $\gamma_{\tilde{Y}} = 0$ and set L to be a large number.

Step 1. Calculate $\gamma_\eta(m)$, $\gamma_\xi(m)$, $\gamma_{\tilde{X}}(m)$ and $\gamma_{(\tilde{Y}-\mu)}(m)$, with $m = 1$ to $m = L$.

Step 2. Set k from 2 to T. Compute Γ_k^{-1} in (4.26).

Step 3. Compute ϕ_k in (4.24) and υ_k in (4.29).

Step 4. Calculate Σ_T^{-1} and $\det \Sigma_T$ in (4.33) and (4.34), respectively.

Step 5. Calculate the updated likelihood function, $\hat{L}_T(a_1^T, a_2^T, \sigma_1^T, \sigma_2^T, H_1^T, H_2^T)$. $\hat{L}_T$ that is minimum provides the efficient parameters, $\{a_1^T, a_2^T, \sigma_1^T, \sigma_2^T, H_1^T, H_2^T\}$ involved in this system.

4.4 Simulation Study

In order to examine the performance of the proposed estimators, we have carried out some simulation experiments. First, we generate the data from model (4.12). We take the parameters $a_1 = 2, a_2 = 2.5, \sigma_1 = 1, \sigma_2 = 0.5, H_1 = 0.7$ and $H_2 = 0.6$, while $\Delta t = \frac{1}{5}$. We simulate the time series from this discrete time model and apply our methodology to estimate the parameters $\vartheta = (a_1, a_2, \sigma_1, \sigma_2, H_1, H_2)\prime$ using the simulated data set. The simulated annealing method is used in order to find the optimal parameters, simultaneously. The simulation is repeated one hundred times. The simulated outcomes of the average value of estimates based on 100 replications, with bias and variances, are reported in Table 4.1. The 5 cases of sample sizes $n = 100, 200, 300, 400, 500$ are considered in the table.

Table 4.1: Average value of estimates based on 100 replications, with bias in ( ) and variance in [ ]

n	100	200	300	400	500
a_1	2.4156 (0.4156) [0.5958]	2.4233 (0.4233) [0.4429]	2.2133 (0.2133) [0.2144]	2.2085 (0.2085) [0.2118]	2.0884 (0.0884) [0.0736]
a_2	2.4829 (−0.0171) [0.1137]	2.5533 (0.0533) [0.0695]	2.5046 (0.0046) [0.0631]	2.4917 (−0.0083) [0.0520]	2.4814 (−0.0186) [0.0527]
σ_1	0.9898 (−0.0102) [0.0951]	1.0329 (0.0329) [0.0313]	0.9879 (−0.0122) [0.0069]	0.9981 (−0.0019) [0.0164]	0.9707 (−0.0293) [0.0059]
σ_2	0.5556 (0.0556) [0.0118]	0.5387 (0.0387) [0.0087]	0.5419 (0.0419) [0.0047]	0.5445 (0.0445) [0.0077]	0.5455 (0.0455) [0.0054]
H_1	0.7037 (0.0037) [0.0012]	0.6998 (−0.00025) [0.000345]	0.7046 (0.0046) [0.00046]	0.7026 (0.0026) [0.00065]	0.7062 (0.0062) [0.00036]
H_2	0.6084 (0.0084) [0.0016]	0.6067 (0.0067) [0.000625]	0.6048 (0.0048) [0.000416]	0.6042 (0.0042) [0.00058]	0.6024 (0.0024) [0.00031]

From the results obtained in Table 4.1, it shows that our methodology is efficient. Most of the biases and variance obtained are within an acceptable tolerance. All of our estimates for σ_1, σ_2, H_1 and H_2 are obviously quite stable and less biased. The performance on the estimation of a_1 and a_2 are fairly satisfactory. This simulation outcomes indicate that our methodology is promising in obtaining statistically efficient estimators for FOU.

4.5 Empirical Results

4.5.1 Data

We used a data set from federal reserve interest rate available online at *http://www.federalreserve.gov*. The interest rate of business day from 2 January 2009 to 31 December 2009 is examined, with 252 observations. The figure of the rate is presented in Figure 4.1. A summary of the time series can be found in Table 4.2, where the mean of this series is 0.1597 and the variance is 0.001502.

Table 4.2: Summary of the interest rate

Min.	1st Qu.	Median	Mean	Var	3rd Qu.	Max.
0.05	0.13	0.16	0.1597	0.0015	0.18	0.25

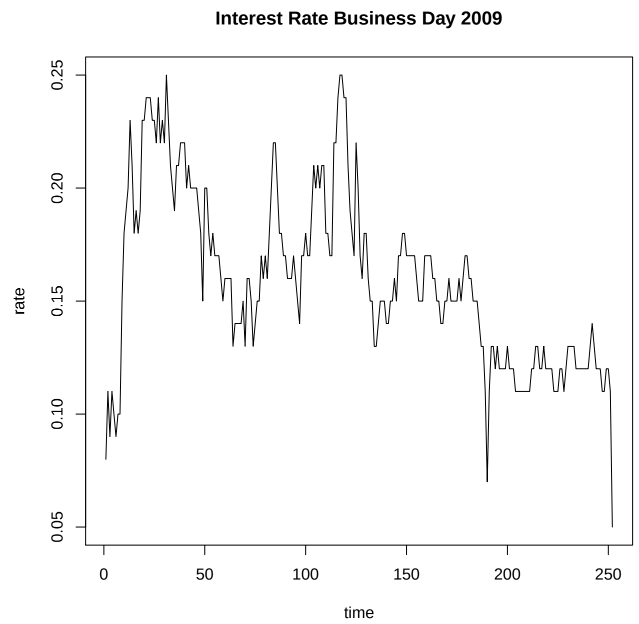

Figure 4.1: Daily (business day) interest rate from 2 January 2009 to 31 December 2009

4.5.2 Estimation

We present in this subsection the results of our study of modeling the data by FOU. Let us describe how the estimation procedure is conducted. First, we use many different cases of the initial values in the model estimation procedure. Each of these cases is computed by using the simulated annealing method. This method produces optimal parameters for each case of initial values considered. Then, we examine the likelihood values for these cases. The likelihood that is the maximum is chosen as the suggested estimates. In Table 4.3, we show some results from these cases, where the initial values are

- i : $a_1 = 2$, $a_2 = 2.5$, $\sigma_1 = 1$, $\sigma_2 = 0.5$, $H_1 = 0.7$ and $H_2 = 0.6$;
- ii : $a_1 = 3$, $a_2 = 5$, $\sigma_1 = 0.5$, $\sigma_2 = 0.7$, $H_1 = 0.8$ and $H_2 = 0.6$;
- iii : $a_1 = 0.1597$, $a_2 = 0.1597$, $\sigma_1 = 0.0387$, $\sigma_2 = 0.0387$, $H_1 = 0.8$ and $H_2 = 0.7$;
- iv : $a_1 = 1$, $a_2 = 2$, $\sigma_1 = 1$, $\sigma_2 = 0.7$, $H_1 = 0.7$ and $H_2 = 0.7$;
- v : $a_1 = 3$, $a_2 = 1$, $\sigma_1 = 0.6$, $\sigma_2 = 0.8$, $H_1 = 0.8$ and $H_2 = 0.8$;
- vi : $a_1 = 3$, $a_2 = 5$, $\sigma_1 = 0.5$, $\sigma_2 = 0.7$, $H_1 = 0.5$ and $H_2 = 0.5$;
- vii : $a_1 = 3$, $a_2 = 5$, $\sigma_1 = 0.5$, $\sigma_2 = 0.7$, $H_1 = 0.51$ and $H_2 = 0.51$;
- viii : $a_1 = 1$, $a_2 = 2$, $\sigma_1 = 1$, $\sigma_2 = 0.7$, $H_1 = 0.5$ and $H_2 = 0.5$;
- ix : $a_1 = 5$, $a_2 = 2$, $\sigma_1 = 1$, $\sigma_2 = 0.5$, $H_1 = 0.7$ and $H_2 = 0.7$;
- x : $a_1 = 5$, $a_2 = 3$, $\sigma_1 = 1$, $\sigma_2 = 0.7$, $H_1 = 0.6$ and $H_2 = 0.8$.

Here, we take $\Delta t = \frac{1}{5}$ which is similar to that of previous section. We can see from Table 4.3 that the suggested estimates are $a_1 = 2.277$, $a_2 = 4.237$, $\sigma_1 = 0.509$, $\sigma_2 = 0.713$, $H_1 = 0.809$ and $H_2 = 0.982$. Obviously the long-memory property is rather strong in interest rate process.

Table 4.3: Likelihood value of different initials

Initial	$\widehat{a_1}$	$\widehat{a_2}$	$\widehat{\sigma_1}$	$\widehat{\sigma_2}$	$\widehat{H_1}$	$\widehat{H_2}$	**likelihood value**
i	4.529	0.557	0.559	0.917	0.699	0.94	-20629.275
ii	1.563	16.11	0.157	0.636	0.773	0.89	27929.3925
iii	0.227	10.167	0.063	4.33	0.865	0.997	-41402.1124
iv	5.999	0.889	0.966	1.073	0.878	0.913	3340.3349
v	0.813	7.112	0.254	3.818	0.711	0.99	-63041.7107
vi	1.601	15.989	0.096	0.585	0.721	0.868	-94979.9325
vii	1.833	15.851	0.081	1.161	0.521	0.9	90338.7164
viii	2.786	1.088	2.681	0.042	0.921	0.593	-1189.4721
ix	**2.277**	**4.237**	**0.509**	**0.713**	**0.809**	**0.982**	**1900437.9156**
x	6.414	1.947	0.545	0.585	0.617	0.846	-45609.6661

4.6 Discussion

In this chapter, we proposed a novel method for estimating the unknown parameters in fractional Ornstein-Uhlenbeck (FOU) model. The likelihood function for FOU is difficult to solve analytically. Its covariance calculation is very expensive. Due to this, we proposed the innovation algorithm approach to simplify this problems. The likelihood function now transformed to an optimization problem with some constraints. The constraints transcription method is applied to append the approximate constraints to the cost function. Then, we solved a standard unconstrained optimization problem, by using simulated annealing method. We carried out some simulation study so as to illustrate the efficiency of our method. We also carried out empirical study to the interest rate data.

Based from these interesting findings on the model-based FBM in these chapters, we will further investigate another important problem in FBM, i.e., the filtering problem in multi-dimensional linear system, in the next chapter.

Chapter 5

Optimal filtering of linear system driven by fractional Brownian motion

5.1 Introduction

The study of filtering has been around for several decades. It goes back as early as 1960, when Kalman [70] dealt with a problem posed by Gauss on the estimation of the satellite orbits. Later in 1961, Kalman and Bucy [71] studied the filtering problem involving linear continuous-time processes. It has been used in various areas arise in physical sciences, engineering, economics and social sciences. Its aim is to extract the best information on the state process based on the measured data. For further details, see, for example, ([3], [6], [22], [54]).

However, most of the filtering problems in the literature are concerned with noises characterized by standard Brownian motion (BM). There are only few results on the filtering problems which are given by fractional Brownian motion (FBM) processes. In ([36], [74], [80]), only one dimensional differential equation is considered. In [4], the study is extended to a multi-dimensional case where both the state and observation are governed by respective linear stochastic differential equations which are driven by FBM processes. In this chapter, we will investigate this problem further, based on the fundamental results established in [4]. We are concerned with a continuous time filtering of a multi-dimensional linear system driven by FBM in control theory.

The filtering problem in the presence of FBM can be transformed (see [4]) into an equivalent deterministic optimal control problem, where its system dynamic is described by nonlinear ordinary differential equations involving convolutional integrals. It is very difficult to solve such an optimal control problem directly. The aim of this chapter is to develop a computational scheme for solving this problem. First, as in [4], this filtering problem is transformed into a deterministic optimal control problem, where its system dynamic is described by nonlinear differential equations involving convolutional integrals. Then, a novel approximation scheme, supported by rigorous mathematical analysis, is developed to solve this optimal control problem.

We first introduce some background knowledge on FBM in continuous time. For further details on FBM, see, for example, ([14], [15], [60], [107]), and for details on filtering problems, see ([3], [135]). Let $(\Omega, \mathcal{F}, P)$ be a probability space and $H \in (0,1)$. $\mathbf{B}$ is an $n-$dimensional Brownian motion with covariance matrix $\mathbf{Q} \in \mathbf{M}_s^+(n \times n)$, where $\mathbf{M}_s^+(n \times n)$ denotes the class of all $n \times n$ real symmetric positive definite matrices. Define

$$\mathbf{B}_H(t) = \int_0^t K_H(t,\theta) d\mathbf{B}(\theta), \tag{5.1}$$

with K_H being a kernel depending on the parameter H. Let it be chosen as:

$$K_H(t,s) = \frac{(t-s)^{H-\frac{1}{2}}}{\Gamma(H+\frac{1}{2})} F(\frac{1}{2}-H, H-\frac{1}{2}, H+\frac{1}{2}, 1-\frac{t}{s}) \mathbf{1}_{(0,t)}(s), \tag{5.2}$$

where Γ is a gamma function and F is a hypergeometric function. $\mathbf{B}_H$ is a R^n-valued Gaussian random process with mean and covariance matrix given by

(i) $E\{\mathbf{B}_H(t)\} = 0$;

(ii) $E\{(\mathbf{B}_H(t), \xi)(\mathbf{B}_H(s), \eta)\} = \int_0^t \int_0^s \varphi_H(\tau-\theta)(\mathbf{Q}\xi, \eta) d\tau d\theta$ for all $\xi, \eta \in \mathbb{R}^n$.

From (ii), it follows that

$$E\{(\mathbf{B}_H(t),\xi)^2\} = t^{2H}(\mathbf{Q}\xi,\xi), \xi \in \mathbb{R}^n, t \in \mathbb{R}_+, \tag{5.3}$$

where $\mathbb{R}_+ = \{t \in \mathbb{R} : t \geq 0\}$.

In this chapter, we will formulate the filtering problem which is driven by a FBM process. As in [3], this filtering problem is shown to be equivalent to a deterministic optimal control problem with convolutional integrals appeared in its system dynamics. Then, we construct a sequence of approximate optimal control problems, where their system dynamics are expressed by ordinary differential equations. Later, the convergence properties of the approximation scheme are established. Numerical simulation is presented to illustrate the efficiency of our method.

5.2 Overview on Linear Optimal Filtering

Filtering is very beneficial. It is one way to eliminate the presence of noises in some signals. A good filter can provide output that is the closest to the correct signal. The linear optimal filter normally associated with Gaussian random processes, and that its minimum mean square error is linear [5].
Consider a dynamical system,

$$d\mathbf{x}(t) = \mathbf{A}(t)\mathbf{x}(t)dt + \boldsymbol{\sigma_1}(t)d\mathbf{B}(t) \tag{5.4a}$$

$$\mathbf{x}(0) = \mathbf{x}_0, \tag{5.4b}$$

where $\{\mathbf{x}(t), t \geq 0\}$ the state process, $\mathbf{A}(t)$ locally integrable and $\boldsymbol{\sigma_1}(t)$ locally square integrable function taking values from matrices $M(n \times n)$ and $M(n \times p)$ respectively. $\{\mathbf{B}(t), t \geq 0\}$ is a Brownian motion (BM) with covariance matrix $\{Q(t) = tQ, t \geq 0\}$, for some $Q \in M_s^+(p \times p)$, given by

$$(Q\xi, \xi) = E(W(1), \xi)^2.$$

The measurement process is governed by

$$d\mathbf{y}(t) = \mathbf{H}(t)\mathbf{x}(t)dt + \boldsymbol{\sigma_2}(t)d\mathbf{B}(t) \tag{5.5a}$$

$$\mathbf{y}(0) = \mathbf{x}_0. \tag{5.5b}$$

The filter is now taking the form given below.

$$d\mathbf{z}(t) = \mathbf{G}(t)\mathbf{z}(t)dt + \mathbf{\Gamma}(t)d\mathbf{y}(t) \tag{5.6a}$$

$$\mathbf{z}_0 = \bar{\mathbf{x}}_0. \tag{5.6b}$$

The problem now is reduced to choosing the matrices $\mathbf{G}$ and $\mathbf{\Gamma}$ so that the estimate $\mathbf{z}$ for the state $\mathbf{x}$ is unbiased and has minimum variance. Readers are referred to [3] for details of this procedure. Motivated by this problem, we study this filtering problem when the Brownian noise is replaced by the fractional Brownian motion (FBM).

5.3 Linear filtering with FBM

Consider the following FBM dynamical system:

$$d\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t)dt + \mathbf{\Xi_1}d\mathbf{B}_{H_1}(t) \tag{5.7a}$$

$$\mathbf{x}(0) = \mathbf{x}_0, \tag{5.7b}$$

where $\mathbf{x}(t) \in \mathbb{R}^n$, and $\{\mathbf{A}, \mathbf{\Xi_1}\}$ are $n \times n$ and $n \times d$ constant matrices, respectively.
The measurement dynamics is given by

$$d\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t)dt + \mathbf{\Xi_2}d\mathbf{B}_{H_2}(t) \tag{5.8a}$$

$$\mathbf{y}(0) = \mathbf{0}, \tag{5.8b}$$

where $\mathbf{y}(t) \in \mathbb{R}^m$, and $\{\mathbf{H}, \mathbf{\Xi_2}\}$ are $m \times n$ and $m \times m$ constant matrices, respectively. $\{\mathbf{B}_{H_1}(t), t \geq 0\}$ and $\{\mathbf{B}_{H_2}(t), t \geq 0\}$ are FBM processes taking values in $\mathbb{R}^d$ and $\mathbb{R}^m$, respectively.

Let $\{\mathcal{F}_t^{\mathbf{y}}, t \geq 0\}$ be an increasing family of subsigma algebras of the sigma algebra $\mathcal{F}$ induced by the random process $\{\mathbf{y}(t), t \geq 0\}$. From [4], the basic filtering problem is to find a process $\mathbf{z}(t)$ so that for each $t \geq 0$, $\mathbf{z}(t)$ is

$\mathcal{F}_t^{\mathbf{y}}$−adapted satisfying

$$(1)E\{\mathbf{z}(t)\} = E\{\mathbf{x}(t)\}, \quad t \geq 0; \text{ and} \tag{5.9a}$$

$$(2)E\{\| \mathbf{x}(t) - \mathbf{z}(t) \|^2\} \text{ is minimum for } t \geq 0, \tag{5.9b}$$

where $\| \cdot \|$ denotes the usual Euclidean norm. That is, for a vector $\mathbf{v} \in \mathbb{R}^n$,

$$\|\mathbf{v}\| = (\sum_{i=1}^{n} (v_i)^2)^{\frac{1}{2}}. \tag{5.10}$$

Furthermore, for a matrix $\mathbf{A} \in \mathbb{R}^{n \times m}$, we define

$$\|\mathbf{A}\| = (\sum_{i=1}^{n} \sum_{j=1}^{m} (A_{ij})^2)^{\frac{1}{2}} \tag{5.11}$$

and

$$\|\mathbf{A}\|_\infty = \max_{1 \leq i,j \leq n} |A_{i,j}|. \tag{5.12}$$

Such a $\mathbf{z}$ is known as the best unbiased-minimum variance (UMV) linear filter driven by the observation process $\mathbf{y}$. It is expressed in the form of the following stochastic differential equations

$$d\mathbf{z}(t) = \mathbf{G}_{\mathbf{\Gamma}}\mathbf{z}(t)dt + \mathbf{\Gamma} d\mathbf{y}(t), \quad t \geq 0, \tag{5.13a}$$

$$\mathbf{z}(0) = \hat{\mathbf{x}}_0 \equiv E\mathbf{x}_0, \tag{5.13b}$$

where $\mathbf{G}_{\mathbf{\Gamma}}$ and $\mathbf{\Gamma} \in \mathcal{D}$ are constant matrices with appropriate dimensions, which are to be determined, and

$$\mathcal{D} = \{\mathbf{\Gamma} : \|\mathbf{\Gamma}\|_\infty = \max_{1 \leq i,j \leq n} |\Gamma_{i,j}| \leq M\}, \tag{5.14}$$

with $M > 0$ being a fixed constant.

Let

$$\mathbf{e}(t) = \mathbf{x}(t) - \mathbf{z}(t), \qquad t \geq 0. \tag{5.15}$$

Then,

$$\begin{aligned} d\mathbf{e} &= d\mathbf{x} - d\mathbf{z} \\ &= (\mathbf{A} - \mathbf{\Gamma H})\mathbf{e}dt + (\mathbf{A} - \mathbf{\Gamma H} - \mathbf{G_\Gamma})\mathbf{z}(t)dt \\ &+ \mathbf{\Xi_1} d\mathbf{B}_{H_1}(t) - \mathbf{\Gamma \Xi_2} d\mathbf{B}_{H_2}(t) \end{aligned} \tag{5.16a}$$

$$\mathbf{e}(0) = \mathbf{e}_0 \equiv \mathbf{x}_0 - \hat{\mathbf{x}}_0. \tag{5.16b}$$

Choose $\mathbf{G_\Gamma} = \mathbf{A} - \mathbf{\Gamma H}$. Then, it follows from (5.13) that

$$d\mathbf{z}(t) = (\mathbf{A} - \mathbf{\Gamma H})\mathbf{z}(t)dt + \mathbf{\Gamma} d\mathbf{y}(t),\ t \geq 0, \tag{5.17a}$$

$$\mathbf{z}(0) = \hat{\mathbf{x}}_0. \tag{5.17b}$$

The error equation (5.16a) with initial condition (5.16b) is reduced to

$$d\mathbf{e} = (\mathbf{A} - \mathbf{\Gamma H})\mathbf{e}(t)dt + \mathbf{\Xi_1} d\mathbf{B}_{H_1}(t) - \mathbf{\Gamma \Xi_2} d\mathbf{B}_{H_2}(t), \tag{5.18a}$$

$$\mathbf{e}(0) = \mathbf{e}_0. \tag{5.18b}$$

Let $\{\mathbf{\Phi}(t,s),\ 0 \leq s \leq t < \infty\}$ denote the transition operator corresponding to $\mathbf{G_\Gamma} = \mathbf{A} - \mathbf{\Gamma H}$. With this operator, the solution of (5.18) can be written as:

$$\mathbf{e}(t) = \mathbf{\Phi}(t,0)\mathbf{e}_0 + \int_0^t \mathbf{\Phi}(t,\theta)\mathbf{\Xi_1} d\mathbf{B}_{H_1}(\theta) - \int_0^t \mathbf{\Phi}(t,\theta)\mathbf{\Gamma \Xi_2} d\mathbf{B}_{H_2}(\theta). \tag{5.19}$$

The transition operator $\mathbf{\Phi}$ is governed by

$$(\frac{\partial}{\partial t})\mathbf{\Phi}(t,s) = \mathbf{G_\Gamma}\mathbf{\Phi}(t,s) \tag{5.20a}$$

and

$$\mathbf{\Phi}(t,t) = \mathbf{I}. \tag{5.20b}$$

Since $\mathbf{G_\Gamma}$ is a constant matrix, we have

$$\mathbf{\Phi}(t,s) = e^{\mathbf{G_\Gamma}(t-s)} = e^{\mathbf{G_\Gamma} t} e^{-\mathbf{G_\Gamma} s}. \tag{5.21}$$

We need Lemma 3.1 of [3], which is quoted below.

Lemma 1. *For each* $\mathbf{\Gamma} \in \mathcal{D}$*, the error covariance matrix* $\mathbf{K}$ *is the solution* *of the system described by the following differential equations.*

$$\begin{aligned}\dot{\mathbf{K}}(t) = \mathbf{G}_{\mathbf{\Gamma}}\mathbf{K}(t) + \mathbf{K}\mathbf{G}_{\mathbf{\Gamma}}^T + e^{\mathbf{G}_{\mathbf{\Gamma}}t}\{\int_0^t \varphi_{H_1}(t-s)e^{-\mathbf{G}_{\mathbf{\Gamma}}s}\mathbf{\Xi}_{\mathbf{1}}ds\}\mathbf{Q}\mathbf{\Xi}_{\mathbf{1}}^T \\ + e^{\mathbf{G}_{\mathbf{\Gamma}}t}\{\int_0^t \varphi_{H_1}(t-s)e^{-\mathbf{G}_{\mathbf{\Gamma}}s}\mathbf{\Xi}_{\mathbf{1}}^T ds\}\mathbf{Q}^T\mathbf{\Xi}_{\mathbf{1}} \\ + e^{\mathbf{G}_{\mathbf{\Gamma}}t}\{\int_0^t \varphi_{H_2}(t-s)e^{-\mathbf{G}_{\mathbf{\Gamma}}s}\mathbf{\Gamma}\mathbf{\Xi}_2 ds\}\mathbf{Q}_0\mathbf{\Xi}_2^T\mathbf{\Gamma}^T \\ + e^{\mathbf{G}_{\mathbf{\Gamma}}t}\{\int_0^t \varphi_{H_2}(t-s)e^{-\mathbf{G}_{\mathbf{\Gamma}}s}\mathbf{\Gamma}^T\mathbf{\Xi}_2^T ds\}\mathbf{Q}_0^T\mathbf{\Xi}_2\mathbf{\Gamma}\end{aligned} \tag{5.22a}$$

with initial condition

$$\mathbf{K}(0) = \mathbf{K}_0, \tag{5.22b}$$

where $\mathbf{Q} \in \mathbf{M}_s^+(d \times d)$ *and* $\mathbf{Q}_0 \in \mathbf{M}_s^+(m \times m)$ *are covariance matrices of* $\mathbf{B}_{H_1}(t)$ *and* $\mathbf{B}_{H_2}(t)$*, respectively, and* $\mathbf{G}_{\mathbf{\Gamma}} = \mathbf{A} - \mathbf{\Gamma}\mathbf{H}$.

Now, the filtering problem is transformed into the following equivalent optimal control problem.

Problem (P). Given the dynamic system

$$\begin{aligned}\dot{\mathbf{K}}(t) = \mathbf{G}_{\mathbf{\Gamma}}\mathbf{K}(t) + \mathbf{K}\mathbf{G}_{\mathbf{\Gamma}}^T + e^{\mathbf{G}_{\mathbf{\Gamma}}t}\{\int_0^t \varphi_{H_1}(t-s)e^{-\mathbf{G}_{\mathbf{\Gamma}}s}\mathbf{\Xi}_{\mathbf{1}}ds\}\mathbf{Q}\mathbf{\Xi}_{\mathbf{1}}^T \\ + e^{\mathbf{G}_{\mathbf{\Gamma}}t}\{\int_0^t \varphi_{H_1}(t-s)e^{-\mathbf{G}_{\mathbf{\Gamma}}s}\mathbf{\Xi}_{\mathbf{1}}^T ds\}\mathbf{Q}^T\mathbf{\Xi}_{\mathbf{1}} \\ + e^{\mathbf{G}_{\mathbf{\Gamma}}t}\{\int_0^t \varphi_{H_2}(t-s)e^{-\mathbf{G}_{\mathbf{\Gamma}}s}\mathbf{\Gamma}\mathbf{\Xi}_{\mathbf{2}} ds\}\mathbf{Q}_0\mathbf{\Xi}_{\mathbf{2}}^T\mathbf{\Gamma}^T \\ + e^{\mathbf{G}_{\mathbf{\Gamma}}t}\{\int_0^t \varphi_{H_2}(t-s)e^{-\mathbf{G}_{\mathbf{\Gamma}}s}\mathbf{\Gamma}^T\mathbf{\Xi}_{\mathbf{2}}^T ds\}\mathbf{Q}_0^T\mathbf{\Xi}_{\mathbf{2}}\mathbf{\Gamma},\end{aligned} \tag{5.23a}$$

$$\mathbf{K}(0) = \mathbf{K}_0, \tag{5.23b}$$

find a $\mathbf{\Gamma} \in \mathcal{D}$ such that the cost function

$$J(\mathbf{\Gamma}) = \int_0^T \mathrm{trace}\{\mathbf{\Sigma}(t)\mathbf{K}(t)\}dt \tag{5.24}$$

is minimized, where $\mathbf{\Sigma}(t)$ is a weighting matrix-valued function, which is positive definite and symmetric for each $t \in [0, T]$, and T is the terminal time.

This optimal control problem is very difficult to solve, because of the appearance of convolutional integrals in its system dynamics. In the next section, an approximation scheme is developed, which will then be used to approximate this optimal control problem into a sequence of optimal control problems involving only ordinary differential equations. Each of these standard approximate optimal control problems can be solved by the control parameterization technique used in conjunction with the time scaling transform reported in ([86], [144]). In particular, the optimal control software package, *MISER 3.3* [67], can be used for this purpose.

5.4 Approximation Method

In this section, we propose an approximation scheme to construct a sequence of approximate problems (*P(N)*) which are governed by ordinary differential equations. First, the kernels $\varphi_{H_1}(t-s)$ and $\varphi_{H_2}(t-s)$ in (5.22) are approximated by using an expansion of Chebyshev series. By doing this, the convolutional integrals are approximated by respective ordinary differential equations. In this way, Problem (P) is approximated by a sequence of optimal control problems involving only ordinary differential equations. Each of which can be solved by using many efficient numerical methods available in the literature.

We approximate the kernels $\varphi_{H_1}(t-s)$ and $\varphi_{H_2}(t-s)$ by a finite expansion of Chebyshev series as follows:

$$\varphi_{H_1}(t) \approx \varphi_{H_1}^N(t) = \sum_{i=1}^{N} \alpha_i^N k_i(t) \tag{5.25}$$

and

$$\varphi_{H_2}(t) \approx \varphi_{H_2}^N(t) = \sum_{i=1}^{N} \beta_i^N k_i(t) \tag{5.26}$$

where

$$k_i(t) = \bar{T}_{i-1}(t) = T_{i-1}(\frac{2t}{T}-1) = \cos[(i-1)\cos^{-1}(\frac{2t}{T}-1)],\ 0 \le t \le T, \tag{5.27}$$

are basis functions obtained from the shifted Chebyshev series. They satisfy the system of ordinary differential equations with constant coefficients,

$$\dot{k}_i(t) = \sum_{j=1}^{N} a_{ij} k_j(t), \tag{5.28}$$

with initial condition

$$k_i(0) = \bar{T}_{i-1}(0) = T_{i-1}(-1) = (-1)^{i-1}, \tag{5.29}$$

where a_{ij}, α_i^N and β_i^N, $i = 1, \cdots, N$, $j = 1, \cdots, N$, are given, respectively, by

$$a_{ij} = \begin{cases} 0, & j \geq i \\ [2(i-1)/T][1-(-1)^{i-j}], & j \neq 1 \\ [(i-1)/T][1+(-1)^i], & j = 1 \end{cases} \tag{5.30}$$

$$\alpha_i^N = \begin{cases} \frac{1}{N}\sum_{j=1}^N \varphi_{H_1}(t_j), & i = 1 \\ \frac{2}{N}\sum_{j=1}^N \bar{T}_{i-1}(t_j)\varphi_{H_1}(t_j), & i = 2, ..., N \end{cases} \tag{5.31}$$

and

$$\beta_i^N = \begin{cases} \frac{1}{N}\sum_{j=1}^N \varphi_{H_2}(t_j), & i = 1 \\ \frac{2}{N}\sum_{j=1}^N \bar{T}_{i-1}(t_j)\varphi_{H_2}(t_j), & i = 2, ..., N, \end{cases} \tag{5.32}$$

while

$$t_j = \frac{T}{2} + \frac{T}{2}\cos\left[\frac{(2j-1)\pi}{2N}\right], \ j = 1, ..., N. \tag{5.33}$$

With $\varphi_{H_1}(t)$ and $\varphi_{H_2}(t)$ being, respectively, approximated by $\varphi_{H_1}^N(t)$ and $\varphi_{H_2}^N(t)$ in (5.23), we obtain

$$\begin{aligned} \dot{\mathbf{K}}^N(t) = \mathbf{G_\Gamma K}^N(t) + \mathbf{K}^N\mathbf{G_\Gamma}^T + e^{\mathbf{G_\Gamma}t}\{\int_0^t \varphi_{H_1}^N(t-s)e^{-\mathbf{G_\Gamma}s}\mathbf{\Xi_1}ds\}\mathbf{Q\Xi_1}^T \\ + e^{\mathbf{G_\Gamma}t}\{\int_0^t \varphi_{H_1}^N(t-s)e^{-\mathbf{G_\Gamma}s}\mathbf{\Xi_1}^T ds\}\mathbf{Q}^T\mathbf{\Xi_1} \\ + e^{\mathbf{G_\Gamma}t}\{\int_0^t \varphi_{H_2}^N(t-s)e^{-\mathbf{G_\Gamma}s}\mathbf{\Gamma\Xi_2}ds\}\mathbf{Q}_0\mathbf{\Xi_2}^T\mathbf{\Gamma}^T \\ + e^{\mathbf{G_\Gamma}t}\{\int_0^t \varphi_{H_2}^N(t-s)e^{-\mathbf{G_\Gamma}s}\mathbf{\Gamma}^T\mathbf{\Xi_2}^T ds\}\mathbf{Q}_0^T\mathbf{\Xi_2\Gamma} \end{aligned} \tag{5.34}$$

Now, Problem (P) is approximated by a sequence of optimal control problems (P(N)) defined below.

Problem (P(N)). Given the dynamical system (5.34) with initial condition

$$\boldsymbol{K}^N(0) = \boldsymbol{K}_0, \tag{5.35}$$

find a $\mathbf{\Gamma} \in \mathcal{D}$ such that the cost function

$$J(\mathbf{\Gamma}) = \int_0^T \text{trace}\{\mathbf{\Sigma}(t)\mathbf{K}^N(t)\}dt \tag{5.36}$$

is minimized.

Define

$$\mathbf{w}_{1,i}(t) = \int_0^t k_i(t-s)e^{-\mathbf{G}_{\mathbf{\Gamma}}s}\mathbf{\Xi_1}ds, \qquad i = 1, \cdots, N, \tag{5.37}$$

$$\mathbf{w}_{2,i}(t) = \int_0^t k_i(t-s)e^{-\mathbf{G}_{\mathbf{\Gamma}}s}\mathbf{\Xi_1^T}ds, \qquad i = 1, \cdots, N, \tag{5.38}$$

$$\mathbf{w}_{3,i}(t) = \int_0^t k_i(t-s)e^{-\mathbf{G}_{\mathbf{\Gamma}}s}\mathbf{\Gamma}\mathbf{\Xi_2}ds, \qquad i = 1, \cdots, N, \tag{5.39}$$

and

$$\mathbf{w}_{4,i}(t) = \int_0^t k_i(t-s)e^{-\mathbf{G}_{\mathbf{\Gamma}}s}\mathbf{\Gamma}^T\mathbf{\Xi_2^T}ds, \qquad i = 1, \cdots, N. \tag{5.40}$$

Taking the time derivative on both sides of (5.37), it follows from (5.28) that

$$\begin{aligned}\dot{\mathbf{w}}_{1,i}(t) &= k_i(0)e^{-\mathbf{G}_{\mathbf{\Gamma}}t}\mathbf{\Xi_1} + \int_0^t \sum_{j=1}^N a_{ij}k_j(t-s)e^{-\mathbf{G}_{\mathbf{\Gamma}}s}\mathbf{\Xi_1}ds \\ &= k_i(0)e^{-\mathbf{G}_{\mathbf{\Gamma}}t}\mathbf{\Xi_1} + \sum_{j=1}^N a_{ij}\mathbf{w}_{1,j}(t),\end{aligned} \tag{5.41a}$$

with

$$\mathbf{w}_{1,i} = 0. \tag{5.41b}$$

Similarly, we have

$$\dot{\mathbf{w}}_{2,i}(t) = k_i(0)e^{-\mathbf{G}_{\mathbf{\Gamma}}t}\mathbf{\Xi_1^T} + \sum_{j=1}^N a_{ij}\mathbf{w}_{2,j}(t), \tag{5.42a}$$

with

$$\mathbf{w}_{2,i} = 0, \tag{5.42b}$$

$$\dot{\mathbf{w}}_{3,i}(t) = k_i(0)e^{-\mathbf{G}_{\mathbf{\Gamma}}t}\mathbf{\Gamma}\mathbf{\Xi_2} + \sum_{j=1}^{N} a_{ij}\mathbf{w}_{3,j}(t), \tag{5.43a}$$

with

$$\mathbf{w}_{3,i} = 0, \tag{5.43b}$$

and

$$\dot{\mathbf{w}}_{4,i}(t) = k_i(0)e^{-\mathbf{G}_{\mathbf{\Gamma}}t}\mathbf{\Gamma}^T\mathbf{\Xi_2}^T + \sum_{j=1}^{N} a_{ij}\mathbf{w}_{4,j}(t), \tag{5.44a}$$

with

$$\mathbf{w}_{4,i} = 0. \tag{5.44b}$$

Since

$$\begin{aligned}\boldsymbol{\eta}_1(t) &= \sum_{i=1}^{N}\alpha_i^N\int_0^t k_i(t-s)e^{-\mathbf{G}_{\mathbf{\Gamma}}s}\mathbf{\Xi_1}ds + \sum_{i=1}^{N}\alpha_i^N\int_0^t k_i(t-s)e^{-\mathbf{G}_{\mathbf{\Gamma}}s}\mathbf{\Xi_1}^T ds\\ &= \sum_{i=1}^{N}\alpha_i^N(\mathbf{w}_{1,i}(t)+\mathbf{w}_{2,i}(t)),\end{aligned} \tag{5.45}$$

and

$$\begin{aligned}\boldsymbol{\eta}_2(t) &= \sum_{i=1}^{N}\beta_i^N\int_0^t k_i(t-s)e^{-\mathbf{G}_{\mathbf{\Gamma}}s}\mathbf{\Gamma}\mathbf{\Xi_2}ds + \sum_{i=1}^{N}\beta_i^N\int_0^t k_i(t-s)e^{-\mathbf{G}_{\mathbf{\Gamma}}s}\mathbf{\Gamma}^T\mathbf{\Xi_2}^T ds\\ &= \sum_{i=1}^{N}\beta_i^N(\mathbf{w}_{3,i}(t)+\mathbf{w}_{4,i}(t)),\end{aligned} \tag{5.46}$$

(5.34) is now approximated by the following system of ordinary differential equations

$$
\begin{aligned}
\dot{\mathbf{K}}^N = \mathbf{G}_{\mathbf{\Gamma}}\mathbf{K}^N(t) + \mathbf{K}^N\mathbf{G}_{\mathbf{\Gamma}}^T + e^{\mathbf{G}_{\mathbf{\Gamma}}t}\sum_{i=1}^{N}\alpha_i^N\mathbf{w}_{1,i}(t)\mathbf{Q}\mathbf{\Xi}_{\mathbf{1}}^T \\
+ e^{\mathbf{G}_{\mathbf{\Gamma}}t}\sum_{i=1}^{N}\alpha_i^N\mathbf{w}_{2,i}(t)\mathbf{Q}^T\mathbf{\Xi}_{\mathbf{1}} \\
+ e^{\mathbf{G}_{\mathbf{\Gamma}}t}\sum_{i=1}^{N}\beta_i^N\mathbf{w}_{3,i}(t)\mathbf{Q}_0\mathbf{\Xi}_{\mathbf{2}}^T\mathbf{\Gamma}^T \\
+ e^{\mathbf{G}_{\mathbf{\Gamma}}t}\sum_{i=1}^{N}\beta_i^N\mathbf{w}_{4,i}(t)\mathbf{Q}_0^T\mathbf{\Xi}_{\mathbf{2}}\mathbf{\Gamma},
\end{aligned} \tag{5.47}
$$

$$
\dot{\mathbf{w}}_{1,i}(t) = k_i(0)e^{-\mathbf{G}_{\mathbf{\Gamma}}t}\mathbf{\Xi}_{\mathbf{1}} + \sum_{j=1}^{N}a_{ij}\mathbf{w}_{1,j}(t), \tag{5.48}
$$

$$
\dot{\mathbf{w}}_{2,i}(t) = k_i(0)e^{-\mathbf{G}_{\mathbf{\Gamma}}t}\mathbf{\Xi}_{\mathbf{1}}^T + \sum_{j=1}^{N}a_{ij}\mathbf{w}_{2,j}(t), \tag{5.49}
$$

$$
\dot{\mathbf{w}}_{3,i}(t) = k_i(0)e^{-\mathbf{G}_{\mathbf{\Gamma}}t}\mathbf{\Gamma}\mathbf{\Xi}_{\mathbf{2}} + \sum_{j=1}^{N}a_{ij}\mathbf{w}_{3,j}(t), \tag{5.50}
$$

$$
\dot{\mathbf{w}}_{4,i}(t) = k_i(0)e^{-\mathbf{G}_{\mathbf{\Gamma}}t}\mathbf{\Gamma}^T\mathbf{\Xi}_{\mathbf{2}}^T + \sum_{j=1}^{N}a_{ij}\mathbf{w}_{4,j}(t), \tag{5.51}
$$

$$
\mathbf{k}(0) = \mathbf{k}_0, \tag{5.52}
$$

$$
\mathbf{w}_{1,j}(0) = \mathbf{0}, \qquad \mathbf{w}_{2,j} = \mathbf{0}, \qquad \mathbf{w}_{3,j}(0) = \mathbf{0}, \qquad \mathbf{w}_{4,j} = \mathbf{0} \tag{5.53}
$$

Let the optimal control problem with its system dynamics governed by (5.47)-(5.51) with initial conditions (5.52)-(5.53) be referred to as Problem $(\hat{P}(N))$. Problem (P(N)) is equivalent to Problem $(\hat{P}(N))$, as (5.34) with initial condition (5.35) is equivalent to (5.47)-(5.51) with initial conditions (5.52)-(5.53).

Remark 1: The system of differential equations (5.47)-(5.51) with initial conditions (5.52)-(5.53) is much easier to solve than the system of differential equations involving convolutional integrals given by (5.34) with initial condition (5.35).

5.5 Analysis of Method

In this section, we shall discuss some convergence properties relating to the approximation of Problem (P) by the sequence of approximate problems $(\hat{P}(N))$. We need Lemma 4.6 of [144], which is quoted below

Lemma 2. *Let* $R_1^N(t) = \varphi_{H_1}(t) - \varphi_{H_1}^N(t)$ *and* $R_2^N(t) = \varphi_{H_2}(t) - \varphi_{H_2}^N(t)$. *Then,*

$$\begin{aligned}\|R_1^N\|_\infty &\leq \frac{C(N)(\frac{\pi T}{4})^m (N-m)!}{N!}\|\varphi_{H_1}^{(m)}\|_\infty \\ &= O(C(N)N^{-m})\end{aligned} \tag{5.54}$$

and

$$\begin{aligned}\|R_2^N\|_\infty &\leq \frac{C(N)(\frac{\pi T}{4})^m (N-m)!}{N!}\|\varphi_{H_2}^{(m)}\|_\infty \\ &= O(C(N)N^{-m})\end{aligned} \tag{5.55}$$

where $m \leq N-1$, $C(N) = \frac{2}{\pi}\log N + 2$ *and* $\|R_i^N\|_\infty = \max_{0\leq t\leq T}\|R_i^N(t)\|$ *for* $\{i = 1, 2\}$.

Theorem 1. *Let* R_1^N *and* R_2^N *be as defined in Lemma 2. Suppose that* $\mathbf{K}(t)$ *and* $\mathbf{K}^N(t)$ *are the error covariance matrices for Problem (P) and (P(N)), respectively. Then*

$$\lim_{N\to\infty}\|\mathbf{K}^N - \mathbf{K}\|_\infty = 0.$$

Proof. Define $\mathbf{u}^N(t) = \mathbf{K}(t) - \mathbf{K}^N(t)$. By taking the time derivative of $\mathbf{u}^N$, and using (5.22) and (5.34), we obtain

$$\begin{aligned}\dot{\mathbf{u}}^N(t) =& \mathbf{G_\Gamma}\mathbf{u}^N(t) + \mathbf{u}^N(t)\mathbf{G_\Gamma}^T + e^{\mathbf{G_\Gamma}t}\{\int_0^t R_1^N(t-s)e^{-\mathbf{G_\Gamma}s}\mathbf{\Xi_1}ds\}\mathbf{Q}\mathbf{\Xi_1}^T \\ &+ e^{\mathbf{G_\Gamma}t}\{\int_0^t R_1^N(t-s)e^{-\mathbf{G_\Gamma}s}\mathbf{\Xi_1}^T ds\}\mathbf{Q}\mathbf{\Xi_1} \\ &+ e^{\mathbf{G_\Gamma}t}\{\int_0^t R_2^N(t-s)e^{-\mathbf{G_\Gamma}s}\mathbf{\Gamma}\mathbf{\Xi_2}ds\}\mathbf{Q}_0\mathbf{\Xi_2}^T\mathbf{\Gamma}^T \\ &+ e^{\mathbf{G_\Gamma}t}\{\int_0^t R_2^N(t-s)e^{-\mathbf{G_\Gamma}s}\mathbf{\Gamma}^T\mathbf{\Xi_2}^T ds\}\mathbf{Q}_0^T\mathbf{\Xi_2}\mathbf{\Gamma}\end{aligned} \tag{5.56a}$$

with initial condition

$$\mathbf{u}^N(0) = \mathbf{0}. \tag{5.56b}$$

Let

$$\begin{aligned}\boldsymbol{\beta}^N(t) = e^{\mathbf{G_\Gamma} t}\{\int_0^t R_1^N(t-s)e^{-\mathbf{G_\Gamma} s}\boldsymbol{\Xi_1} ds\}\mathbf{Q}\boldsymbol{\Xi_1}^T \\ + e^{\mathbf{G_\Gamma} t}\{\int_0^t R_1^N(t-s)e^{-\mathbf{G_\Gamma} s}\boldsymbol{\Xi_1}^T ds\}\mathbf{Q}\boldsymbol{\Xi_1} \\ + e^{\mathbf{G_\Gamma} t}\{\int_0^t R_2^N(t-s)e^{-\mathbf{G_\Gamma} s}\boldsymbol{\Gamma}\boldsymbol{\Xi_2} ds\}\mathbf{Q}_0\boldsymbol{\Xi_2}^T\boldsymbol{\Gamma}^T \\ + e^{\mathbf{G_\Gamma} t}\{\int_0^t R_2^N(t-s)e^{-\mathbf{G_\Gamma} s}\boldsymbol{\Gamma}^T\boldsymbol{\Xi_2}^T ds\}\mathbf{Q}_0^T\boldsymbol{\Xi_2}\boldsymbol{\Gamma}\end{aligned} \tag{5.57}$$

Then, we have

$$\begin{aligned}&\|\boldsymbol{\beta}^N(t)\| \\ &\leq \|e^{\mathbf{G_\Gamma} t}\|\{\int_0^t \|R_1^N(t-s)\|\|e^{-\mathbf{G_\Gamma} s}\|\|\boldsymbol{\Xi_1}\| ds\}\|\mathbf{Q}\|\|\boldsymbol{\Xi_1}^T\| \\ &+ \|e^{\mathbf{G_\Gamma} t}\|\{\int_0^t \|R_1^N(t-s)\|\|e^{-\mathbf{G_\Gamma} s}\|\|\boldsymbol{\Xi_1}^T\| ds\}\|\mathbf{Q}\|\|\boldsymbol{\Xi_1}\| \\ &+ \|e^{\mathbf{G_\Gamma} t}\|\{\int_0^t \|R_2^N(t-s)\|\|e^{-\mathbf{G_\Gamma} s}\|\|\boldsymbol{\Gamma}\|\|\boldsymbol{\Xi_2}\| ds\}\|\mathbf{Q}_0\|\|\boldsymbol{\Xi_2}^T\|\|\boldsymbol{\Gamma}^T\| \\ &+ \|e^{\mathbf{G_\Gamma} t}\|\{\int_0^t \|R_2^N(t-s)\|\|e^{-\mathbf{G_\Gamma} s}\|\|\boldsymbol{\Gamma}^T\|\|\boldsymbol{\Xi_2}^T\| ds\}\|\mathbf{Q}_0^T\|\|\boldsymbol{\Xi_2}\|\|\boldsymbol{\Gamma}\|.\end{aligned} \tag{5.58}$$

From Lemma 2, it follows that for any $\varepsilon > 0$, there exists an N_0 such that for all $N > N_0$, we have

$$\|R_1^N\|_\infty \leq \varepsilon \tag{5.59}$$

and

$$\|R_2^N\|_\infty \leq \varepsilon. \tag{5.60}$$

Furthermore, since $\mathbf{A}, \mathbf{H}, \boldsymbol{\Xi_1}$, and $\boldsymbol{\Xi_2}$ are all constant matrices and $e^{\mathbf{G_\Gamma} t}$, $e^{-\mathbf{G_\Gamma} t}$ and $\boldsymbol{\beta}^N(t)$ are all continuous on $[0, T]$, there exists a constant $\tilde{M}$ such that

$$\|\mathbf{A}\|_\infty = \max_{1\leq i,j\leq n} |A_{i,j}| \leq \tilde{M}, \qquad \|\mathbf{H}\|_\infty = \max_{1\leq i,j\leq n} |H_{i,j}| \leq \tilde{M},$$

$$\|e^{\mathbf{G_\Gamma}}\|_\infty = \max_{0\leq t\leq T} \|e^{\mathbf{G_\Gamma} t}\| \leq \tilde{M}, \qquad \|e^{-\mathbf{G_\Gamma}}\|_\infty = \max_{0\leq t\leq T} \|e^{-\mathbf{G_\Gamma} t}\| \leq \tilde{M},$$

$$\|\boldsymbol{\Xi_1}\|_\infty = \max_{1\leq i,j\leq n} |\Xi_{1(i,j)}| \leq \tilde{M}, \qquad \|\boldsymbol{\Xi_2}\|_\infty = \max_{1\leq i,j\leq n} |\Xi_{2(i,j)}| \leq \tilde{M}$$

$$\text{and} \qquad \|\boldsymbol{\beta}^N\|_\infty = \max_{0\leq t\leq T} \|\boldsymbol{\beta}^N(t)\| \leq \tilde{M},$$

where T is the terminal time of the process $x(t)$.

Thus, (5.58) is reduced to

$$\begin{aligned}
&\|\boldsymbol{\beta}^N(t)\| \\
&\leq \|e^{\mathbf{G_\Gamma}t}\|\{\int_0^t \varepsilon\|e^{-\mathbf{G_\Gamma}s}\|\|\mathbf{\Xi_1}\|ds\}\|\mathbf{Q}\|\|\mathbf{\Xi_1^T}\| \\
&+ \|e^{\mathbf{G_\Gamma}t}\|\{\int_0^t \varepsilon\|e^{-\mathbf{G_\Gamma}s}\|\|\mathbf{\Xi_1^T}\|ds\}\|\mathbf{Q}\|\|\mathbf{\Xi_1}\| \\
&+ \|e^{\mathbf{G_\Gamma}t}\|\{\int_0^t \varepsilon\|e^{-\mathbf{G_\Gamma}s}\|\|\mathbf{\Gamma}\|\|\mathbf{\Xi_2}\|ds\}\|\mathbf{Q}_0\|\|\mathbf{\Xi_2^T}\|\|\mathbf{\Gamma}^T\| \\
&+ \|e^{\mathbf{G_\Gamma}t}\|\{\int_0^t \varepsilon\|e^{-\mathbf{G_\Gamma}s}\|\|\mathbf{\Gamma}^T\|\|\mathbf{\Xi_2^T}\|ds\}\|\mathbf{Q}_0^T\|\|\mathbf{\Xi_2}\|\|\mathbf{\Gamma}\|. \\
&\leq 4T\varepsilon\tilde{M} = \varepsilon\bar{M},
\end{aligned} \tag{5.61}$$

for all $N > N_0$, where $\bar{M} = 4T\tilde{M}$.

From (5.56) and (5.57), we have

$$\begin{aligned}
&\dot{\mathbf{u}}^N(t) = \mathbf{G_\Gamma}\mathbf{u}^N(t) + \mathbf{u}^N(t)\mathbf{G_\Gamma^T} + \boldsymbol{\beta}^N(t) \\
&\mathbf{u}^N(0) = \mathbf{0},
\end{aligned} \tag{5.62}$$

with $\mathbf{G_\Gamma} = \mathbf{A} - \mathbf{\Gamma}\mathbf{H}$. Taking integration, we obtain

$$\mathbf{u}^N(t) = \int_0^t \mathbf{G_\Gamma}\mathbf{u}^N(s)ds + \int_0^t \mathbf{u}^N(s)\mathbf{G_\Gamma^T}ds + \int_0^t \boldsymbol{\beta}^N(s)ds. \tag{5.63}$$

Note that

$$\begin{aligned}
\|\mathbf{G_\Gamma}\|_\infty &= \max_{1\leq i,j\leq n} |G_{\mathbf{\Gamma}(i,j)}| \\
&\leq \|\mathbf{A}\|_\infty + M\|\mathbf{H}\|_\infty,
\end{aligned} \tag{5.64}$$

where M is as defined in (5.14). Thus, for $N > N_0$,

$$\begin{aligned}\|\mathbf{u}^N(t)\| &= \|\int_0^t \mathbf{G}_{\mathbf{\Gamma}}\mathbf{u}^N(s)ds + \int_0^t \mathbf{u}^N(s)\mathbf{G}_{\mathbf{\Gamma}}^T ds + \int_0^t \boldsymbol{\beta}^N(s)ds\| \\ &\leq \int_0^t \|\mathbf{G}_{\mathbf{\Gamma}}\|\|\mathbf{u}^N(s)\|ds + \int_0^t \|\mathbf{u}^N(s)\|\|\mathbf{G}_{\mathbf{\Gamma}}^T\|ds + \int_0^t \|\boldsymbol{\beta}^N(s)\|ds. \\ &\leq 2(\|\mathbf{A}\|_\infty + M\|\mathbf{H}\|_\infty)\int_0^t \|\mathbf{u}^N(s)\|ds + \varepsilon\bar{M}.\end{aligned} \tag{5.65}$$

Therefore, by Gronwall's inequality, we have

$$\|\mathbf{u}^N\|_\infty \leq \bar{M}\varepsilon\exp(|2(\|\mathbf{A}\|_\infty + M\|\mathbf{H}\|_\infty)|T)$$

for $N > N_0$. Since $\varepsilon > 0$ is arbitrary, it follows that $\|\mathbf{u}^N\|_\infty \to \infty$ as $N \to \infty$. This completes the proof.

From Theorem 1, we observe that Problem (*P(N)*) converges to the original problem (P) as $N \to \infty$. Thus, to solve Problem (P) we will solve a sequence of Problem (*P(N)*), where each of which involves only system of ordinary differential equations. There are several efficient optimization techniques that can be used. For example, the optimal control software, *MISER 3.3*, is applicable for this purpose.

5.6 Numerical Simulation

In this section, we present some numerical examples to illustrate the applicability of our proposed method.

Example 1. Consider a system which is governed by the following stochastic differential equation:

$$dx = xdt + dB_{H_1}(t), \; t \in [0,1] \tag{5.66a}$$

$$x(0) = 0, \tag{5.66b}$$

The measurement dynamics is given by

$$dy = 2xdt + dB_{H_2}(t), \ t \in [0,1] \tag{5.67a}$$

$$y(0) = 0. \tag{5.67b}$$

Then, the filter system becomes

$$dz = (1-2r)\, z(t)\, dt + r dy(t), \ t \geq 0, \tag{5.68a}$$

$$z(0) = \hat{x}_0 \equiv E\{x_0\} \tag{5.68b}$$

where $r \in [-10, 10]$ is the parameter to be determined.

Suppose that the statistics of this system are given as follows:

$$H_1 = H_2 = 0.8, \ Q = Q_0 = 0.01.$$

Then, r can be obtained by solving the following optimal control problem.

$$\min_r J(r) = \int_0^1 k(t)\, dt \tag{5.69}$$

subject to

$$\dot{k}(t) = 2(1-2r)\, k(t) + 0.04 \int_0^t \varphi_{H_1}(t-s)\, e^{(1-2r)(t-s)} ds \tag{5.70a}$$

$$k(0) = E\left\{(x_0 - \hat{x}_0)^2\right\}. \tag{5.70b}$$

Let this problem be referred to as Problem (Q). We construct the corresponding approximate optimal control Problem (Q(N))with $N = 7$, 13 and 19. The corresponding values of the optimal parameter r^* and the cost obtained are: $\{9.59496, 0.000428209\}$,$\{9.7218, 0.000433229\}$ and $\{9.9543, 0.000403976\}$, respectively. The time histories of the approximate states are plotted in Figure 5.1. From Figure 5.1, we see that the convergence is very fast with respect to N, and K^N with $N = 19$ can be regarded as the solution of system (5.70) with $r^* = 9.9543$.

Example 2. Consider a dynamical system given by the following stochas-

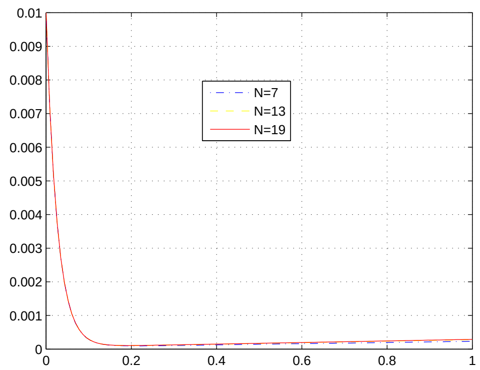

Figure 5.1: The state $k(t)$ with $N = 7$, $N = 13$ and $N = 19$.

tic differential equations defined on $[0, 1]$.

$$d\mathbf{x} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \mathbf{x} dt + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} d\mathbf{B}_{H_1}(t), \tag{5.71a}$$

$$\mathbf{x}(0) = [0\ 0]^T. \tag{5.71b}$$

The measurement system is given by

$$d\mathbf{y} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{x} dt + \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} d\mathbf{B}_{H_2}(t), \tag{5.72a}$$

$$\mathbf{y}(0) = [0\ 0]^T. \tag{5.72b}$$

Then, the filter system becomes

$$d\mathbf{z} = \left(\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} - \begin{bmatrix} \gamma_1 & \gamma_2 \\ \gamma_3 & \gamma_4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \mathbf{z} dt + \begin{bmatrix} \gamma_1 & \gamma_2 \\ \gamma_3 & \gamma_4 \end{bmatrix} d\mathbf{y}, \tag{5.73a}$$

$$\mathbf{z}(0) = \hat{\mathbf{x}}_0 = E\{\mathbf{x}_0\}. \tag{5.73b}$$

Suppose that $H_1 = H_2 = 0.8$, and that $\mathbf{Q} = \mathbf{Q}_0$ is the 2×2 identity matrix. Then,

$$\tilde{\mathbf{Q}}(s, t) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } \tilde{\mathbf{Q}}_0(s, t) = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}. \tag{5.74}$$

Now we suppose that

$$\mathbf{\Gamma} = \begin{bmatrix} \gamma_1 & \gamma_2 \\ \gamma_3 & \gamma_4 \end{bmatrix}, \ \mathbf{\Sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{5.75}$$

where $\gamma_i \in [-10, 10]$, $i = 1, \cdots, 4$, are the parameters to be selected by solving the following optimal control problem.

$$\min_{\mathbf{\Gamma}} J(\mathbf{\Gamma}) = \int_0^1 \mathrm{trace}\{\mathbf{\Sigma K}\} dt \tag{5.76}$$

subject to

$$\begin{aligned}\dot{\mathbf{K}}(t) = \mathbf{G}_{\mathbf{\Gamma}}(t)\mathbf{K}(t) + \mathbf{K}\mathbf{G}_{\mathbf{\Gamma}}^T(t) + e^{\mathbf{G}_{\mathbf{\Gamma}} t}\{\int_0^t \varphi_{H_1}(t-s) e^{-\mathbf{G}_{\mathbf{\Gamma}} s} \tilde{\mathbf{Q}}(s,t) ds\} \\ + e^{\mathbf{G}_{\mathbf{\Gamma}} t}\{\int_0^t \varphi_{H_1}(t-s) e^{-\mathbf{G}_{\mathbf{\Gamma}} s} \tilde{\mathbf{Q}}^T(s,t) ds\} \\ + e^{\mathbf{G}_{\mathbf{\Gamma}} t}\{\int_0^t \varphi_{H_2}(t-s) e^{-\mathbf{G}_{\mathbf{\Gamma}} s} \mathbf{\Gamma}(s) \tilde{\mathbf{Q}}_0(s,t) ds\} \mathbf{\Gamma}^T(t) \\ + e^{\mathbf{G}_{\mathbf{\Gamma}} t}\{\int_0^t \varphi_{H_2}(t-s) e^{-\mathbf{G}_{\mathbf{\Gamma}} s} \mathbf{\Gamma}^T(s) \tilde{\mathbf{Q}}_0^T(s,t) ds\} \mathbf{\Gamma}(t),\end{aligned} \tag{5.77a}$$

$$\mathbf{K}(0) = \mathbf{K}_0, \tag{5.77b}$$

where

$$\mathbf{G}_{\mathbf{\Gamma}}(t) = \begin{bmatrix} -\gamma_1 & 1-\gamma_2 \\ 1-\gamma_3 & 1-\gamma_4 \end{bmatrix}, \ e^{\mathbf{G}_{\mathbf{\Gamma}} t} = \exp\{\int_0^t \begin{bmatrix} -\gamma_1 & 1-\gamma_2 \\ 1-\gamma_3 & 1-\gamma_4 \end{bmatrix}\} dt,$$

$$e^{-\mathbf{G}_{\mathbf{\Gamma}} t} = \exp\{-\int_0^t \begin{bmatrix} -\gamma_1 & 1-\gamma_2 \\ 1-\gamma_3 & 1-\gamma_4 \end{bmatrix}\} dt,$$

$\tilde{\mathbf{Q}}$, $\tilde{\mathbf{Q}}_0$ and Γ from (5.74) and (5.75), respectively.

Let this problem be referred to as Problem (Q). We now construct the approximate optimal control Problem (*Q(N)*) with $N = 5$. Then, the optimal control software package, *MISER 3.3*, is used to solve such an approximate optimal control problem. The optimal Γ^* obtained is:

$$\boldsymbol{\Gamma}^*(t) = \begin{bmatrix} 9.9965 & 1.12784 \\ 1.14684 & 9.9965 \end{bmatrix},$$

and the optimal cost obtained is 0.104372705.

The time histories of the components of the state $\mathbf{K}^*$, i.e., the solution of the system (5.77) with $N = 5$ corresponding to $\boldsymbol{\Gamma} = \boldsymbol{\Gamma}^*$ are plotted in Figure 5.2 and Figure 5.3. For $N = 13$, the optimal cost obtained is 0.105569817 and the optimal $\boldsymbol{\Gamma}^*$ solution obtained is

$$\boldsymbol{\Gamma}^* = \begin{bmatrix} 9.9987 & 1.1603 \\ 1.11603 & 9.9987 \end{bmatrix}.$$

The time histories of the components of the corresponding $\mathbf{K}^*$ are plotted in Figure 5.4 and Figure 5.5.

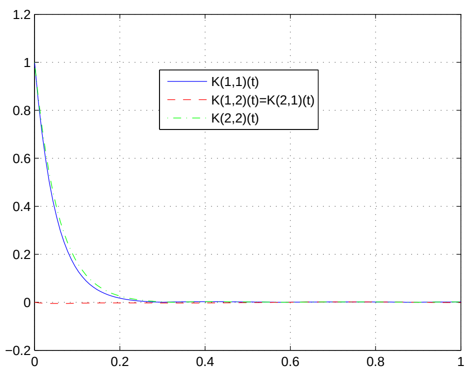

Figure 5.2: The state $K^*(t)$ with $N = 5$.

From these examples, we can say that the method proposed is efficient. The figures show that the error covariance matrices converge very fast with respect to the observation data. These imply that if more observation data is available, then the estimation of $x(t)$ will be more accurate. The large error that can be seen occurs earlier are basically caused by the fact that the initial error covariance matrices considered in these examples are large.

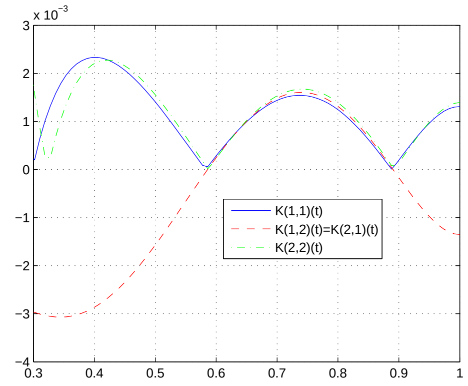

Figure 5.3: Zoom of the state $K^*(t)$ with $N = 5$.

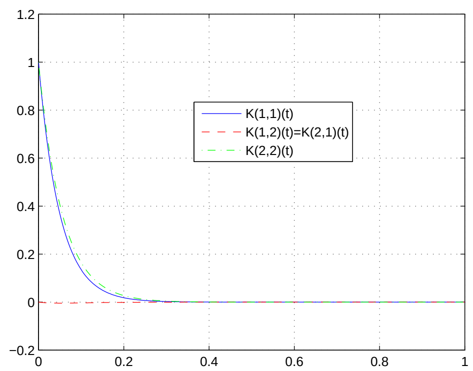

Figure 5.4: The state $K^*(t)$ with $N = 13$.

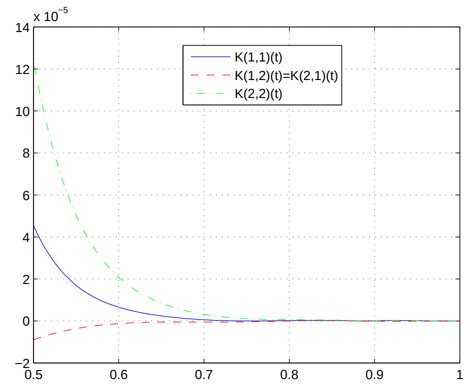

Figure 5.5: Zoom of the state $K^*(t)$ with $N = 13$.

5.7 Discussion

In this chapter, we studied a continuous time filtering of a multi-dimensional stochastic differential system driven by a fractional Brownian motion process. We showed that this filtering problem is equivalent to an optimal control problem involving convolutional integrals in its dynamical system. Then, a novel approximation scheme is developed and applied to this optimal control problem. It yields a sequence of standard optimal control problems. The convergence of the approximate standard optimal control problem to the optimal control problem involving convolutional integrals in its system dynamics is established. Further, two numerical examples were solved by using the method proposed. The results obtained clearly demonstrate its efficiency and effectiveness.

Chapter 6

Conclusion and Outlook

Long memory processes, i.e., the fractional Brownian motion (FBM), in financial modeling, is the subject of this thesis. Intuitively, the financial markets behave quite differently from the Brownian motion (BM), as BM is independent of the past. Thus, by using BM, it will mean that the financial markets have no memory of historical shocks in the financial time series. Despite of this, BM has always been a backbone in most finance models. By using FBM, the influence of the whole history of data must be taken into account [123]. This thesis addresses some related problems in financial environment.

In Chapter 2, we presented some Hurst estimators available in the literature. We chose two most prominent estimators, the R/S analysis and Whittle method, to estimate the Hurst index H. H is the parameter in FBM. It indicates how weak/strong the dependency is in our data. First, a fractional integrated (FI) model and fractional integrated GARCH (FI-GARCH) model are generated. These models are selected because they allow the dependency in their data. In simulation study, we found that the use of the R/S analysis is sufficient. The differences between the performance of these two estimators are not significant. Later, the R/S analysis was applied to analyze the Malaysia financial data. The advantages and disadvantages of these methods were also discussed.

Motivated by the drawbacks found in Hurst estimators, we investigated the model based estimation of long memory indexes in Chapter 3 and Chapter 4. In reality, financial markets do have a long memory [107]. Thus, the FBM processes are applied to model these markets. With this in mind, Chapter 3 is devoted to the fractional Black-Scholes (FBS) model. We use the geometric fractional Brownian motion (GFBM) to represent the asset prices. Our aim is to estimate unknown parameters in GFBM. The estimations of these parameters are of great interest to researchers and practitioners. The parameters obtained from this method are used in the FBS model. FBS can capture the dependency in assets. This is an advantage over the standard Black-Scholes model. We illustrated the efficiency of our method by some simulation study. The results obtained showed that our propose method is efficient. We also conducted some empirical investigation to stock indices, and applied our method to European option pricing.

Chapter 4 studied another important model in financial modeling, i.e. the fractional Ornstein-Uhlenbeck (FOU) model. This model is important in the modeling of the interest rate. We considered the Vasicěk model which is disturbed by FBM. The dynamic measurement error is manipulated in the development of the likelihood function by using the innovation algorithm. We then used the simulation results obtained in our discussion on the efficiency of our method.

Chapter 5 investigated the filtering problem, where a continuous time linear system driven by FBM is considered. We showed that this filtering problem is equivalent to an optimal control problem involving convolutional integrals in its dynamical system. We proposed an approximation scheme to cater for this problem. We also established the convergence of the approximate optimal control problem to the optimal control problem involving convolutional integrals in this system. Finally, two numerical examples were solved by using our proposed method. The results obtained showed that the method is highly efficient and effective.

6.1 Future Research Problems

There are some challenging problems that arise from the work in this thesis for future research. In Chapter 3, we assume that the volatility in GFBM model is constant. Currently, there are no results available in the literature on time varying volatility based on GFBM model. This is an interesting and highly significant research topic. The model would be more accurate in describing a real market behavior. The estimates of these parameters that are obtained will be more significant. This is an important future research problem.

In Chapter 4, we use the innovation algorithm to help construct the likelihood function. We then use the simulated annealing method to search for the optimal parameters. A more natural approach is to search for the complete maximum likelihood estimators, where the analytical gradient formulas of the likelihood function with respect to its parameters are to be derived and used in the optimization process. The task appears to be highly complicated and we have yet to resolve this problem.

In Chapter 5, the system matrices of the dynamical system, and the measurement dynamics are assumed to be constant matrices. This assumption is required so that the proposed approximation scheme can be carried out. It does not seem possible to obtain similar approximation scheme for the case when the system matrices are time dependent. This is an interesting future research problem.

Appendix A

Derivation of complete maximum likelihood estimation

We will now derive the estimators for μ_1 and σ_1^2 in regards to the log likelihood obtained in (3.25). The partial derivatives of $\ell_n(\theta)$ with respect to μ_1 and σ_1^2 are given by

$$
\begin{aligned}
\frac{\partial \ell}{\partial \mu_1} &= -\frac{1}{2\sigma_1^2}\{(Z-\mu_1\mathbf{1}+\frac{1}{2}\sigma_1^2\mathbf{x}_H)'\Sigma_0^{-1}(-\mathbf{1})+(-\mathbf{1})'\Sigma_0^{-1}(Z-\mu_1\mathbf{1}+\frac{1}{2}\sigma_1^2\mathbf{x}_H)\} \\
&= \frac{1}{\sigma_1^2}\mathbf{1}'\Sigma_0^{-1}(Z-\mu_1\mathbf{1}+\frac{1}{2}\sigma_1^2\mathbf{x}_H), \\
\frac{\partial \ell}{\partial \sigma_1^2} &= -\frac{n}{2\sigma_1^2}-\frac{1}{2}\frac{1}{(\sigma_1^2)^2}\{\sigma_1^2\{(Z-\mu_1\mathbf{1}+\frac{1}{2}\sigma_1^2\mathbf{x}_H)'\Sigma_0^{-1}\frac{1}{2}\mathbf{x}_H+\frac{1}{2}\mathbf{x}_H'\Sigma_0^{-1}(Z-\mu_1\mathbf{1}+\frac{1}{2}\sigma_1^2\mathbf{x}_H)\} \\
&\quad -\{(Z-\mu_1\mathbf{1}+\frac{1}{2}\sigma_1^2\mathbf{x}_H)'\Sigma_0^{-1}(Z-\mu_1\mathbf{1}+\frac{1}{2}\sigma_1^2\mathbf{x}_H)\}\} \\
&= -\frac{n}{2\sigma_1^2}+\frac{1}{2(\sigma_1^2)^2}\{(Z-\mu_1\mathbf{1}-\frac{1}{2}\sigma_1^2\mathbf{x}_H)'\Sigma_0^{-1}(Z-\mu_1\mathbf{1}+\frac{1}{2}\sigma_1^2\mathbf{x}_H)\},
\end{aligned}
$$

Setting them equal to zero yields

$$
\mu_1 = \frac{1}{\mathbf{1}'\Sigma_0^{-1}\mathbf{1}}(\mathbf{1}'\Sigma_0^{-1}Z+\frac{1}{2}\sigma_1^2\mathbf{1}'\Sigma_0^{-1}\mathbf{x}_H) \tag{A.1}
$$

and

$$
\begin{aligned}
\sigma_1^2 &= \frac{1}{n}(Z - \mu_1\mathbf{1} - \frac{1}{2}\sigma_1^2\mathbf{x}_H)'\Sigma_0^{-1}(Z - \mu_1\mathbf{1} + \frac{1}{2}\sigma_1^2\mathbf{x}_H) \\
&= \frac{1}{n}(Z - \frac{1}{2}\sigma_1^2\mathbf{x}_H)'\Sigma_0^{-1}(Z - (\frac{\mathbf{1}'\Sigma_0^{-1}Z + \frac{1}{2}\sigma_1^2\mathbf{1}'\Sigma_0^{-1}\mathbf{x}_H}{\mathbf{1}'\Sigma_0^{-1}\mathbf{1}})\mathbf{1} + \frac{1}{2}\sigma_1^2\mathbf{x}_H) \\
&= \frac{1}{n}(Z - \frac{1}{2}\sigma_1^2\mathbf{x}_H)'\Sigma_0^{-1}(\mathbf{I} - (\frac{\mathbf{1}\mathbf{1}'\Sigma_0^{-1}}{\mathbf{1}'\Sigma_0^{-1}\mathbf{1}}))(Z + \frac{1}{2}\sigma_1^2\mathbf{x}_H),
\end{aligned}
$$

respectively. By substituting $\Sigma_1 = \Sigma_0^{-1}(\mathbf{I} - (\frac{\mathbf{1}\mathbf{1}'\Sigma_0^{-1}}{\mathbf{1}'\Sigma_0^{-1}\mathbf{1}}))$, the above equation is simplified to

$$
\begin{aligned}
\sigma_1^2 &= \frac{1}{n}(Z - \frac{1}{2}\sigma_1^2\mathbf{x}_H)'\Sigma_1(Z + \frac{1}{2}\sigma_1^2\mathbf{x}_H) \\
&= \frac{1}{n}Z'\Sigma_1 Z - \sigma_1^2(\frac{\mathbf{x}_H'\Sigma_1 Z - Z'\Sigma_1\mathbf{x}_H}{2n}) - \sigma_1^4(\frac{\mathbf{x}_H'\Sigma_1\mathbf{x}_H}{4n}).
\end{aligned}
$$

Notice from this equality that the solution for σ^2 can be obtained by solving the quadratic equation

$$
\sigma_1^4(\frac{1}{4n}\mathbf{x}_H'\Sigma_1\mathbf{x}_H) + \sigma_1^2 - \frac{1}{n}Z'\Sigma_1 Z = 0,
$$

which, when $\mathbf{x}_H'\Sigma_1\mathbf{x}_H \neq 0$, gives

$$
\begin{aligned}
\widehat{\sigma}_1^2 &= \frac{\sqrt{1 + \frac{1}{n^2}\mathbf{x}_H'\Sigma_1\mathbf{x}_H Z'\Sigma_1 Z} - 1}{\frac{1}{2n}\mathbf{x}_H'\Sigma_1\mathbf{x}_H} \\
&= \frac{1 + \frac{1}{n^2}\mathbf{x}_H'\Sigma_1\mathbf{x}_H Z'\Sigma_1 Z - 1}{\frac{1}{2n}\mathbf{x}_H'\Sigma_1\mathbf{x}_H(\sqrt{1 + \frac{1}{n^2}\mathbf{x}_H'\Sigma_1\mathbf{x}_H Z'\Sigma_1 Z} + 1)} \\
&= \frac{2Z'\Sigma_1 Z}{\sqrt{n^2 + \mathbf{x}_H'\Sigma_1\mathbf{x}_H Z'\Sigma_1 Z} + n}. \qquad \text{(A.2)}
\end{aligned}
$$

Note that even when $\mathbf{x}_H'\Sigma_1\mathbf{x}_H = 0$, the equality (A.2) still remains valid. Therefore (3.29) and (3.30) in Subsection 3.5 follow from (A.2) and (A.1), respectively.

References

[1] O.O. Aalen and H.K. Gjessing. Survival models based on the Ornstein-Uhlenbeck process. *Lifetime Data Analysis*, 10:407–423, 2004.

[2] P. Abry, P. Flandrin, M. Taqqu, and D. Veitch. *Self-Similar Network Traffic and Performance Evaluation*, chapter Wavelets for the analysis, estimation and synthesis of scaling data, pages 39–88. Wiley, New York, 2000.

[3] N.U. Ahmed. *Linear and Nonlinear Filtering for Scientists and Engineers*. World Scientific, 1998.

[4] N.U. Ahmed and C.D. Charalambous. Filtering for linear systems driven by fractional Brownian motion. *SIAM J. Control Optim.*, 41(1):313–330, 2002.

[5] B.D.O. Anderson and J.B. Moore. *Optimal Filtering*. Prentice Hall, Inc, 1979.

[6] B.O.O. Anderson and J.B. Moore. *Linear Optimal Control*. Prentice-Hall, Englewood Cliffs, N.J., 1979.

[7] D.W.K. Andrew and Y. Sun. Adaptive local polynomial Whittle estimation of long-rang dependence. *Econometrica*, 72:569–614, 2004.

[8] D.W.K. Andrews and Y. Sun. Adaptive local polynomial Whittle estimation of long-range dependence. *Econometrica*, 72:569–614, 2004.

[9] M. Ausloos. Statistical physics in foreign exchange currency and stock markets. *Physica A*, 258:48–65, 2000.

[10] L. Bachelier. *Théorie de la Spéculation*. PhD thesis, Annales Scientiques de l'École Normale Supérieure, 1900.

[11] R.T. Baillie, T. Bollerslev, and H.O. Mikkelsen. Fractionally integrated autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 74:3–30, 1996.

[12] O.E. Barndorff-Nielson and N. Shephard. Non-gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society, Series B*, 63:167–241, 2001.

[13] D. Bendel, W.D. Hamman, and E.M. Smith. Some evidence of persistence in South African financial time series. *Journal for Studies in Economics & Econometrics*, 20(1):9–83, 1996.

[14] J. Beran. *Statistics for Long-Memory Processes*. Chapman & Hall, New York, 1994.

[15] F. Biagini, Y. Hu, B. Øksendal, and T. Zhang. *Stochastic Calculus for Fractional* B*rownian Motion and Applications*. Springer-Verlag London Ltd., 2008.

[16] T. Bjork and B.J. Christensen. Interest rate dynamics and consistent forward rate curves. *Mathematical Finance*, 9:223–248, 1999.

[17] T. Bjork and H. Hult. A note on Wick products and the fractional Black Scholes model. *Finance and Stochastics*, 9(2):197–209, 2005.

[18] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654, 1973.

[19] T BOLLERSLEV. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327, 1986.

[20] P. Bourke. An introduction to fractals. http://astronomy.swin.edu.au/ pbourke/fractals/fractintro, May 1991.

[21] P.J. Brockwell and R.A. Davis. *Time Series: Theory and Methods*. Springer-Verlag, 1987.

[22] R.S. Bucy and P.D. Joseph. *Filtering for Stochastic Processes with Applications to Guidance*. New York: Interscience, 1968.

[23] D.O. Cajueiro and J.F. Barbachan. Volatility estimation and option pricing with fractional Brownian motion. SSRN:http://ssrn.com/abstract=837765, 2005.

[24] D.O. Cajueiro and B.M. Tabak. The Hurst exponent over time: Testing the assertion that emerging markets are becoming more efficient. *Physica A*, 336:521–537, 2004.

[25] D.O. Cajueiro and B.M. Tabak. Ranking efficiency of emerging equity market ii. *Chaos, Solitons and Fractals*, 23:671–675, 2005.

[26] M.J. Cannon, D.B. Percival, D.C. Caccia, G.M. Raymond, and J.B. Bassingthwaighte. Evaluationg scaled windowed variance methods for estimating the Hurst coefficient in time series. *Physica A*, 241:606–626, 1996.

[27] I. Casas and J. Gao. Econometric estimation in long-range dependent volatility models: Theory and practice. *Journal of Econometrics*, 147:72–83, 2008.

[28] N. Cassola and L.J. Barros. A two-factor model of the German term structure of interest rates. ECB Working Paper 46, 2001.

[29] C.W. Cheong. Long persistence volatility and links between national stock market indices. *Int. Research Journal of Finance and Economics*, 7:15–195, 2007.

[30] P. Cheridito, H. Kawaguchi, and M. Maejima. Fractional Ornstein-Uhlenbeck processes. *Electron. J. Probab.*, 8:1–14, 2003.

[31] S.H. Chun, K.J. Kim, and S.H. Kim. Chaotic analysis of predictability versus knowledge discovery techniques: Case study of the Polish stock market. *Expert Systems*, 19:264–272, 2002.

[32] R.G. Clegg. A practical guide to measuring the Hurst parameter. *Int. J. of Simulation: Systems, Science & Technology*, 7(2):3–14, 2006.

[33] F. Comte. Simulation and estimation of long memory continuous time models. *J. Time Ser. Anal.*, 17(1):19–36, 1996.

[34] F. Comte and E. Renault. Long memory in continuous-time stochastic volatility models. *Mathematical Finance*, 8:291–323, 1998.

[35] M. Couillard and M. Davison. A comment on measuring the Hurst exponent of financial time series. *Physica A*, 348:404–418, 2005.

[36] L. Coutin and L. Decreusefond. Abstract nonlinear filtering theory in the presence of fractional brownian motion. *Ann. Appl. Probab.*, 9:1058–1090, 1999.

[37] R. Dahlhaus. Efficient parameter estimation for self-similar processes. *Ann. Statist.*, 17(4):1749–1766, 1989.

[38] F. de Jong and P. Santa-Clara. The dynamics of the forward interest rate curve: a formulation with state variables. *Journal of Financial and Quantitative Analysis*, 34:131–157, 1999.

[39] G. De Rossi. Kalman filtering of consistent forward rate curves: a tool to estimate and model dynamically the term structure. *Journal of Empirical Finance*, 11:277–308, 2004.

[40] D.K. Ding, F.H. Harris, S.T. Lau, and T.H. McInish. An investigation of price discovery in informationally-linked market: Equity trading in Malaysia and Singapore. *Journal of Multinational Financial Management*, 9:317–329, 1999.

[41] Z. Ding, C.W.J. Granger, and R.F. Engle. A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1:83–106, 1993.

[42] H.J. Edisan and C.M Reinhart. Capital controls during financial crises: The case of Malaysia and Thailand. Board of Governers of he Federal Reserve System Int. Finance Discussion Papers 662, 2000.

[43] A.G. Ellinger. *The Art of Investment.* Bowers & Bowers, London, 1971.

[44] R.J. Elliot and L Chan. Perpetual american options with fractional Brownian motion. *Quantitative Finance*, 4(2):123–128, 2004.

[45] R.J. Elliott and P.E. Kopp. *Mathematics of Financial Markets.* Springer, 2005.

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

[47] R.F. Engle. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica*, 50:987–1008, 1982.

[48] R. Fox and M. Taqqu. Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Statist.*, 14(2):517–532, 1986.

[49] J. Geweke and S. Porter-Hudak. The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4:221–238, 1983.

[50] L. Giraitis and M. Taqqu. Whittle estimator for finite-variance non gaussian time series with long memory. *The Annals of Statistics*, 27:178–203, 1999.

[51] W.N. Goetzmann. Patterns in three centuries of stock market prices. *The Journal of Business*, 66(2):249–270, 1993.

[52] S.K. Goh, M.H. Alias, and N. Olekalns. New evidence of financial openness in Malaysia. *Journal of Asian Economics*, 4:311–325, 2003.

[53] M.T. Greene and B.D. Fielitz. Long term dependence in common stock returns. *Journal of Financial Econometrics*, 4:339–349, 1977.

[54] S. Haykin. *Adaptive Filter Theory.* Prentice Hall, 2001.

[55] T. Higuchi. Approach to an irregular time series on the basis of the fractal theory. *Physica D*, 31:277–283, 1988.

[56] E.P. Hog and P.H. Frederiksen. The fractional Ornstein-Uhlenbeck process: Term structure theory and application. Scandinavian Working Papers in Business Administration, 2006. No:F-2006-01.

[57] H.A.A.B. Hoque, J.H. Kim, and C.S. Pyun. A comparison of variance ratio tests of random walk: A case of Asian emerging stock markets. *Int. Rev. of Economics and Finance*, 2006.

[58] J.S. Howe, W.D. Martin, and B.D.J. Wood. Fractal structure in the Pacific rim. Southwestern Finance Assoc, Annual Meeting, New Orleans, March 1997.

[59] Y. Hu. Option pricing in a market where the volatility is driven by fractional Brownian motions. In *Recent Development in Mathematical Finance*, 2001.

[60] Y. Hu and B. Øksendal. Fractional white noise calculus and applications to finance. *Inf. Dim. Anal. Quant. Probab.*, 6:1–32, 2003.

[61] H. Hult. Approximating some Volterra type stochastic integral with applications to parameter estimation. *Stochastic Processes and their Applications*, 105(1):1–21, 2002.

[62] H. Hult. *Topics on Fractional* B*rownian Motion and Regular Variation for Stochastic Processes.* Doctoral dissertation, Royal Institute of Technology, Stockholm, 2003.

[63] H.E. Hurst. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770–808, 1951.

[64] H.E. Hurst. Methods of using long-term storage in reservoirs. *Proceedings of the Institution of Civil Engineers*, 5:519–590, 1956.

[65] M. Izzeldin and A. Murphy. Bootstrapping the small sample critical values of the rescaled range statistics. *The Economic and Social Review*, 31(4):351–359, 2000.

[66] Gao. J. Modelling long-range dependent Gaussian processes with application in continuous-time financial models. *Journal of Applied Probability*, 41:467–482, 2004.

[67] L.S. Jennings, M.E. Fisher, K.L. Teo, and C.J. Goh. *Miser 3.3-Optimal Control Software: Theory and User Manual.* Department of Mathematics and Statistics, The University of Western Australia, Perth, Australia, 2004.

[68] L.S. Jennings and K.L. Teo. A computational algorithm for functional inequality constrained optimization problems. *Automatica*, 26:371–375, 1990.

[69] K.S. Jomo. Financial liberalization, crises, and Malaysian policy responses. *World Development*, 26:1563–1574, 1998.

[70] R.E. Kalman. A new approach to linear filtering and prediction problems. *Trans. ASME, J. Basic Eng.*, 82:35–45, 1960.

[71] R.E. Kalman and R.S. Bucy. New results in linear filtering and prediction theory. *Trans. ASME, J. Basic Eng.*, 83:95–108, 1961.

[72] E. Kaplan and D. Rodrik. Did the Malaysia capital controls work? NBER Working Paper Series, 2001.

[73] J. Karuppiah and C.A. Los. Wavelet multiresolution analysis of high-frequency Asian fx rates. *Int. Rev. of Financial Analysis*, 14:211–246, 2005.

[74] M.L. Kleptsyna, M.L. Kloeden, P.E. Ahn, and V.V. Ahn. Linear filtering with fractional Brownian motion. *Stochastic Anal. Appl.*, 16:907–914, 1998.

[75] M.L. Kleptsyna and A. Le Breton. Statistical inference for the fractional Ornstein-Uhlenbeck process. *Stat. Inference Stoch. Process.*, 5(3):229–248, 2002.

[76] A.N. Kolmogorov. Wienerische spiralen und einige andere interessante Kurven in Hilbertschen raum. *C.R. (Doklady) Acad. Sci. USSR (N.S.)*, 26:115–118, 1940.

[77] H.L. Koul and D. Surgailis. Asymptotic normality of the Whittle estimator in linear regression models with long memory errors. *Statistical Inference for Stochastic Processes*, 3:129–147, 2000.

[78] A. Kukush, Y. Mishura, and E. Valkeila. Statistical inference with fractional brownian motion. *Statistical Inference for Stochastic Processes*, 8:71–93, 2005.

[79] S.T. Lau, C.T. Lee, and T.H. McInish. Stock returns and beta, firm size, e/p, cf/p, book-to-market, and sales growth: Evidence from Singapore and Malaysia. *Journal of Multinational Financial Management*, 12:207–222, 2002.

[80] A. Le Breton. Filtering and parameter estimation in a simple linear model driven by a fractional Brownian motion. *Statist. Probab. Lett.*, 38:263–24, 1998.

[81] N.N. Leonenko and L.M. Sakhno. On the Whittle estimators for some classes of continuous parameter random processes and fields. *Statistics and Probability Letters*, 76:781–795, 2006.

[82] A. Lindner and R.A. Maller. Lévy integrals and the stationarity of generalised Ornstein-Uhlenbeck processes. *Stoch. Proc. Appl*, 115:1701–1722, 2005.

[83] R. Liu and T. Lux. Long memory in financial time series: Estimation of bivariate multi-fractal model and its application for value-at-risk. Manuscript, University of Kiel, 2005.

[84] P. Llyod and P. Simth. Global economic challenges to ASEAN integration and competitiveness: A prospective look. REPF Project Report No. 03/006a, 2004.

[85] A.W. Lo. Long-term memory in stock market prices. *Econometrica*, 59:1279–1313, 1991.

[86] M.A. Lukas and K.L. Teo. A computational method for a general class of optimal control problems involving integrodifferential equations. *Optimal Control Applications & Methods*, 12:141–162, 1991.

[87] C. H. Luo, C. Y. Wen, J. J. Liaw, S. H. Chiu, and W. M. Lee. Texture characterization of atmospheric fine particles by fractional Brownian motion analysis. *Atmospheric Environment,*, 38:935940, 2004.

[88] C.H. Luo, C.Y. Wen, J.J. Liaw, S.H. Chiu, and W.G. Lee. Texture characterization of atmospheric fine particles by fractional Brownian motion analysis. *Atmospheric Environment*, 38(6):935–940, 2004.

[89] T. Lux. Long-term stochastic dependence in financial prices: evidence from the German stock market. *Applied Economics Letters*, 3:701–706, 1996.

[90] B.B. Mandelbrot. The variation of certain speculative prices. *J. Business*, 36:394–419, 1963.

[91] B.B. Mandelbrot. How long is the coast of Britain? statistical self-similarity and fractional dimension. *Science*, 156:636–638, 1967.

[92] B.B. Mandelbrot. Long-run interdependence in price records and other economic time series. *Econometrica*, 38:122–123, 1970.

[93] B.B. Mandelbrot. Statistical methodology for non-periodic cycles: From the covariance to R/S analysis. *Annals of Economic and Social Measurement*, 1:259–290, 1972.

[94] B.B. Mandelbrot. Limit theorems on the self-normalized range for weakly and strongly dependent processes. *Z. Wahrscheinlichkeitstheorie uenv. Gebiete*, 31:271–285, 1975.

[95] B.B. Mandelbrot. *The fractal geometry of nature*. New York: W.H. Freeman & Co., 1982.

[96] B.B. Mandelbrot. *Fractals and scaling in finance: Discontinuity, concentration, risk*. New York: Springer Verlag, 1997.

[97] B.B. Mandelbrot and R.L. Hudson. *The Misbehavior of Markets*. Basic Books, 2004.

[98] B.B. Mandelbrot and M. Taqqu. Robust R/S analysis of long-run serial correlation. *Bulletin of the International Statistical Institute*, 48(Book 2):59–104, 1979.

[99] B.B. Mandelbrot and J.W. Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Review*, 10:422–437, 1968.

[100] B.B. Mandelbrot and J. Wallis. Noah, Joseph and operational hydrology. *Water Resources Research*, 4:909–918, 1968.

[101] B.B. Mandelbrot and J. Wallis. Computer experiments with fractional Gaussian noises. *Water Resources Research*, 5:228–267, 1969.

[102] B.B. Mandelbrot and J. Wallis. Some long run properties of geophysical records. *Water Resources Research*, 5:967–988, 1969.

[103] B.B. Mandelbrot and J.R. Wallis. Robustness of the rescaled range R/S in the measurement of noncyclic long-run statistical dependence. *Water Resources Research*, 5:967–988, 1969.

[104] H. Masuda. On multidimensional Ornstein-Uhlenbeck processes driven by a general Lévy process. *Bernoulli*, 10:97–120, 2004.

[105] D. McKenzie. Non periodic Australian stock market cycles: Evidence from rescaled range analysis. *Economic Record*, 77:393–406, 2001.

[106] R.C. Merton. *Continuous-time finance*. Basil Blackwell, Inc., 1992.

[107] Y. Mishura. *Stochastic Calculus for Fractional* B*rownian Motion and Related Processes*. Springer, 2008.

[108] J. Moody and W. Lizhong. Price behaviour and Hurst exponents of tick-by-tick interbank foreign exchange rates. In *Proceedings of the IEEE/IAFE 1995 Computational Intelligence for Financial Engineering*, pages 26–30, 1995.

[109] E. Moulines, F. Roueff, and M. Taqqu. A wavelet Whittle estimator of the memory parameter of a non-stationary Gaussian time series. Tech. Rep. Ecole Nationale Superieure des Telecommunications et Boston University, 2005.

[110] C. Muckley. Empirical asset return distributions: Is chaos the culprit? *Applied Economics Letters*, 11(2):81–86, 2004.

[111] S.V. Muniandy, S.C. Lim, and R. Murugan. Inhomogeneous scaling behaviours in Malaysian foreign currency exchange rates. *Physica A*, 301:407–428, 2001.

[112] R.J. Navarro, R. Tamangan, N.G. Natan, E. Ramos, and A. Guzman. The identification of long memory process in the Asean-4 stock markets by fractional and multifractional Brownian motion. *The Philippine Statistician*, 55:65–83, 2006.

[113] D. Nawrocki. R/S analysis and long term dependence in stock market indices. *Managerial Finance*, 21(7):78–91, 1995.

[114] K. Opong, G. Mullholland, A. Fox, and K. Farahmand. The behaviour of some UK equity indices, an application of Hurst and BDS tests. *Journal of Empirical Finance*, 6:267–282, 1999.

[115] M.S. Pan, R.C.W. Fok, and Y.A. Liu. Dynamic linkages between exchange rates and stock prices: Evidence from east Asian markets. *Int. Rev. of Economics and Finance*, 2006.

[116] C.K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, and A.L. Goldberger. Mosaic organization of DNA nucleodites. *Physical Review E*, 49:1685–1689, 1994.

[117] Fischer R. and Akay M. Improved estimators for fractional Brownian motion via the expectation-maximization algorithm. *Medical Engineering and Physics*, 24:77–83, 2002.

[118] G.O. Roberts, O. Papaspiliopoulus, and P. Dellaportas. Bayesian inference for non-Gaussian Ornstein-Uhlenbeck stochastic volatiity processes. *Journal of the Royal Statistical Society, Series B*, 66(2):369–393, 2004.

[119] P.M. Robinson. Gaussian semiparametric estimation of long range dependence. *The Annals of Statistics*, 23:1630–1661, 1995.

[120] P.M. Robinson. *Time Series with Long Memory*. Oxford University Press, 2003.

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

[122] Sheldon M. Ross. *An Introduction to Mathematical Finance: Options and Other Topics*. Cambridge University Press, 1999.

[123] S. Rostek. *Option Pricing in Fractional* B*rownian Markets*. Springer, 2009.

[124] S. Sadique and P. Silvapulle. Long term memory in stock market returns. *Int. J. Fin. Econ.*, 6:59–67, 2001.

[125] P.A. Samuelson. *The Random Character of Stock Market Prices*, chapter Rational Theory of Warrant Pricing, pages 506–532. Cambridge MIT Press, 1964.

[126] W.F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442, 1964.

[127] A.N. Shiryaev. *Essentials of Stochastic Finance. Facts, Models, Theory.* World Scientific, Singapore, 1999.

[128] R.H. Shumway and D.S. Stoffer. *Time Series Analysis and Its Applications.* Springer, 2006.

[129] F.B. Sowell. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics*, 53:165–188, 1992.

[130] D.B. Stephenson, V. Pavan, and R. Bojariu. Is the North Atlantic oscillation a random walk? *Int. J. Clim.*, 20:1–18, 2000.

[131] C.S. Tai. Market integration and currency risk in Asian emerging markets. *Research in International Business and Finance*, 53:165–188, 2007.

[132] Nassim Nicholas Taleb. *The Black Swan.* Random House, 2007.

[133] M. Taqqu, V. Teverosky, and W. Willinger. Estimation for long-range dependence. *Fractals*, 3:785–798, 1995.

[134] M. Taqqu and S. Teverovsky. Robustness of Whittle-type estimators for time series with long range dependence. *Commun. Statist.-Stochastic Models*, 13:723–757, 1997.

[135] K.L. Teo, C.J. Goh, and K.H. Wong. *A Unified Computational Approach to Optimal Control Problems.* Longman Scientific & Technical, 1991.

[136] J.L. Teugels and B. Sundt. *Encyclopedia of Actuarial Science.* Wiley, 2004.

[137] V. Teverosky, M.S. Taqqu, and W. Willinger. A critical look at Lo's modified R/S statistic. *Journal of Statistical Planning and Inference*, 80:211–228, 1999.

[138] B. Tsybakov and N. Georganas. Self-similar processes in communication networks. *IEEE Trans. Inform. Theory*, 44:1713–1725, 1998.

[139] G.E. Uhlenbeck and L.S. Ornstein. On the theory of the Brownian motion. *Phys. Rev*, 36:823–841, 1930.

[140] L. Valdivieso, W. Schoutens, and F. Tuerlinckx. Maximum likelihood estimation in processes of Ornstein-Uhlenbeck type. *Stat. Infer. Stoch. Process*, 12:1–19, 2009.

[141] D. Veitch and P. Abry. A wavelet-based joint estimator of the parameters of long-range dependence. *Information Theory*, 45, 1999.

[142] R. Weron. Estimating long-range dependence: Finance sample properties and confidence intervals. *Physica A*, 312:462–468, 2002.

[143] R. Weron and Przybylowicz. Hurst analysis of electricity price dynamics. *Physica A*, pages 462–468, 2000.

[144] C.Z. Wu, K.L. Teo, Y. Zhao, and W.Y. Yan. An optimal control problem involving impulsive integrodifferential systems. *Optimization Methods and Software*, 22(3):531–549, 2007.

[145] Wei-Guo Zhang, Wei-Lin Xiao, and Chun-Xiong He. Equity warrants pricing model under fractional Brownian motion and an empirical study. *Expert Syst. Appl.*, 36(2):3056–3065, 2009.