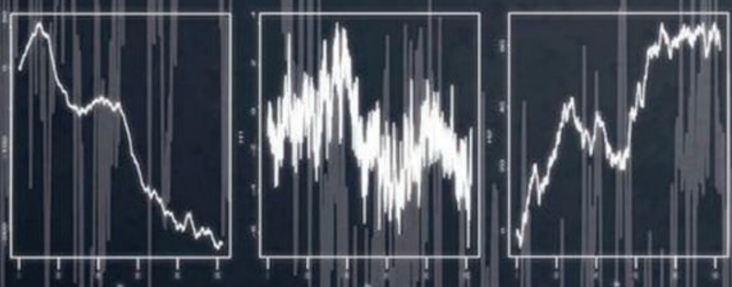
WILEY

Statistical Inference for Fractional Diffusion Processes



B. L. S. Prakasa Rao

WILEY SERIES IN PROBABILITY AND STATISTICS

Statistical Inference for Fractional Diffusion Processes

WILEY SERIES IN PROBABILITY AND STATISTICS

Established by WALTER A. SHEWHART and SAMUEL S. WILKS

A complete list of the titles in this series appears at the end of this volume.

Statistical Inference for Fractional Diffusion Processes

B. L. S. Prakasa Rao

Department of Mathematics and Statistics,
University of Hyderabad, India

WILEY

A John Wiley and Sons, Ltd., Publication

This edition first published 2010

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

Library of Congress Cataloging-in-Publication Data

Prakasa Rao, B. L. S.
Statistical inference for fractional diffusion processes / B.L.S. Prakasa Rao.
p. cm. – (Wiley series in probability and statistics)
Summary: "Statistical Inference for Fractional Diffusion Processes looks at statistical inference for stochastic processes modeled by stochastic differential equations driven by fractional Brownian motion. Other related topics, such as sequential inference, nonparametric and non parametric inference and parametric estimation are also discussed"– Provided by publisher.
Includes bibliographical references and index.
ISBN 978-0-470-66568-8
1. Fractional calculus. 2. Probabilities. I. Title.
QA314.P73 2010
515'.83–dc22

2010010075

A catalogue record for this book is available from the British Library.
ISBN: 978-0-470-66568-8

Typeset in 10/12 Times-Roman by Laserwords Private Limited, Chennai, India
Printed in Singapore by Markono Print Media Pte Ltd.

In memory of my maternal grandfather
Kanchinadham Venkata Subrahmanyam

for teaching me the three 'R's
(Reading, Writing and Arithmetic)
with love and affection

Contents

Preface

In his study of long-term storage capacity and design of reservoirs based on investigations of river water levels along the Nile, Hurst observed a phenomenon which is invariant to changes in scale. Such a scale-invariant phenomenon was also observed in studies of problems connected with traffic patterns of packet flows in high-speed data networks such as the Internet. Mandelbrot introduced a class of processes known as self-similar processes and studied applications of these processes to understand the scale-invariant phenomenon. Long-range dependence is connected with the concept of self-similarity in that the increments of a self-similar process with stationary increments exhibit long-range dependence under some conditions. A long-range dependence pattern is observed in modeling in macroeconomics and finance. Mandelbrot and van Ness introduced the term *fractional Brownian motion* for a Gaussian process with a specific covariance structure and studied its properties. This process is a generalization of classical Brownian motion also known as the Wiener process. Translation of such a process occurs as a limiting process of the log-likelihood ratio in studies on estimation of the location of a cusp of continuous density by Prakasa Rao. Kolmogorov introduced this process in his paper on the Wiener skewline and other interesting curves in Hilbert spaces. Levy discussed the properties of such a process in his book *Processus Stochastiques et Movement Brownien*. Increments of fractional Brownian motion exhibit long-range dependence.

Most of the books dealing with fractional Brownian motion look at the probabilistic properties. We look at the statistical inference for stochastic processes, modeled by stochastic differential equations driven by fractional Brownian motion, which we term as *fractional diffusion processes*. Since fractional Brownian motion is not a semimartingale, it is not possible to extend the notion of the Ito integral for developing stochastic integration for a large class of random processes with fractional Brownian motion as the integrator. Several methods have been developed to overcome this problem. One of them deals with the notion of the Wick product and uses the calculus developed by Malliavin and others. We avoid this approach as it is not in the toolbox of most statisticians. Kleptsyna, Le Breton and their co-workers introduced another method by using the notion of fundamental martingale associated with fractional Brownian motion. This method turns out to be very useful in the context of statistical inference for fractional diffusion processes. Our aim in this book is to consider parametric and nonparametric inference problems for fractional diffusion processes when a complete path of the process over a finite interval is observable. There is no significant work in the area of statistical inference for fractional diffusion processes when discrete data or sampled data from the process is only available.

It is a pleasure to thank Professor V. Kannan and his colleagues in the Department of Mathematics and Statistics, University of Hyderabad, for inviting me to visit their university after I retired from the Indian Statistical Institute and for providing me with excellent facilities for continuing my research work during the last five years leading to this book. Professor M. N. Mishra, presently with the Institute of Mathematics and Applications at Bhuvaneswar, collaborated with me during the last several years in my work on inference for stochastic processes. I am happy to acknowledge the same. Figures on the cover page were reproduced with the permission of Professor Ton Dieker from his Master's thesis "Simulation of fractional Brownian motion". Thanks are due to him.

Thanks are due to my wife Vasanta Bhagavatula for her unstinted support in all of my academic pursuits.

B. L. S. Prakasa Rao
Hyderabad, India
January 18, 2010

1

Fractional Brownian motion and related processes

1.1 Introduction

In his study of long-term storage capacity and design of reservoirs based on investigations of river water levels along the Nile, Hurst (1951) observed a phenomenon which is invariant to changes in scale. Such a scale-invariant phenomenon was recently observed in problems connected with traffic patterns of packet flows in high-speed data networks such as the Internet (cf. Leland *et al*. (1994), Willinger *et al*. (2003), Norros (2003)) and in financial data (Willinger *et al*. (1999)). Lamperti (1962) introduced a class of stochastic processes known as semi-stable processes with the property that, if an infinite sequence of contractions of the time and space scales of the process yield a limiting process, then the limiting process is semi-stable. Mandelbrot (1982) termed these processes as self-similar and studied applications of these models to understand scale-invariant phenomena. Long-range dependence is related to the concept of self-similarity for a stochastic process in that the increments of a self-similar process with stationary increments exhibit long-range dependence under some conditions. A long-range dependence pattern is also observed in macroeconomics and finance (cf. Henry and Zafforoni (2003)). A fairly recent monograph by Doukhan *et al*. (2003) discusses the theory and applications of long-range dependence. Before we discuss modeling of processes with long-range dependence, let us look at the consequences of long-range dependence phenomena. A brief survey of self-similar processes, fractional Brownian motion and statistical inference is given in Prakasa Rao (2004d).

Suppose $\{X_i, 1 \leq i \leq n\}$ are independent and identically distributed (i.i.d.) random variables with mean μ and variance σ^2. It is well known that the sample mean $\bar{X}_n = n^{-1} \sum_{i=1}^{n} X_i$ is an unbiased estimator of the mean μ and its variance is σ^2/n which is proportional to n^{-1}. In his work on yields of agricultural experiments, Smith (1938) studied mean yield $\bar{X}_n$ as a function of the number n of plots and observed that $Var(\bar{X}_n)$ is proportional to n^{-a} where $0 < a < 1$. If $a = 0.4$, then approximately 100 000 observations are needed to achieve the same accuracy of $\bar{X}_n$ as from 100 i.i.d. observations. In other words, the presence of long-range dependence plays a major role in estimation and prediction problems.

Long-range dependence phenomena are said to occur in a stationary time series $\{X_n, n \geq 0\}$ if $Cov(X_0, X_n)$ of the time series tends to zero as $n \to \infty$ and yet the condition

$$\sum_{n=0}^{\infty} |Cov(X_0, X_n)| = \infty \tag{1.1}$$

holds. In other words, the covariance between X_0 and X_n tends to zero as $n \to \infty$ but so slowly that their sums diverge.

1.2 Self-similar processes

A real-valued stochastic process $Z = \{Z(t), -\infty < t < \infty\}$ is said to be *self-similar* with index $H > 0$ if, for any $a > 0$,

$$\mathcal{L}(\{Z(at), -\infty < t < \infty\}) = \mathcal{L}(\{a^H Z(t), -\infty < t < \infty\}) \tag{1.2}$$

where $\mathcal{L}$ denotes the class of all finite-dimensional distributions and the equality indicates the equality of the finite-dimensional distributions of the process on the right of Equation (1.2) with the corresponding finite-dimensional distributions of the process on the left of Equation (1.2). The index H is called the *scaling exponent* or the *fractal index* or the *Hurst index* of the process. If H is the scaling exponent of a self-similar process Z, then the process Z is called an H-*self-similar* process or H-ss process for short. A process Z is said to be degenerate if $P(Z(t) = 0) = 1$ for all $t \in R$. Hereafter, we write $X \stackrel{\Delta}{=} Y$ to indicate that the random variables X and Y have the same probability distribution.

Proposition 1.1: A non-degenerate H-ss process Z cannot be a stationary process.

Proof: Suppose the process Z is a stationary process. Since the process Z is non degenerate, there exists $t_0 \in R$ such that $Z(t_0) \neq 0$ with positive probability and, for all $a > 0$,

$$Z(t_0) \stackrel{\Delta}{=} Z(at_0) \stackrel{\Delta}{=} a^H Z(t_0)$$

by stationarity and self-similarity of the process Z. Let $a \to \infty$. Then the family of random variables on the right diverge with positive probability, whereas the random variable on the left is finite with probability one, leading to a contradiction. Hence the process Z is not a stationary process.

Proposition 1.2: Suppose that $\{Z(t), -\infty < t < \infty\}$ is an H-ss process. Define

$$Y(t) = e^{-tH} Z(e^t), \; -\infty < t < \infty. \tag{1.3}$$

Then the process $\{Y(t), -\infty < t < \infty\}$ is a stationary process.

Proof: Let $k \geq 1$ be an integer. For a_i, $1 \leq j \leq k$ real and $h \in R$,

$$\begin{aligned}
\sum_{j=1}^{k} a_j Y(t_j + h) &\stackrel{\Delta}{=} \sum_{j=1}^{k} a_j e^{-(t_j+h)H} Z(e^{t_j+h}) \\
&\stackrel{\Delta}{=} \sum_{j=1}^{k} a_j e^{-(t_j+h)H} e^{hH} Z(e^{t_j}) \\
&\quad \text{(by self-similarity of the process } Z) \\
&\stackrel{\Delta}{=} \sum_{j=1}^{k} a_j e^{-t_j H} Z(e^{t_j}) \\
&\stackrel{\Delta}{=} \sum_{j=1}^{k} a_j Y(t_j).
\end{aligned} \tag{1.4}$$

Since the above relation holds for every $(a_1, \ldots, a_k) \in R^k$, an application of the Cramér–Wold theorem shows that the finite-dimensional distribution of the random vector $(Y(t_1 + h), \ldots, Y(t_k + h))$ is the same as that of the random vector $(Y(t_1), \ldots, Y(t_k))$. Since this property holds for all t_i, $1 \leq i \leq k$, $k \geq 1$ and for all h real, it follows that the process $Y = \{Y(t), -\infty < t < \infty\}$ is a stationary process.

The transformation defined by (1.3) is called the *Lamperti transformation*. By retracing the arguments given in the proof of Proposition 1.2, the following result can be proved.

Proposition 1.3: Suppose $\{Y(t), -\infty < t < \infty\}$ is a stationary process. Let $X(t) = t^H Y(\log t)$ for $t > 0$. Then $\{X(t), t > 0\}$ is an H-ss process.

Proposition 1.4: Suppose that a process $\{Z(t), -\infty < t < \infty\}$ is a second-order process, that is, $E[Z^2(t)] < \infty$ for all $t \in R$, and it is an H-ss process with stationary increments, that is,

$$Z(t + h) - Z(t) \stackrel{\Delta}{=} Z(h) - Z(0)$$

for $t \in R, h \in R$. Let $\sigma^2 = Var(Z(1))$. Then the following properties hold:

(i) $Z(0) = 0$ a.s.

(ii) If $H \neq 1$, then $E(Z(t)) = 0, -\infty < t < \infty$.

(iii) $Z(-t) \stackrel{\Delta}{=} -Z(t), -\infty < t < \infty$.

(iv) $E(Z^2(t)) = |t|^{2H} E(Z^2(1)), -\infty < t < \infty$.

(v) Suppose $H \neq 1$. Then

$$Cov(Z(t), Z(s)) = (\sigma^2/2)\{|t|^{2H} + |s|^{2H} - |t-s|^{2H}\}.$$

(vi) $0 < H \leq 1$.

(vii) If $H = 1$, then $Z(t) \stackrel{\Delta}{=} tZ(1), -\infty < t < \infty$.

Proof:

(i) Note that $Z(0) = Z(a.0) \stackrel{\Delta}{=} a^H Z(0)$ for any $a > 0$ by the self-similarity of the process Z. It is easy to see that this relation holds only if $Z(0) = 0$ a.s.

(ii) Suppose $H \neq 1$. Since $Z(2t) \stackrel{\Delta}{=} 2^H Z(t)$, it follows that

$$\begin{aligned} 2^H E(Z(t)) &= E(Z(2t)) \\ &= E(Z(2t) - Z(t)) + E(Z(t)) \\ &= E(Z(t) - Z(0)) + E(Z(t)) \\ &\quad \text{(by stationarity of the increments)} \\ &= 2\, E(Z(t)) \end{aligned} \tag{1.5}$$

for any $t \in R$. The last equality follows from the observation that $Z(0) = 0$ a.s. from (i). Hence $E(Z(t)) = 0$ since $H \neq 1$.

(iii) Observe that, for any $t \in R$,

$$\begin{aligned} Z(-t) &\stackrel{\Delta}{=} Z(-t) - Z(0) \\ &\stackrel{\Delta}{=} Z(0) - Z(t) \ \text{(by stationarity of the increments)} \\ &\stackrel{\Delta}{=} -Z(t) \ \text{(by Property (i))}. \end{aligned} \tag{1.6}$$

Therefore $Z(-t) \stackrel{\Delta}{=} -Z(t)$ for every $t \in R$.

(iv) It is easy to see that, for any $t \in R$,

$$\begin{aligned} E(Z^2(t)) &= E(Z^2(|t| \ sgn\ t)) \\ &= |t|^{2H} E(Z^2(sgn\ t)) \ \text{(by self-similarity)} \\ &= |t|^{2H} E(Z^2(1)) \ \text{(by Property (iii))} \\ &= \sigma^2 |t|^{2H} \ \text{(by Property (ii))}. \end{aligned} \tag{1.7}$$

Here the function $sgn\ t$ is equal to 1 if $t \geq 0$ and is equal to -1 if $t < 0$. If $\sigma^2 = 1$, the process Z is called a standard H-ss process with stationary increments.

(v) Let $R_H(t, s)$ be the covariance between $Z(t)$ and $Z(s)$ for any $-\infty < t$, $s < \infty$. Then

$$\begin{aligned} R_H(t,s) &\equiv Cov(Z(t), Z(s)) \\ &= E[Z(t)Z(s)] \quad \text{(by Property (ii))} \\ &= \frac{1}{2}\{E(Z^2(t)) + E(Z^2(s)) - E[(Z(t) - Z(s))^2]\} \\ &= \frac{1}{2}\{E(Z^2(t)) + E(Z^2(s)) - E[(Z(t-s) - Z(0))^2]\} \\ &\quad \text{(by stationarity of the increments)} \\ &= \frac{1}{2}\{E(Z^2(t)) + E(Z^2(s)) - E(Z^2(t-s))\} \quad \text{(by Property (ii))} \\ &= \frac{\sigma^2}{2}\{|t|^{2H} + |s|^{2H} - |t-s|^{2H}\} \quad \text{(by Property (iv)).} \end{aligned} \tag{1.8}$$

In particular, it follows that the function $R_H(t, s)$ is nonnegative definite as it is the covariance function of a stochastic process.

(vi) Note that

$$\begin{aligned} E(|Z(2)|) &= E(|Z(2) - Z(1) + Z(1)|) \\ &\leq E(|Z(2) - Z(1)|) + E(|Z(1)|) \\ &= E(|Z(1) - Z(0)|) + E(|Z(1)|) \\ &\quad \text{(by stationarity of the increments)} \\ &= 2E(|Z(1)|) \quad \text{(by Property (i)).} \end{aligned} \tag{1.9}$$

Self-similarity of the process Z implies that

$$E(|Z(2)|) = 2^H E(|Z(1)|). \tag{1.10}$$

Combining relations (1.9) and (1.10), we get

$$2^H E(|Z(1)|) \leq 2E(|Z(1)|)$$

which, in turn, implies that $H \leq 1$ since the process Z is a non degenerate process.

(vii) Suppose $H = 1$. Then $E(Z(t)Z(1)) = tE(Z^2(1))$ and $E(Z^2(t)) = t^2E(Z^2(1))$ by the self-similarity of the process Z. Hence

$$\begin{aligned} E(Z(t) - tZ(1))^2 &= E(Z^2(t)) - 2t\ E(Z(t)Z(1)) + t^2E(Z^2(1)) \\ &= (t^2 - 2t^2 + t^2)E(Z^2(1)) \\ &= 0. \end{aligned} \tag{1.11}$$

This relation shows that $Z(t) = tZ(1)$ a.s. for every $t \in R$.

Remarks: As was mentioned earlier, self-similar processes have been used for stochastic modeling in such diverse areas as hydrology (cf. Montanari (2003)), geophysics, medicine, genetics and financial economics (Willinger *et al*. (1999)) and more recently in modeling Internet traffic patterns (Leland *et al*. (1994)). Additional applications are given in Goldberger and West (1987), Stewart *et al*. (1993), Buldyrev *et al*. (1993), Ossandik *et al*. (1994), Percival and Guttorp (1994) and Peng *et al*. (1992, 1995a,b). It is important to estimate the Hurst index H for modeling purposes. This problem has been considered by Azais (1990), Geweke and Porter-Hudak (1983), Taylor and Taylor (1991), Beran and Terrin (1994), Constantine and Hall (1994), Feuverger *et al*. (1994), Chen *et al*. (1995), Robinson (1995), Abry and Sellan (1996), Comte (1996), McCoy and Walden (1996), Hall *et al*. (1997), Kent and Wood (1997), and more recently in Jensen (1998), Poggi and Viano (1998), and Coeurjolly (2001).

It was observed that there are some phenomena which exhibit self-similar behavior locally but the nature of self-similarity changes as the phenomenon evolves. It was suggested that the parameter H must be allowed to vary as a function of time for modeling such data. Goncalves and Flandrin (1993) and Flandrin and Goncalves (1994) propose a class of processes which are called *locally self-similar* with dependent scaling exponents and discuss their applications. Wang *et al*. (2001) develop procedures using wavelets to construct local estimates for time-varying scaling exponent $H(t)$ of a locally self-similar process.

A second-order stochastic process $\{Z(t), t > 0\}$ is called *wide-sense H-self-similar* if it satisfies the following conditions for every $a > 0$:

(i) $E(Z(at)) = a^H E(Z(t)), t > 0$;

(ii) $E(Z(at)Z(as)) = a^{2H} E(Z(t)Z(s)), t > 0, s > 0$.

This definition can be compared with the definition of (strict) H-self-similarity which is that the processes $\{Z(at)\}$ and $\{a^H Z(t)\}$ have the same finite-dimensional distributions for every $a > 0$. The wide-sense definition is more general. However, it excludes self-similar processes with infinite second moments such as non-Gaussian stable processes. Given a wide-sense H-ss process Z, it is possible to form a wide-sense stationary process Y via the Lamperti transformation

$$Y(t) = e^{-Ht} Z(e^t).$$

The Lamperti transformation helps in using the techniques developed for the study of wide-sense stationary processes in wide-sense self-similar processes. Yazici and Kashyap (1997) introduced the concept of wide-sense H-ss processes. Nuzman and Poor (2000, 2001) discuss linear estimation of self-similar processes via the Lamperti transformation and generalize reproducing kernel Hilbert space methods for wide-sense self-similar processes. These results were applied to solve linear problems including projection, polynomial signal detection and polynomial amplitude estimation for general wide-sense self-similar processes.

1.3 Fractional Brownian motion

A Gaussian H-ss process $W^H = \{W_H(t), -\infty < t < \infty\}$ with stationary increments and with fractal index $0 < H < 1$ is termed *fractional Brownian motion* (fBm). Note that $E[W^H(t)] = 0, -\infty < t < \infty$. It is said to be standard if $Var(W^H(1)) = 1$.

For standard fractional Brownian motion,

$$Cov(W^H(t), W^H(s)) = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H}).$$

If $H = \frac{1}{2}$, then fBm reduces to the Brownian motion known as the Wiener process. It is easy to see that if $\{X(t), -\infty < t < \infty\}$ is a Gaussian process with stationary increments with mean zero, with $X(0) = 0$ and $E(X^2(t)) = \sigma^2|t|^{2H}$ for some $0 < \sigma < \infty$ and $0 < H < 1$, then the process $\{X(t), -\infty < t < \infty\}$ is fBm. The following theorem gives some properties of standard fBm.

Theorem 1.5: Let $\{W^H(t), -\infty < t < \infty\}$ be standard fBm with Hurst index H for some $0 < H < 1$. Then:

(i) There exists a version of the process $\{W^H(t), -\infty < t < \infty\}$ such that the sample paths of the process are continuous with probability one.

(ii) The sample paths of the process $\{W^H(t), -\infty < t < \infty\}$ are nowhere differentiable in the L^2-sense.

(iii) For any $0 < \lambda < H$, there exist constants $h > 0$ and $C > 0$ such that, with probability one,

$$|W^H(t) - W^H(s)| < C|t-s|^{\lambda},\ 0 \le t, s \le 1, |t-s| \le h.$$

(iv) Consider the standard fBm $W^H = \{W^H(t), 0 \le t \le T\}$ with Hurst index H. Then

$$\begin{aligned}\lim_{n\to\infty} \sum_{j=0}^{2^n-1} \left| W^H\left(\frac{j+1}{2^n}T\right) - W^H\left(\frac{j}{2^n}T\right)\right|^p &= 0 \text{ a.s. if } pH > 1 \\ &= \infty \text{ a.s. if } pH < 1 \\ &= T \text{ a.s. if } pH = 1. \quad (1.12)\end{aligned}$$

Property (i) stated above follows from Kolmogorov's sufficient condition for a.s. continuity of the sample paths of a stochastic process and the fact that

$$E[|W^H(t) - W^H(s)|^\alpha] = E[|W^H(1)|^\alpha]|t - s|^{\alpha H}, \ -\infty < t, s < \infty$$

for any $\alpha > 0$. The equation given above follows from the observation that fBm is an H-ss process with stationary increments. The constant $\alpha > 0$ can be chosen so that $\alpha H > 1$ satisfies Kolmogorov's continuity condition.

Property (ii) is a consequence of the relation

$$E\left[\left|\frac{W^H(t) - W^H(s)}{t - s}\right|^2\right] = E[|W^H(1)|^2]|t - s|^{2H-2}$$

and the last term tends to infinity as $t \to s$ since $H < 1$. Hence the paths of fBm are not L^2-differentiable.

For a discussion and proofs of Properties (iii) and (iv), see Doukhan *et al.* (2003) and Decreusefond and Ustunel (1999). If the limit

$$\lim_{n\to\infty} \sum_{j=0}^{2^n-1} \left|W^H\left(\frac{j+1}{2^n}T\right) - W^H\left(\frac{j}{2^n}T\right)\right|^p$$

exists a.s., then the limit is called the p-th variation of the process W^H over the interval $[0, T]$. If $p = 2$, then it is called the quadratic variation over the interval $[0, T]$. If $H = \frac{1}{2}$ and $p = 2$, in (iv), then the process W^H reduces to the standard Brownian motion W and we have the well-known result

$$\lim_{n\to\infty} \sum_{j=0}^{2^n-1} \left|W\left(\frac{j+1}{2^n}T\right) - W^H\left(\frac{j}{2^n}T\right)\right|^2 = T \ \text{ a.s.}$$

for the quadratic variation of the standard Brownian motion on the interval $[0, T]$. If $H < \frac{1}{2}$, then, for $p = 2$, we have $pH < 1$ and the process has infinite quadratic variation by Property (iv). If $H > \frac{1}{2}$, then, for $p = 2$, we have $pH > 1$ and the process has zero quadratic variation by Property (iv). Such a process is called a *Dirichlet process*. Furthermore, the process W^H has finite p-th variation if $p = 1/H$. In other words,

$$\lim_{n\to\infty} \sum_{j=0}^{2^n-1} \left|W^H\left(\frac{j+1}{2^n}T\right) - W^H\left(\frac{j}{2^n}T\right)\right|^{1/H} = T \ \text{ a.s.} \qquad (1.13)$$

Let us again consider standard fBm W^H with Hurst index $H > \frac{1}{2}$ over an interval $[0, T]$. Let $0 = t_0^{(n)} < t_1^{(n)} < \cdots < t_n^{(n)} = T$ be a subdivision of the interval $[0, T]$ such that

$$\max_{0\le j\le n-1} |t_{j+1}^{(n)} - t_j^{(n)}| \to 0$$

as $n \to \infty$. Note that

$$E[(W^H(t) - W^H(s))^2] = |t-s|^{2H}, 0 \le t, s \le T$$

and hence

$$\begin{aligned}\sum_{j=0}^{n-1} E[(W^H(t_{j+1}^{(n)}) - W^H(t_j^{(n)}))^2] &= \sum_{j=0}^{n-1} [t_{j+1}^{(n)} - t_j^{(n)}]^{2H} \\ &\le \max_{0 \le j \le n-1} [t_{j+1}^{(n)} - t_j^{(n)}]^{2H-1} T \\ &= \left(\max_{0 \le j \le n-1} [t_{j+1}^{(n)} - t_j^{(n)}] \right)^{2H-1} T. \quad (1.14)\end{aligned}$$

Therefore

$$\lim_{n \to \infty} \sum_{j=0}^{n-1} E[(W^H(t_{j+1}^{(n)}) - W^H(t_j^{(n)}))^2] = 0$$

or equivalently

$$\lim_{n \to \infty} E\left[\sum_{j=0}^{n-1} (W^H(t_{j+1}^{(n)}) - W^H(t_j^{(n)}))^2 \right] = 0.$$

This statement in turn implies that

$$\sum_{j=0}^{n-1} (W^H(t_{j+1}^{(n)}) - W^H(t_j^{(n)}))^2 \xrightarrow{p} 0 \text{ as } n \to \infty.$$

As a consequence of this fact, it can be shown that the process W^H is not a semimartingale from the results in Liptser and Shiryayev (1989). For the definition of a semimartingale and its properties, see Prakasa Rao (1999b).

Representation of fBm as a stochastic integral

Suppose $\{W(t), -\infty < t < \infty\}$ is standard Brownian motion and $H \in \left(\frac{1}{2}, 1\right)$. Define a process $\{Z(t), -\infty < t < \infty\}$ with $Z(0) = 0$ by the relation

$$\begin{aligned}Z(t) - Z(s) &= \lim_{a \to -\infty} \left(c_H \int_a^t (t-\tau)^{H-\frac{1}{2}} dW_\tau - c_H \int_a^s (s-\tau)^{H-\frac{1}{2}} dW_\tau \right) \\ &= c_H \int_s^t (t-\tau)^{H-\frac{1}{2}} dW_\tau + c_H \int_{-\infty}^s [(t-\tau)^{H-\frac{1}{2}} - (s-\tau)^{H-\frac{1}{2}}] dW_\tau \quad (1.15)\end{aligned}$$

for $t > s$, where

$$c_H = \left(2H\,\Gamma\left(\frac{3}{2} - H\right) / \Gamma\left(\frac{1}{2} + H\right)\Gamma(2 - 2H)\right)^{1/2} \tag{1.16}$$

and $\Gamma(.)$ is the gamma function. Here the integrals are defined as Wiener integrals and the resulting process Z is a mean zero Gaussian process. In order to show that the process Z is in fact fBm, we have to prove that its covariance function is that of fBm. We will come back to this discussion later in this section.

Integration with respect to fBm

It is known that, in order to develop the theory of stochastic integration of a random process with respect to another stochastic process satisfying the usual properties of integrals such as linearity and dominated convergence theorem, it is necessary for the integrator to be a semimartingale. This can be seen from Theorem VIII.80 in Dellacherie and Meyer (1982). Semimartingales can also be characterized by this property. Since fBm is not a semimartingale, it is not possible to define stochastic integration of a random process with respect to fBm starting with the usual method of limiting arguments based on Riemann-type sums for simple functions as in the case of Ito integrals. However, the special case of a stochastic integration of a deterministic integrand with respect to fBm as the integrator can be developed using the theory of integration with respect to general Gaussian processes as given, say, in Huang and Cambanis (1978) and more recently in Alos *et al*. (2001). There are other methods of developing stochastic integration of a random process with respect to fBm using the notion of Wick product and applying the techniques of Malliavin calculus. We do not use these approaches throughout this book and unless specified otherwise, we consider fBm with Hurst index $H > \frac{1}{2}$ throughout this book. The reason for such a choice of H for modeling purposes will become clear from our discussion later in this section.

Let $\{Z(t), -\infty < t < \infty\}$ be standard fBm with Hurst index $H > \frac{1}{2}$ and suppose $\{Y(t), -\infty < t < \infty\}$ is a simple process in the sense that $Y(t) = \sum_{j=1}^{k} X_j I_{(T_{j-1},T_j]}(t)$ where $-\infty < T_0 < T_1 < \cdots < T_k < \infty$. We define the stochastic integral of the process Y with respect to Z by the relation

$$\int_{-\infty}^{\infty} Y(t)dZ(t) = \sum_{j=1}^{k} X_j(Z(T_j) - Z(T_{j-1})). \tag{1.17}$$

If the process Y is of locally bounded variation, then we can define the integral by using the integration by parts formula

$$\int_a^b Y(t)dZ(t) = Y(b)Z(b) - Y(a)Z(a) - \int_a^b Z(t)dY(t) \tag{1.18}$$

and the integral on the right of Equation (1.18) can be defined using the theory of Lebesgue–Stieltjes integration. Suppose the process Y is non random, that is, deterministic. Under suitable conditions on the non random function Y, the integral on the left of (1.18) can be defined as an L^2-limit of Riemann sums of the type defined in (1.17) with nonrandom sequence T_j. Gripenberg and Norros (1996) give an example of a random process Y illustrating the problem of non continuity in extending the above method of stochastic integration with respect to fBm for random integrands. We will consider integrals of non random functions only with integrator as fBm throughout this book unless otherwise stated.

An alternate way of defining the stochastic integral of a non random function f with respect to fBm Z is by the formula

$$\int_{-\infty}^{\infty} f(t)dZ(t)$$

$$= c_H\left(H-\frac{1}{2}\right)\int_{-\infty}^{\infty}\left(\int_{\tau}^{\infty}(t-\tau)^{H-\frac{3}{2}}f(t)dt\right)dW_\tau \tag{1.19}$$

where W is a standard Wiener process and the constant c_H is as defined in (1.16). The integral defined on the right of (1.19) exists provided the function

$$\int_{\tau}^{\infty}(t-\tau)^{H-\frac{3}{2}}f(t)dt$$

as a function of τ is square integrable. A sufficient condition for this to hold is that $f \in L^2(R)\cap L^1(R)$. It is easy to see that the random variable

$$\int_{-\infty}^{\infty} f(t)dZ(t)$$

is Gaussian and

$$E\left(\int_{-\infty}^{\infty} f(t)dZ(t)\right) = 0$$

whenever $f \in L^2(R)\cap L^1(R)$. We now obtain the covariance formula for two such integrals.

Theorem 1.6: For functions $f, g \in L^2(R)\cap L^1(R)$,

$$E\left(\int_{-\infty}^{\infty} f(t)dZ(t)\int_{-\infty}^{\infty} g(t)dZ(t)\right)$$

$$= H(2H-1)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t)g(s)|t-s|^{2H-2}dtds. \tag{1.20}$$

Proof: Note that

$$
\begin{aligned}
&E\left(\int_{-\infty}^{\infty} f(t)dZ(t) \int_{-\infty}^{\infty} g(t)dZ(t)\right)\\
&= c_H^2\left(H-\frac{1}{2}\right)^2 E\left[\left(\int_{-\infty}^{\infty}\left(\int_{\tau}^{\infty}(t-\tau)^{H-\frac{3}{2}} f(t)dt\right)dW_\tau\right)\right.\\
&\qquad\qquad \left.\left(\int_{-\infty}^{\infty}\left(\int_{\tau}^{\infty}(t-\tau)^{H-\frac{3}{2}} g(t)dt\right)dW_\tau\right)\right]\\
&= c_H^2\left(H-\frac{1}{2}\right)^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t)g(s)\left[\int_{-\infty}^{\min(s,t)}(s-\tau)^{H-\frac{3}{2}}(t-\tau)^{H-\frac{3}{2}}d\tau\right]dtds
\end{aligned}
$$

by the properties of Wiener integrals. From the results in Abramowitz and Stegun (1970), 6.2.1, 6.2.2, it follows that

$$
\begin{aligned}
\int_{-\infty}^{\min(s,t)}(s-\tau)^{H-\frac{3}{2}}(t-\tau)^{H-\frac{3}{2}}d\tau &= \int_0^{\infty}(|t-s|+\tau)^{H-\frac{3}{2}}\tau^{H-\frac{3}{2}}d\tau\\
&= |t-s|^{2H-2}\int_0^{\infty}(1+\tau)^{H-\frac{3}{2}}\tau^{H-\frac{3}{2}}d\tau\\
&= |t-s|^{2H-2}\frac{\Gamma(H-\frac{1}{2})\Gamma(2-2H)}{\Gamma(\frac{3}{2}-H)}. \qquad (1.21)
\end{aligned}
$$

From the equations derived above, it follows that

$$
\begin{aligned}
&E\left(\int_{-\infty}^{\infty} f(t)dZ(t)\int_{-\infty}^{\infty} g(t)dZ(t)\right)\\
&= H(2H-1)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t)g(s)|t-s|^{2H-2}dtds. \qquad (1.22)
\end{aligned}
$$

As a consequence of Theorem 1.6, we obtain that

$$
E\left[\int_{-\infty}^{\infty} f(t)dZ(t)\right]^2 = H(2H-1)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t)f(s)|t-s|^{2H-2}dtds. \qquad (1.23)
$$

Our discussion of integration with respect to fBm given here is based on Gripenberg and Norros (1996). Another way of defining the integral of a non random function with respect to fBm is given later in this section. Zahle (1998) has defined path wise integration with respect to fBm when the integrand is possibly random using the methods of fractional calculus. We will briefly review this approach in the last section of this chapter.

Inequalities for moments of integrals with respect to fBm

We assume as before that $H > \frac{1}{2}$. Let Ψ denote the integral operator

$$\Psi f(t) = H(2H-1)\int_0^\infty f(s)|t-s|^{2H-2}ds \tag{1.24}$$

and define the inner product

$$\langle f, g\rangle_\Psi = \langle f, \Psi g\rangle = H(2H-1)\int_0^\infty \int_0^\infty f(t)g(s)|t-s|^{2H-2}dtds \tag{1.25}$$

where $\langle ., .\rangle$ denote the usual inner product in $L^2([0,\infty))$. Let L^2_Ψ be the space of functions f such that $\langle f, f\rangle_\Psi < \infty$. Let $L^2_\Psi([0,T])$ be the space of functions f such that $\langle f I_{[0,T]}, f I_{[0,T]}\rangle_\Psi < \infty$. Here I_A denotes the indicator function of the set A. The mapping $Z(t) \to I_{[0,t]}$ can be extended to an isometry between the Gaussian subspace of the space generated by the random variables $Z(t), t \geq 0$, and the function space L^2_Ψ, as well as to an isometry between a subspace of the space generated by the random variables $Z(t), 0 \leq t \leq T$, and the function space $L^2_\Psi([0,T])$ (cf. Huang and Cambanis (1978)). For $f \in L^2_\Psi$, the integral

$$\int_0^\infty f(t)dZ(t)$$

is defined as the image of f by this isometry. In particular, for $f, g \in L^2_\Psi([0,T])$,

$$E\left(\int_0^T f(t)dZ(t)\int_0^T g(t)dZ(t)\right) = H(2H-1)\int_0^T\int_0^T f(t)g(s)|t-s|^{2H-2}dtds \tag{1.26}$$

and

$$E\left(\int_u^v f(t)dZ(t)\right)^2 = H(2H-1)\int_u^v\int_u^v f(t)f(s)|t-s|^{2H-2}dtds. \tag{1.27}$$

Let

$$||f||_{L^p((u,v))} = \left(\int_u^v |f(t)|^p\right)^{1/p} \tag{1.28}$$

for $p \geq 1$ and $0 \leq u < v \leq \infty$.

Theorem 1.7: Let Z be standard fBm with Hurst index $H > \frac{1}{2}$. Then, for every $r > 0$, there exists a positive constant $c(H,r)$ such that for every $0 \leq u < v < \infty$,

$$E\left|\int_u^v f(t)dZ(t)\right|^r \leq c(H,r)||f||^r_{L^{1/H}((u,v))} \tag{1.29}$$

and

$$E\left|\int_u^v f(t)dZ(t)\int_u^v g(t)dZ(t)\right|^r \leq c(H,r)||f||^r_{L^{1/H}((u,v))}||g||^r_{L^{1/H}((u,v))}. \tag{1.30}$$

Proof: Since the random variable

$$\int_0^T f(t)dZ(t)$$

is a mean zero Gaussian random variable, for every $r > 0$, there exists a constant $k(r) > 0$ such that

$$\begin{aligned}E\left|\int_0^T f(t)dZ(t)\right|^r &\leq k(r)\left(E\left|\int_0^T f(t)dZ(t)\right|^2\right)^{r/2}\\ &= k(r)\left[H(2H-1)\int_0^T\int_0^T f(t)f(s)|t-s|^{2H-2}dtds\right]^{r/2}.\end{aligned} \tag{1.31}$$

Furthermore

$$\begin{aligned}&\int_0^T\int_0^T |f(t)f(s)||t-s|^{2H-2}dtds\\ &= \int_0^T |f(t)|\left(\int_0^T |f(s)||t-s|^{2H-2}ds\right)dt\\ &\leq \left(\int_0^T |f(u)|^{1/H}du\right)^H\left(\int_0^T du\left(\int_0^T |f(v)||u-v|^{2H-2}dv\right)^{1/(1-H)}\right)^{1-H}\\ &\qquad \text{(by Holder's inequality)}\\ &\leq A\left(\frac{1}{H},\frac{1}{1-H}\right)\left(\int_0^T |f(u)|^{1/H}du\right)^{2H}\end{aligned} \tag{1.32}$$

for some positive constant $A(1/H, 1/1-H)$. The last inequality follows from the Hardy–Littlewood inequality (cf. Stein (1971), p. 119) stated below in Proposition 1.8. Combining the inequalities in (1.31) and (1.32), the inequality (1.29) stated in Theorem 1.7 is proved.

Proposition 1.8: Let $0 < \alpha < 1$ and $1 < p < q < \infty$, $1/q = (1/p) - \alpha$. Suppose $f \in L^p((0,\infty))$. Define

$$I^\alpha f(x) = \frac{1}{\Gamma(\alpha)}\int_0^\infty |x-y|^{\alpha-1}f(y)dy. \tag{1.33}$$

Then there exists a positive constant $A(p, q)$ such that

$$||I^{\alpha} f||_{L^q((0,\infty))} \leq A(p, q)||f||_{L^p((0,\infty))}. \tag{1.34}$$

We do not prove this inequality. For details, see Stein (1971), pp. 117–120.

The results stated above can be reformulated in the following manner. Let $\{W_t, t \geq 0\}$ be standard fBm with Hurst index $H > \frac{1}{2}$. Suppose a function $f : [0, T] \to R$ satisfies the condition

$$||f||_{L^{1/H}([0,T])} = \left(\int_0^T |f(s)|^{1/H} ds\right)^H < \infty. \tag{1.35}$$

Then the Wiener integral

$$Y_t^H = \int_0^t f(s) dW_s^H, \ 0 \leq t \leq T$$

exists and is a mean zero Gaussian process, and for every $r > 0$, there exists a positive constant $c(r, H)$ such that

$$E|Y_{t_2}^H - Y_{t_1}^H|^r \leq c(r, H) \left(\int_{t_1}^{t_2} |f(s)|^{1/H} ds\right)^{rH}, \ 0 \leq t_1 \leq t_2 \leq T.$$

The inequalities proved in Theorem 1.7 are due to Memin *et al*. (2001). Extensions of these inequalities are discussed in Slominski and Ziemkiewicz (2005).

We now discuss some maximal inequalities for fBm due to Novikov and Valkeila (1999). Let $\{Z(t), t \geq 0\}$ be standard fBm. Let $\mathcal{F} = \{\mathcal{F}_t, t \geq 0\}$ be the filtration generated by fBm Z where $\mathcal{F}_t$ is the σ-algebra generated by the family $\{Z(s), 0 \leq s \leq t\}$. For any process X, define $X_t^* = \sup_{0 \leq s \leq t} |X_s|$. Since fBm is a self-similar process, it follows that $Z(at) \stackrel{\Delta}{=} a^H Z(t)$ for every $a > 0$. This in turn implies that $Z^*(at) \stackrel{\Delta}{=} a^H Z^*(t)$ for every $a > 0$. In particular, we have the following important result.

Proposition 1.9: Let $T > 0$ and Z be fBm with Hurst index H. Then, for every $p > 0$,

$$E[(Z_T^*)^p] = K(p, H) T^{pH} \tag{1.36}$$

where $K(p, H) = E[(Z_1^*)^p]$.

Proof: The result follows from the observation $Z_T^* \stackrel{\Delta}{=} T Z_1^*$ by self-similarity.

Novikov and Valkeila (1999) proved the following result. Recall that a random variable τ is said to be a stopping time with respect to the filtration $\mathcal{F} = \{\mathcal{F}_t, t \geq 0\}$ if the event $[\tau \leq t] \in \mathcal{F}_t$ for every $t \geq 0$.

Proposition 1.10: Let τ be any stopping time with respect to the filtration $\mathcal{F}$ defined above. Then, for any $p > 0$ and $H > \frac{1}{2}$, there exist positive constants $c_1(p, H)$ and $c_2(p, H)$ depending only on the parameters p and H such that

$$c_1(p, H)E(\tau^{pH}) \leq E[(Z_\tau^*)^p] \leq c_2(p, H)E(\tau^{pH}).$$

This result is the analogue of the classical Burkholder–Davis–Gundy inequality for martingales. However, recall that fBm is not a semimartingale. We point out that if $\{B_t, t \geq 0\}$ is standard Brownian motion, then, for any stopping time τ with respect to the filtration generated by the Brownian motion B and for any $p > 0$, there exist positive constants $c_1(p)$ and $c_2(p)$ depending only on p such that

$$c_1(p)E(\tau^{p/2}) \leq E[(B_\tau^*)^p] \leq c_2(p)E(\tau^{p/2}).$$

For a proof of Proposition 1.10, see Novikov and Valkeila (1999).

Deconvolution of fBm

Samorodnitsky and Taqqu (1994) proved that there is an integral transformation of standard Brownian motion W which provides a representation for standard fBm W^H with Hurst index $H \in (0, 1)$. It is the moving average representation of fBm and is given by

$$\{W^H(t), -\infty < t < \infty\} \stackrel{\Delta}{=} \left\{\int_{-\infty}^{\infty} f_H(t, u)dW(u), -\infty < t < \infty\right\} \tag{1.37}$$

in the sense that the processes on both sides have the same finite-dimensional distributions where

$$f_H(t, u) = \frac{1}{C_1(H)}((t-u)_+^{H-\frac{1}{2}} - (-u)^{H-\frac{1}{2}}) \tag{1.38}$$

and

$$C_1(H) = \left[\int_0^{\infty} ((1+u)^{H-\frac{1}{2}} - u^{H-\frac{1}{2}})^2 du + \frac{1}{2H}\right]^{1/2} \tag{1.39}$$

where $a_+^\alpha = a^\alpha$ if $a > 0$ and $a_+^\alpha = 0$ if $a \leq 0$. Pipiras and Taqqu (2002) obtained a generalization of this result. They proved the following theorem.

Theorem 1.11: Let W^{H_1} and W^{H_2} be two standard fBms with the Hurst indices $H_i \in (0, 1), i = 1, 2$ respectively. Further suppose that $H_1 \neq H_2$. Then

$$\{W^{H_2}(t), -\infty < t < \infty\} \stackrel{\Delta}{=} \left\{\int_{-\infty}^{\infty} f_{H_1,H_2}(t, u)dW^{H_1}(u), -\infty < t < \infty\right\} \tag{1.40}$$

where

$$f_{H_1,H_2}(t,u) = \frac{C_1(H_1)\Gamma(H_2+\frac{1}{2})}{C_1(H_2)\Gamma(H_1+\frac{1}{2})\Gamma(H_2-H_1+1)}((t-u)_+^{H_2-H_1} - (-u)_+^{H_2-H_1}) \tag{1.41}$$

with $C_1(\frac{1}{2}) = 1$.

By taking $H_2 = \frac{1}{2}$ in the above theorem, we get the following deconvolution formula or autoregressive representation for fBm proved in Pipiras and Taqqu (2002).

Theorem 1.12: Let W^H be standard fBm with index $H \in (0, 1)$ with $H \neq \frac{1}{2}$. Let W be standard Brownian motion. Then

$$\{W(t), -\infty < t < \infty\} \stackrel{\Delta}{=} \left\{\int_{-\infty}^{\infty} f_{H,\frac{1}{2}}(t,u)dW^H(u), -\infty < t < \infty\right\} \tag{1.42}$$

where

$$f_{H,\frac{1}{2}}(t,u) = \frac{C_1(H)}{\Gamma(H+\frac{1}{2})\Gamma(\frac{3}{2}-H)}((t-u)_+^{\frac{1}{2}-H} - (-u)_+^{\frac{1}{2}-H}). \tag{1.43}$$

Let $\mathcal{F}_{H,t}$ denote the σ-algebra generated by the process $\{W^H(s), 0 \leq s \leq t\}$ and $\mathcal{F}_{\frac{1}{2},t}$ denote the σ-algebra generated by the process $\{W(s), 0 \leq s \leq t\}$. Pipiras and Taqqu (2002) proved that the inversion formula

$$W(t) = \int_{-\infty}^{\infty} f_{H,\frac{1}{2}}(t,u)dW^H(u) \tag{1.44}$$

holds for each $t \in R$ almost everywhere and hence the σ-algebras $\mathcal{F}_{H,t}$ and $\mathcal{F}_{\frac{1}{2},t}$ are the same up to sets of measure zero for $t > 0$. The equality in (1.44) does not hold for $t < 0$.

The fundamental martingale

We noted earlier that fBm Z is not a semimartingale and its paths are continuous and locally of unbounded variation a.s. but with zero quadratic variation whenever $H > \frac{1}{2}$. However, we will now show that such fBm can be transformed into a martingale by an integral transformation following the work of Norros *et al.* (1999) (cf. Molchan (1969)). We will first prove a lemma dealing with some equations for integrals of fractional powers.

Lemma 1.13: Let $B(\alpha, \beta)$ denote the beta function with parameters α and β given by

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx. \tag{1.45}$$

The following identities hold:

(i) For $\mu > 0$, $\nu > 0$ and $c > 1$,

$$\int_0^1 t^{\mu-1}(1-t)^{\nu-1}(c-t)^{-\mu-\nu}\,dt = c^{-\nu}(c-1)^{-\mu}B(\mu,\nu). \tag{1.46}$$

(ii) For $\mu \in R$, $\nu > -1$ and $c > 1$,

$$\int_1^c t^{\mu}(t-1)^{\nu}\,dt = \int_0^{1-1/c} s^{\nu}(1-s)^{-\mu-\nu-2}\,ds. \tag{1.47}$$

(iii) Suppose that $\mu > 0$, $\nu > 0$ and $c > 1$. Then

$$\begin{aligned}\int_0^1 & t^{\mu-1}(1-t)^{\nu-1}(c-t)^{-\mu-\nu+1}\,dt \\ &= B(\mu,1-\mu)-(\mu+\nu-1)B(\mu,\nu) \\ &\quad \times \int_0^{1-1/c} s^{-\mu}(1-s)^{\mu+\nu-2}\,ds \quad (\text{if } \mu < 1) \\ &= (\mu+\nu-1)B(\mu,\nu)c^{-\nu+1}\int_0^1 s^{\mu+\nu-2}(c-s)^{-\mu}\,ds \quad (\text{if } \mu+\nu > 1).\end{aligned} \tag{1.48}$$

(iv) For $\mu \in (0,1)$ and for $x \in (0,1)$,

$$\int_0^1 t^{-\mu}(1-t)^{-\mu}|x-t|^{2\mu-1}dt = B(\mu,1-\mu). \tag{1.49}$$

We give a proof of this lemma following Norros *et al*. (1999).

Proof: The identities in (i) and (ii) can be proved by using the transformations $t = cs/(c-1+s)$ and $t = 1/(1-s)$ respectively. We now prove (iii).

Suppose $\mu < 1$. Then

$$\begin{aligned}\int_0^1 & t^{\mu-1}(1-t)^{\nu-1}(c-t)^{-\mu-\nu+1}\,dt \\ &= \int t^{\mu-1}(1-t)^{\nu-1}\left[(1-t)^{-\mu-\nu+1}+(-\mu-\nu+1)\int_1^c (v-t)^{-\mu-\nu}dv\right]dt \\ &= B(\mu,1-\mu)-(\mu+\nu-1)\int_1^c\left[\int_0^1 t^{\mu-1}(1-t)^{\nu-1}(v-t)^{-\mu-\nu}\,dt\right]dv \\ &= B(\mu,1-\mu)-(\mu+\nu-1)B(\mu,\nu)\int_1^c v^{-\nu}(v-1)^{-\mu}dv \quad (\text{by(i)}) \\ &= B(\mu,1-\mu)-(\mu+\nu-1)B(\mu,\nu)\int_0^{1-1/c} s^{-\mu}(1-s)^{\mu+\nu-2}ds \quad (\text{by(ii)})\end{aligned}$$

$$= B(\mu, 1-\mu) - (\mu+\nu-1)B(\mu,\nu) \times \left[B(1-\mu, \mu+\nu-1) - \int_0^{1/c} s^{\mu+\nu-2}(1-s)^{-\mu} ds\right]$$

$$= (\mu+\nu-1)B(\mu,\nu)\int_0^{1/c} s^{\mu+\nu-2}(1-s)^{-\mu} ds. \tag{1.50}$$

The last equality follows from the identity

$$(\mu+\nu-1)B(\mu,\nu)B(1-\mu,\mu+\nu-1) = B(\mu, 1-\mu). \tag{1.51}$$

Since the first term and the last term in (1.50) are analytic in μ for $\mu > 0$, the statement in (iii) holds for all $\mu > 0$. The last result given in (iv) follows from (iii) and (1.51) by elementary but tedious calculations. We omit the details.

Let Ψ denote the integral operator

$$\Psi f(t) = H(2H-1)\int_0^\infty f(s)|s-t|^{2H-2} ds.$$

Lemma 1.14: Suppose that $H > \frac{1}{2}$. Let $w(t,s)$ be the function

$$\begin{aligned} w(t,s) &= c_1 s^{\frac{1}{2}-H}(t-s)^{\frac{1}{2}-H} \quad \text{for } 0 < s < t \\ &= 0 \quad \text{for } s > t \end{aligned} \tag{1.52}$$

where

$$c_1 = \left[2H\Gamma\left(\frac{3}{2}-H\right)\Gamma\left(H+\frac{1}{2}\right)\right]^{-1}. \tag{1.53}$$

Then

$$\begin{aligned} \Psi w(t,.)(s) &= 1 \quad \text{for } s \in [0,t) \\ &= \frac{(H-\frac{1}{2})s^{H-\frac{1}{2}}}{(\frac{3}{2}-H)B(H+\frac{1}{2}, 2-2H)}\int_0^t u^{1-2H}(s-u)^{H-\frac{3}{2}} du \quad \text{for } s > t. \end{aligned} \tag{1.54}$$

Proof: Recall that $H > \frac{1}{2}$. For $s \in [0,t]$, the result follows from (iv) of Lemma 1.13 by choosing $\mu = H - \frac{1}{2}$. For $s > t$, the result is obtained from (iii) of Lemma 1.13.

Let

$$M_t = \int_0^t w(t,s) dZ_s, \ t \geq 0. \tag{1.55}$$

Theorem 1.15: The process $\{M_t, t \geq 0\}$ is a mean zero Gaussian process with independent increments with $E(M_t^2) = c_2^2 t^{2-2H}$ where

$$c_2 = \frac{c_H}{2H(2-2H)^{1/2}} \tag{1.56}$$

and

$$c_H = \left[\frac{2H\Gamma\left(\frac{3}{2} - H\right)}{\Gamma(H + \frac{1}{2})\Gamma(2 - 2H)} \right]^{1/2} \tag{1.57}$$

In particular, the process $\{M_t, t \geq 0\}$ is a zero mean martingale.

Proof: From the properties of Wiener integrals, it follows that the process $\{M_t, t \geq 0\}$ is a Gaussian process with mean zero. Suppose that $s < t$. In view of Lemma 1.14, it follows that

$$\begin{aligned} cov(M_s, M_t) &= \langle w(s, .), \Psi w(t.) \rangle \\ &= \langle w(s, .), I_{[0,t]} \rangle \\ &= \int_0^s w(s, u) du \\ &= c_1 B\left(\frac{3}{2} - H, \frac{3}{2} - H\right) s^{2-2H} \\ &= c_2^2 s^{2-2H}. \end{aligned} \tag{1.58}$$

Here $\langle . \rangle$ denotes the usual inner product in $L^2_{((0,\infty))}$. Note that the last term is independent of t which shows that the process M has uncorrelated increments. Since it is a mean zero Gaussian process, the increments will be mean zero independent random variables. Let $\mathcal{F}_t$ denote the σ-algebra generated by the random variables $\{Z_s, 0 \leq s \leq t\}$. It is now easy to see that the process $\{M_t, \mathcal{F}_t, t \geq 0\}$ is a zero mean Gaussian martingale.

The martingale M defined above is called the *fundamental martingale* associated with fBm Z.

It is easy to check that the process

$$W_t = \frac{2H}{c_H} \int_0^t s^{H-\frac{1}{2}} dM_s \tag{1.59}$$

is standard Brownian motion. Stochastic integration with respect to a martingale is defined in the obvious way as in the case of Ito integrals. For details, see Prakasa Rao (1999b).

Furthermore, for $0 \leq s \leq t$,

$$\begin{aligned} cov(Z_s, M_t) &= \langle I_{[0,s]}, \Psi w(t, .)\rangle \\ &= \langle I_{[0,s]}, I_{[0,t]}\rangle \\ &= s. \end{aligned} \tag{1.60}$$

In particular, it follows that the increment $M_t - M_s$ is independent of $\mathcal{F}_s$ for $0 \leq s \leq t$. Let

$$Y_t = \int_0^t s^{\frac{1}{2}-H} dZ_s.$$

Observing that the process $\{Y_t, t \geq 0\}$ is a Gaussian process, it can be seen that

$$Z_t = \int_0^t s^{H-\frac{1}{2}} dY_s.$$

In fact, the process Y generates the same filtration $\{\mathcal{F}_t, t \geq 0\}$ as the filtration generated by the process Z. It can be shown that

$$\begin{aligned} E[M_t Y_T] &= \frac{c_H^2}{2H} \int_0^t (T-s)^{H-\frac{1}{2}} s^{1-2H} ds \;\text{ if }\; t < T \\ &= \frac{T^{\frac{3}{2}-H}}{\frac{3}{2}-H} \;\text{ if }\; t \geq T \end{aligned} \tag{1.61}$$

For proof, see Proposition 3.2 in Norros *et al*. (1999). We leave it to the reader to check that

$$M_t = \frac{c_H}{2H} \int_0^t s^{\frac{1}{2}-H} dW_s. \tag{1.62}$$

The martingale M is the fundamental martingale associated with fBm Z in the sense that the martingale M generates, up to sets of measure zero, the same filtration as that generated by the process Z. Furthermore, the same holds for the related processes W and Y defined above. In fact, the process Y has the representation

$$Y_T = 2H \int_0^T (T-t)^{H-\frac{1}{2}} dM_t$$

and the martingale M_t can be represented in the form

$$M_t = c_1 \int_0^t (t-s)^{\frac{1}{2}-H} dY_s.$$

For detailed proofs of these results, see Norros *et al*. (1999). Let $\{\mathcal{F}_t^W\}$ denote the filtration generated by the process W defined by (1.59). It is known that all right-continuous square integrable $\{\mathcal{F}_t^W\}$-martingales can be expressed as stochastic integrals with respect to the process W. Since the filtrations $\{\mathcal{F}_t^W\}$ and $\{\mathcal{F}_t\}$ coincide up to sets of measure zero, it follows that all the right-continuous square integrable $\{\mathcal{F}_t\}$-martingales can also be expressed as stochastic integrals with respect to W.

Baxter-type theorem for fBm

In a fairly recent paper on the estimation of the Hurst index for fBm, Kurchenko (2003) derived a Baxter-type theorem for fBm.

Let $f : (a, b) \to R$ be a function and let k be a positive integer. Let $\Delta_h^{(k)} f(t)$ denote the increment of k-th order of the function f in an interval $[t, t+h] \subset (a, b)$ as defined by

$$\Delta_h^{(k)} f(t) = \sum_{i=0}^{k} (-1)^i {}^kC_i f\left(t + \frac{i}{k}h\right).$$

For any $m \geq 0$, positive integer $k \geq 1$ and $0 < H < 1$, define

$$V_k(m, H) = \frac{1}{2} \sum_{i,j=0}^{k} (-1)^{i+j+1} {}^kC_i {}^kC_j \left| m + \frac{i-j}{k} \right|^{2H}.$$

It can be checked that $V_1(0, H) = 1$ and $V_2(0, H) = 2^{2-2H} - 1$. Note that

$$\Delta_1^{(2)} f(t) = f(t) - 2\, f\left(t + \frac{1}{2}\right) + f(t+1).$$

Kurchenko (2003) proved the following Baxter-type theorem for second-order increments for fBm among several other results.

Theorem 1.16: Let $\{W_H(t), t \geq 0\}$ be standard fBm with Hurst index $H \in (0, 1)$ as defined above. Then, with probability one,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} (\Delta_1^{(2)} W_H(m))^2 = V_2(0, H) \text{ a.s.}$$

In other words,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} \left[W_H(m) - 2\, W_H\left(m + \frac{1}{2}\right) + W_H(m+1) \right]^2 = V_2(0, H) \text{ a.s.}$$

for any standard fBm with Hurst index $H \in (0, 1)$.

For a proof of Theorem 1.16, see Kurchenko (2003).

Singularity of fBms for different Hurst indices

It is well known that if P and Q are probability measures generated by two Gaussian processes, then these measures are either equivalent or singular with respect to each other (cf. Feldman (1958), Hajek (1958)). For a proof, see Rao (2000), p. 226.

Let $\{W_{H_i}(t), t \geq 0\}, i = 1, 2,$ be two standard fBms with Hurst indices $H_1 \neq H_2$. From the result stated above, it follows that the probability measures generated by these processes are either equivalent or singular with respect to each other. We will now prove that they are singular with respect to each other (cf. Prakasa Rao (2008c)).

Theorem 1.17: Let $\{W_{H_i}(t), t \geq 0\}, i = 1, 2,$ be two standard fBms with Hurst indices $H_1 \neq H_2$. Let P_i be the probability measure generated by the process $\{W_{H_i}(t, t \geq 0\}$ for $i = 1, 2$. Then the probability measures P_1 and P_2 are singular with respect to each other.

Proof: Applying Theorem 1.16, we obtain that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} \left[W_{H_i}(m) - 2W_{H_i}\left(m + \frac{1}{2}\right) + W_{H_i}(m+1) \right]^2$$
$$= V_2(0, H_i) \text{ a.s.}[P_i], i = 1, 2.$$

Since $V_2(0, H_1) \neq V_2(0, H_2)$ if $H_1 \neq H_2$, and since the convergence stated above is a.s. convergence under the corresponding probability measures, it follows that the measures P_1 and P_2 are singular with respect to each other.

Long-range dependence

Suppose $\{Z(t), -\infty < t < \infty\}$ is fBm with Hurst index H for some $0 < H < 1$. Define $X_k = Z(k+1) - Z(k)$ for any integer $k \in R$. The process $\{X_k\}$ is called *fractional Gaussian noise*. Since the process Z is H-ss with stationary increments, it is easy to see that the discrete time parameter process $\{X_k, -\infty < k < \infty\}$ is stationary with mean $E(X_k) = 0$, $E(X_k^2) = E(Z^2(1)) = \sigma^2$ (say) and the auto covariance

$$\gamma(k) = E(X_i X_{k+i}) = \gamma(-k) = \frac{\sigma^2}{2}(|k+1|^{2H} + |k-1|^{2H} - 2|k|^{2H}). \quad (1.63)$$

Suppose $k \neq 0$. Then

$$\begin{aligned} \gamma(k) &= 0 \text{ if } H = \frac{1}{2} \\ &< 0 \text{ if } 0 < H < \frac{1}{2} \\ &> 0 \text{ if } \frac{1}{2} < H < 1. \end{aligned} \quad (1.64)$$

This can be checked from the strict convexity of the function $f(x) = x^{2H}, x > 0$, for $\frac{1}{2} < H < 1$ and strict concavity of the function $f(x) = x^{2H}, x > 0$, for $0 < H < \frac{1}{2}$. Furthermore, if $H \neq \frac{1}{2}$, then

$$\gamma(k) \simeq \sigma^2 H(2H-1)|k|^{2H-2}$$

as $|k| \to \infty$. In particular, $\gamma(k) \to 0$ as $|k| \to \infty$ if $0 < H < 1$. Observe that, if $\frac{1}{2} < H < 1$, then $\sum_{k=-\infty}^{\infty} \gamma(k) = \infty$ and the process $\{X_k, -\infty < k < \infty\}$ exhibits long-range dependence. In this case, the auto covariance tends to zero but so slowly that the sum of auto covariances diverges. If $0 < H < \frac{1}{2}$, then $\sum_{k=-\infty}^{\infty} |\gamma(k)| < \infty$. This is the reason why the class of processes, driven by fBm with Hurst index $H > \frac{1}{2}$, is used for modeling phenomena with long-range dependence.

1.4 Stochastic differential equations driven by fBm

Fundamental semimartingale

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ be a stochastic basis satisfying the usual conditions. The natural filtration of a process is understood as the P-completion of the filtration generated by this process.

Let $W^H = \{W_t^H, t \geq 0\}$ be standard fBm with Hurst parameter $H \in (0, 1)$, that is, a Gaussian process with continuous sample paths such that $W_0^H = 0$, $E(W_t^H) = 0$ and

$$E(W_s^H W_t^H) = \frac{1}{2}[s^{2H} + t^{2H} - |s-t|^{2H}], t \geq 0, s \geq 0. \tag{1.65}$$

Let us consider a stochastic process $Y = \{Y_t, t \geq 0\}$ defined by the stochastic integral equation

$$Y_t = \int_0^t C(s)ds + \int_0^t B(s)dW_s^H, t \geq 0 \tag{1.66}$$

where $C = \{C(t), t \geq 0\}$ is an $(\mathcal{F}_t)$-adapted process and $B(t)$ is a non vanishing, non random function. For convenience, we write the above integral equation in the form of a stochastic differential equation

$$dY_t = C(t)dt + B(t)dW_t^H, t \geq 0 \tag{1.67}$$

driven by fBm W^H. Recall that the stochastic integral

$$\int_0^t B(s)dW_s^H \tag{1.68}$$

is not a stochastic integral in the Ito sense, but one can define the integral of a deterministic function with respect to fBm as the integrator in a natural sense (cf. Gripenberg and Norros (1996), Norros *et al*. (1999)) as we discussed earlier. Even though the process Y is not a semimartingale, one can associate a semimartingale $Z = \{Z_t, t \geq 0\}$ which is called a *fundamental semimartingale* such that the natural filtration $(\mathcal{Z}_t)$ of the process Z coincides with the natural filtration $(\mathcal{Y}_t)$ of the process Y (Kleptsyna *et al*. (2000a)). Define, for $0 < s < t$,

$$k_H = 2H \;\Gamma\left(\frac{3}{2} - H\right)\; \Gamma\left(H + \frac{1}{2}\right), \tag{1.69}$$

$$k_H(t,s) = k_H^{-1} s^{\frac{1}{2}-H}(t-s)^{\frac{1}{2}-H}, \tag{1.70}$$

$$\lambda_H = \frac{2H\;\Gamma(3-2H)\;\Gamma(H+\frac{1}{2})}{\Gamma(\frac{3}{2}-H)}, \tag{1.71}$$

$$w_t^H = \lambda_H^{-1} t^{2-2H}, \tag{1.72}$$

and

$$M_t^H = \int_0^t k_H(t,s)dW_s^H, t \geq 0. \tag{1.73}$$

The process M^H is a Gaussian martingale, called the *fundamental martingale* (cf. Norros *et al*. (1999)), and its quadratic variation $\langle M^H \rangle_t = w_t^H$. Furthermore, the natural filtration of the martingale M^H coincides with the natural filtration of fBm W^H. In fact the stochastic integral

$$\int_0^t B(s)dW_s^H \tag{1.74}$$

can be represented in terms of the stochastic integral with respect to the martingale M^H. For a measurable function f on $[0, T]$, let

$$K_H^f(t,s) = -2H\frac{d}{ds}\int_s^t f(r)r^{H-\frac{1}{2}}(r-s)^{H-\frac{1}{2}}dr, 0 \leq s \leq t \tag{1.75}$$

when the derivative exists in the sense of absolute continuity with respect to the Lebesgue measure (see Samko *et al*. (1993) for sufficient conditions). The following result is due to Kleptsyna *et al*. (2000a).

Theorem 1.18: Let M^H be the fundamental martingale associated with fBm W^H. Then

$$\int_0^t f(s)dW_s^H = \int_0^t K_H^f(t,s)dM_s^H, t \in [0, T] \tag{1.76}$$

a.s. $[P]$ whenever both sides are well defined.

Suppose the sample paths of the process $\{C(t)/B(t), t \geq 0\}$ are smooth enough (see Samko *et al*. (1993)) so that

$$Q_H(t) = \frac{d}{dw_t^H} \int_0^t k_H(t,s) \frac{C(s)}{B(s)} ds, t \in [0,T] \tag{1.77}$$

is well defined, where w^H and $k_H(t,s)$ are as defined in (1.72) and (1.70) respectively and the derivative is understood in the sense of absolute continuity. The following theorem due to Kleptsyna *et al*. (2000a) associates a *fundamental semimartingale* Z associated with the process Y such that the natural filtration $(\mathcal{Z}_t)$ coincides with the natural filtration $(\mathcal{Y}_t)$ of Y.

Theorem 1.19: Suppose the sample paths of the process Q_H defined by (1.77) belong P-a.s. to $L^2([0,T], dw^H)$ where w^H is as defined by (1.72). Let the process $Z = \{Z_t, t \in [0,T]\}$ be defined by

$$Z_t = \int_0^t k_H(t,s)[B(s)]^{-1} dY_s \tag{1.78}$$

where the function $k_H(t,s)$ is as defined in (1.70). Then the following results hold:

(i) The process Z is an $(\mathcal{F}_t)$-semimartingale with the decomposition

$$Z_t = \int_0^t Q_H(s) dw_s^H + M_t^H \tag{1.79}$$

where M^H is the fundamental martingale defined by (1.73).

(ii) The process Y admits the representation

$$Y_t = \int_0^t K_H^B(t,s) dZ_s \tag{1.80}$$

where the function $K_H^B(.,.)$ is as defined in (1.75).

(iii) The natural filtrations $(\mathcal{Z}_t)$ and $(\mathcal{Y}_t)$ coincide.

Kleptsyna *et al*. (2000a) derived the following Girsanov-type formula as a consequence of Theorem 1.19.

Theorem 1.20: Suppose the assumptions of Theorem 1.19 hold. Define

$$\Lambda_H(T) = \exp\left(-\int_0^T Q_H(t) dM_t^H - \frac{1}{2}\int_0^t Q_H^2(t) dw_t^H\right). \tag{1.81}$$

Suppose that $E(\Lambda_H(T)) = 1$. Then the measure $P^* = \Lambda_H(T)P$ is a probability measure and the probability measure of the process Y under P^* is the same as that of the process V defined by

$$V_t = \int_0^t B(s)dW_s^H, 0 \leq t \leq T \tag{1.82}$$

under the P-measure.

Stochastic differential equations

It is possible to define the stochastic integral of a random process $\{\sigma(t, X_t), t \geq 0\}$ with respect to fBm W^H as the integrator for some class of stochastic processes and to define a stochastic differential equation of the type

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t^H, X_0, t > 0.$$

Sufficient conditions for the existence and uniqueness of solutions of such stochastic differential equations driven by fBm are discussed in Mishura (2008), p. 197. We do not go into the details here. The following result due to Nualart and Rascanu (2002) gives sufficient conditions for the existence and uniqueness of the solution.

For any $\lambda \in (0, 1]$, let $C^\lambda[0, T]$ be the space of continuous functions f defined on the interval $[0, T]$ such that

$$\sup_{0 \leq x_1 \neq x_2 \leq T} \frac{|f(x_1) - f(x_2)|}{|x_1 - x_2|^\lambda} < \infty.$$

Define the norm

$$||f||_{C^\lambda} = \max_{x \in [0,T]} |f(x)| + \sup_{0 \leq x_1 \neq x_2 \leq T} \frac{|f(x_1) - f(x_2)|}{|x_1 - x_2|^\lambda} < \infty$$

on the space $C^\lambda[0, T]$. Let

$$C^{\mu-}[0, T] = \cap_{\lambda < \mu} C^\lambda[0, T].$$

Define

$$C_0 = \frac{1}{2}\left[\left(H - \frac{1}{2}\right) H(1 - H)\, B\left(\frac{3}{2} - H, \frac{3}{2} - H\right)\, B\left(H - \frac{1}{2}, \frac{3}{2} - H\right)\right]^{-1/2} \tag{1.83}$$

and

$$C_1 = C_0\, B\left(\frac{3}{2} - H, \frac{3}{2} - H\right). \tag{1.84}$$

Let

$$z(t,u) = C_0 u^{\frac{1}{2}-H}(t-u)^{\frac{1}{2}-H} \tag{1.85}$$

and

$$w(t,u) = C_0 u^{\frac{3}{2}-H}(t-u)^{\frac{1}{2}-H}. \tag{1.86}$$

Norros *et al*. (1999) proved that

$$M_t = \int_0^t z(t,u)dW_u^H \tag{1.87}$$

is well defined as a pathwise integral and is a martingale with respect to the filtration $\{\mathcal{F}_t, t \geq 0\}$ generated by the process W^H. The quadratic variation of the martingale M is $\langle M \rangle_t = \frac{t^{2-2H}}{2-2H}$.

Furthermore,

$$W_t = \int_0^t u^{H-\frac{1}{2}} dM_u \tag{1.88}$$

is a Wiener process W adapted to the same filtration.

Theorem 1.21: Let a function $S : [0, T] \times R \to R$ be such that:

(i) for all $N \geq 0$, there exists $L_N > 0$ such that

$$|S(t,x) - S(t,y)| \leq L_N|x-y|, |x| \leq N, |y| \leq N, 0 \leq t \leq T; \tag{1.89}$$

(ii) and there exists $M > 0$ such that

$$|S(t,x)| \leq M(1+|x|), x \in R, 0 \leq t \leq T. \tag{1.90}$$

Then the stochastic integral equation

$$X_t = x_0 + \int_0^t S(u, X_u)du + \epsilon W_t^H, 0 \leq t \leq T \tag{1.91}$$

or equivalently the stochastic differential equation

$$dX_t = S(t, X_t)dt + \epsilon dW_t^H, X_0 = x_0, 0 \leq t \leq T \tag{1.92}$$

has a unique solution $\{X_t, 0 \leq t \leq T\}$ and the sample paths of this process belong to $C^{H-}[0, T]$ with probability one.

Theorem 1.22: Suppose that the function $S(t, x)$ satisfies the conditions stated in Theorem 1.21. Furthermore, suppose that the constant L_N in Equation (1.89) does not depend on N, that is, $L_N = L$ for some L for every $N \geq 1$. Let $\{X_t, 0 \leq t \leq T\}$ be the solution of Equation (1.92) and

$$x_t = x_0 + \int_0^t S(u, x_u)du, 0 \leq t \leq T. \tag{1.93}$$

Then

$$\sup_{0\leq t\leq T} |X_t - x_t| \leq \epsilon C \sup_{0\leq t\leq T} |W_t^H| \tag{1.94}$$

where $C = e^{LT}$.

This inequality is a consequence of the Gronwall lemma (see Chapter 5).

Absolute continuity of measures

Consider the stochastic differential equations (SDEs)

$$dX_t = S_i(t, X_t)dt + \epsilon dW_t^H, X_0 = x_0, 0 \leq t \leq T, i = 1, 2. \tag{1.95}$$

Suppose that sufficient conditions are satisfied by the functions S_i so that there exist unique solutions for the SDEs defined above. Let X^i be the solution of the equation for $i = 1, 2$. Let P_i^T be the probability measure generated by the process X_i on the space $C[0, T]$ associated with the Borel σ-algebra induced by the supremum norm on the space $C[0, T]$. The following theorem, due to Androshchuk (2005), gives sufficient conditions under which the probability measures $P_i^T, i = 1, 2$, are equivalent to each other and gives a formula for the Radon–Nikodym derivative. An alternate form for the Radon–Nikodym derivative via the fundamental semimartingale is discussed in Theorem 1.20.

Theorem 1.23: Suppose the functions $S_i(t, x), i = 1, 2$, satisfy the following conditions: (i) $S_i(t, x) \in C^1([0, T] \times R)$; (ii) there exists a constant $M > 0$ such that $|S_i(t, x)| \leq M(1 + |x|), x \in R, 0 \leq t \leq T$. Then Equation (1.95) has unique solutions for $i = 1, 2$ and these solutions belong to $C^{H-}[0, T]$ a.s. In addition, the probability measures $P_i^T, i = 1, 2$, are absolutely continuous with respect to each other and

$$\frac{dP_2^T}{dP_1^T}(X^1) = \exp\left(\frac{1}{\epsilon}L_T - \frac{1}{2\epsilon^2}\langle L\rangle_T\right) \tag{1.96}$$

where

$$\begin{aligned} L_T = \int_0^T \Bigg\{ &(2-2H)t^{\frac{1}{2}-H} \\ &\times\left(C_1\Delta S(0, x_0) + \int_0^t u^{2H-3}\left[\int_0^u w(u, v)\, d(\Delta S(v, X_v^1))\right] du\right) \\ &+ t^{H-\frac{3}{2}} \int_0^t w(t, u)\, d(\Delta S(u, X_u^1))\Bigg\} dW_t \end{aligned} \tag{1.97}$$

with

$$\Delta S(t, x) = S_2(t, x) - S_1(t, x), \tag{1.98}$$

C_1 as defined by (1.84), $w(t, u)$ as given in (1.86) and W the Wiener process constructed from the process W^H using (1.88).

We omit the proof of this theorem. For details, see Androshchuk (2005).

1.5 Fractional Ornstein–Uhlenbeck-type process

We now study the fractional analogue of the Ornstein–Uhlenbeck process, that is, a process which is the solution of a one-dimensional homogeneous linear SDE driven by fBm W^H with Hurst index $H \in [\frac{1}{2}, 1)$.

Langevin (1908) suggested the following method to study the movement of a particle immersed in a liquid. He modeled the particle's velocity v by the equation

$$\frac{dv(t)}{dt} = -\frac{f}{m}v(t) + \frac{F(t)}{m} \tag{1.99}$$

where m is the mass of the particle, $f > 0$ is a friction coefficient and $F(t)$ is the fluctuating force resulting from the impact of the particles with the surrounding medium. Uhlenbeck and Ornstein (1930) studied a random version of the model by treating $F(t), t \geq 0$, as a random process and then derived that, for $v(0) = x$, the random variable $v(t)$ has a normal distribution with mean $xe^{-\lambda t}$ and variance $(\sigma^2/2\lambda)(1-e^{-2\lambda t})$ for $\lambda = f/m$ and $\sigma^2 = 2fkT/m^2$ where k is the Boltzmann constant and T is the temperature. Doob (1942) observed that, if $v(0)$ is a Gaussian random variable with mean zero and variance $\sigma^2/2\lambda$ independent of the stochastic process $\{F(t), t \geq 0\}$, then the solution $\{v(t), t \geq 0\}$ of (1.99) is stationary and the process

$$\left\{ I_{[t>0]} t^{1/2} v\left(\frac{1}{2\lambda}\log t\right), t \geq 0 \right\}$$

is Brownian motion. Since the sample paths of Brownian motion are nowhere differentiable a.s., the differential equation (1.99) has to be interpreted as a stochastic integral equation formulation or as a stochastic differential equation of the form

$$dX_t = -\lambda X_t dt + dW_t, X_0 = x, t \geq 0 \tag{1.100}$$

where $\{W_t, t \geq 0\}$ is Brownian motion. It can be shown that this equation has the unique solution

$$X_t = e^{-\lambda t}\left(x + \int_0^t e^{\lambda s} dW_s\right), t \geq 0 \tag{1.101}$$

by an application of Ito's lemma (cf. Prakasa Rao (1999a)). Such a process is called the Ornstein–Uhlenbeck process. In analogy with this formulation, consider the SDE

$$dX_t = -\lambda X_t dt + \sigma dW_t^H, X_0 = 0, t \geq 0 \tag{1.102}$$

where λ and σ^2 are constants. The existence and uniqueness of the solution of this SDE are discussed in Cheridito *et al*. (2003). This process is called a *fractional Ornstein–Uhlenbeck-type process*.

Existence and uniqueness

Theorem 1.24: Let $\{W_t^H, -\infty < t < \infty\}$ be fBm with index $H \in (0, 1]$ defined on a probability space $(\Omega, \mathcal{F}, P)$ and $X(0, \omega) = \eta(\omega) \in R$. Let $-\infty \leq a < \infty$ and $\lambda, \sigma > 0$. Then, for almost every $\omega \in \Omega$, the following hold:

(a) for all $t > a$,

$$\int_a^t e^{\lambda u} dW_u^H(\omega)$$

exists as a Riemann–Stieltjes integral and is equal to

$$e^{\lambda t} W_t^H(\omega) - e^{\lambda a} W_a^H(\omega) - \lambda \int_a^t W_u^H(\omega) e^{\lambda u} du;$$

(b) the function

$$\int_a^t e^{\lambda u} dW_u^H(\omega), t > a$$

is continuous in t; and

(c) the unique continuous function $x(t)$, that is, the solution of the integral equation

$$x(t) = \eta(\omega) - \lambda \int_0^t x(s) ds + \sigma W_t^H(\omega), t \geq 0$$

or equivalently of the SDE

$$dX(t) = -\lambda X(t) dt + \sigma dW_t^H, X(0) = \eta, t \geq 0$$

is given by

$$x(t) = e^{-\lambda t} \left[\eta(\omega) + \sigma \int_0^t e^{\lambda u} dW_u^H(\omega) \right], t \geq 0.$$

In particular, the unique continuous solution of the equation

$$x(t) = \sigma \int_{-\infty}^0 e^{\lambda u} dW_u^H(\omega) - \lambda \int_0^t x(s) ds + \sigma W_t^H(\omega), t \geq 0$$

is given by

$$x(t) = \sigma \int_{-\infty}^{t} e^{-\lambda(t-u)} dW_u^H(\omega), t \geq 0.$$

For a proof of this theorem, see Cheridito *et al.* (2003).

Let

$$Y_t^{H,\eta} = e^{-\lambda t}\left(\eta + \sigma \int_0^t e^{\lambda u} dW_u^H\right)$$

where the stochastic integral is defined as the pathwise Riemann–Stieltjes integral. As a consequence of the above theorem, it follows that $\{Y_t^{H,\eta}, t \geq 0\}$ is the unique a.s. continuous process that is a solution of the SDE

$$dX(t) = -\lambda X(t)dt + \sigma dW_t^H, t \geq 0, X(0) = \eta. \tag{1.103}$$

In particular, the process

$$Y_t^H = \sigma \int_{-\infty}^{t} e^{-\lambda(t-u)} dW_u^H, 0 \leq t < \infty$$

is the a.s. continuous solution of (1.103) with the initial condition

$$Y_0^H = \eta = \sigma \int_{-\infty}^{0} e^{\lambda u} dW_u^H.$$

Note that the process $\{Y_t^H, -\infty < t < \infty\}$ is a Gaussian process and is a stationary process as the increments of fBm W^H are stationary. Furthermore, for every η,

$$Y_t^H - Y_t^{H,\eta} \triangleq e^{-\lambda t}(Y_0^H - \eta)$$

and

$$e^{-\lambda t}(Y_0^H - \eta) \to 0 \text{ a.s. as } t \to \infty.$$

The process $\{Y_t^{H,\eta}, t \geq 0\}$ is a fractional Ornstein–Uhlenbeck-type process with the initial condition η and $\{Y_t^H, t \geq 0\}$ is a stationary fractional Ornstein–Uhlenbeck-type process. It can be checked that, for any fixed t, s,

$$\begin{aligned} cov(Y_t^H, Y_{t+s}^H) &= cov(Y_0^H, Y_s^H) \\ &= E\left[\sigma^2 \int_{-\infty}^{0} e^{\lambda u} dW_u^H \int_{-\infty}^{s} e^{-\lambda(s-v)} dW_v^H\right] \\ &= \frac{1}{2}\sigma^2 \sum_{n=1}^{N} \lambda^{-2n}(\Pi_{k=0}^{2n-1}(2H-k))s^{2H-2n} + O(s^{2H-2N-2}) \end{aligned} \tag{1.104}$$

and the last equality holds as $s \to \infty$ (for details, see Cheridito *et al*. (2003)). It can be shown that the process $\{Y_t^H, -\infty < t < \infty\}$ is ergodic and exhibits long-range dependence for $H \in (\frac{1}{2}, 1]$.

1.6 Mixed fBm

Cheridito (2001) introduced the class of mixed fBms. They are linear combinations of different fBms. We first consider a special case.

Let a and b be real numbers not both zero. *Mixed fractional Brownian motion* (mfBm) of parameters a, b and H is a process $M^H(a,b) = \{M_t^H(a,b), t > 0\}$ defined by

$$M_t^H(a,b) = aW_t + bW_t^H, t > 0$$

where $\{W_t, t \geq 0\}$ is Brownian motion and the process $\{W_t^H, t \geq 0\}$ is *independent* fBm with Hurst index H. It is easy to check that the process $M^H(a,b)$ is a zero mean Gaussian process with $E[(M_t^H(a,b))^2] = a^2 t + b^2 t^{2H}$, and

$$cov(M_t^H(a,b), M_s^H(a,b)) = a^2 \min(t,s) + \frac{1}{2} b^2 (|t|^{2H} + |s|^{2H} - |t-s|^{2H}),$$
$$t > 0, s > 0.$$

Furthermore, the increments of the process $M^H(a,b)$ are stationary and it is mixed-self-similar in the sense that, for any $h > 0$,

$$\{M_{ht}^H(a,b)\} \stackrel{\Delta}{=} \{M_t^H(ah^{1/2}, bh^H)\}.$$

Recall that the notation $\{X_t\} \stackrel{\Delta}{=} \{Y_t\}$ means that the processes specified on both sides have the same finite-dimensional distributions.

Proposition 1.25: For all $0 < H < 1$, $H \neq \frac{1}{2}$, and $b \neq 0$, the process $M^H(a,b)$ is *not* a Markov process.

Proof: For notational convenience, we write M^H for $M^H(a,b)$. If the process M^H is Markov, then, for all $s < t < u$,

$$cov(M_s^H, M_u^H) \; var(M_t^H) = cov(M_s^H, M_t^H) \; cov(M_t^H, M_u^H)$$

from the results in Revuz and Yor (1991). It can be checked that the above identity does not hold, for instance, for $s = \frac{1}{2}$, $t = 1$ and $u = \frac{3}{2}$ whenever $0 < H < 1$ and $H \neq \frac{1}{2}$. For details, see Zili (2006).

Suppose $b \neq 0$. We leave it to the reader to check that the increments of the process M^H are long-range dependent if $H > \frac{1}{2}$.

Proposition 1.26: For all $T > 0$ and $0 < \gamma < \min(H, \frac{1}{2})$, the mixed fBm M^H has a version of the process with sample paths which are Holder-continuous of order γ in the interval $[0, T]$ with probability one.

Proof: Let $\alpha > 0$ and $0 \le s \le t \le T$. By the stationarity of the increments and mixed-self-similarity of the process M^H, it follows that

$$\begin{aligned} E(|M_t^H - M_s^H|^\alpha) &= E(|M_{t-s}^H|^\alpha) \\ &= E(|M_1^H(a(t-s)^{1/2}, b(t-s)^H)|^\alpha). \end{aligned} \tag{1.105}$$

(i) Suppose $H \le \frac{1}{2}$. Then, there exist positive constants C_1 and C_2 depending on α such that

$$\begin{aligned} E(|M_t^H - M_s^H|^\alpha) &\le (t-s)^{\alpha H} E(|M_1^H(a(t-s)^{\frac{1}{2}-H}, b)|^\alpha) \\ &\le (t-s)^{\alpha H}[C_1|a|^\alpha (t-s)^{a(\frac{1}{2}-H)} E|W_1|^\alpha \\ &\quad + C_2|b|^\alpha E(|W_1^H|^\alpha)] \\ &\le C_\alpha (t-s)^{\alpha H} \end{aligned} \tag{1.106}$$

where

$$C_\alpha = C_1|a|^\alpha (t-s)^{a(\frac{1}{2}-H)} E|W_1|^\alpha + C_2|b|^\alpha E(|W_1^H|^\alpha).$$

(ii) Suppose $H > \frac{1}{2}$. Then, there exist positive constants C_3 and C_4 depending on α such that

$$\begin{aligned} E(|M_t^H - M_s^H|^\alpha) &\le (t-s)^{\alpha/2} E(|M_1^H(a, b(t-s)^{H-\frac{1}{2}})|^\alpha) \\ &\le (t-s)^{\alpha/2}[C_3|a|^\alpha E(|W_1|^\alpha) \\ &\quad + C_4|b|^\alpha (t-s)^{\alpha(H-\frac{1}{2})} E(|W_1^H|^\alpha)] \\ &\le C'_\alpha (t-s)^{\alpha/2} \end{aligned} \tag{1.107}$$

where

$$C'_\alpha = C_3|a|^\alpha E(|W_1|^\alpha) + C_4|b|^\alpha (t-s)^{\alpha(H-\frac{1}{2})} E(|W_1^H|^\alpha).$$

Hence, for every $\alpha > 0$, there exists $C_\alpha > 0$ such that

$$E(|M_t^H - M_s^H|^\alpha) \le C_\alpha |t-s|^{\alpha \min(\frac{1}{2}, H)}.$$

The result stated in the theorem now follows from Kolmogorov's theorem (cf. Revuz and Yor (1991), p. 25).

Miao *et al*. (2008) introduced the class of *fractional mixed fractional Brownian motion process* $\{Z_t, t > 0\}$ given by

$$Z_t^H = aW_t^{H_1} + bW_t^{H_2},$$

where $\{W_t^{H_1}, t \geq 0\}$ and $\{W_t^{H_2}, t \geq 0\}$ are independent fBms with Hurst indices H_1 and H_2 respectively, and studied the properties of such processes. For details, see Miao *et al*. (2008).

Suppose $b \neq 0$. Cheridito (2001) proved that the mixed fBm $M_t^H(a, b)$ is not a semimartingale if $H \in (0, \frac{1}{2}) \cup (\frac{1}{2}, \frac{3}{4}]$. Furthermore, it is equivalent to a multiple of the Wiener process if $H = \frac{1}{2}$ and equivalent to a Wiener process if $H \in (\frac{3}{4}, 1]$. For details, see Cheridito (2001).

A function $f(t), t \geq 0$, is said to belong to the lower class of a process X defined on a probability space $(\Omega, \mathcal{F}, P)$ if, for almost all $\omega \in \Omega$, there exists a function $t_0 = t_0(\omega)$ such that $X(t) \geq f(t)$ for every $t > t_0$. El-Nouty (2001, 2002) characterized such classes for fBm and extended the results to fractional mixed fBm in El-Nouty (2003a) and to integrated fBm in El-Nouty (2003b).

1.7 Donsker-type approximation for fBm with Hurst index $H > \frac{1}{2}$

Let $Z = \{Z(t), t \geq 0\}$ be standard fBm with Hurst index $H > \frac{1}{2}$. Norros *et al*. (1999) obtained the following kernel representation of the process Z with respect to standard Brownian motion W:

$$Z(t) = \int_0^t z(t, s) dW_s \tag{1.108}$$

where

$$z(t, s) = c_H \left(H - \frac{1}{2}\right) s^{\frac{1}{2}-H} \int_s^t u^{H-\frac{1}{2}} (u - s)^{H-\frac{1}{2}} du, \quad s \leq t \tag{1.109}$$

with

$$c_H = \left(\left[2H\Gamma\left(\frac{3}{2} - H\right)\right] / \left[\Gamma\left(\frac{1}{2} + H\right)\Gamma(2 - 2H)\right]\right)^{1/2}. \tag{1.110}$$

The function $z(t, s)$ is defined to be zero if $s \geq t$. We now briefly discuss an analogue of the Donsker-type approximation theorem for fBm as a limit of a random walk. This result is due to Sottinen (2001).

Let $\psi_i^{(n)}$ be independent and identically distributed (i.i.d.) random variables with mean zero and variance one and define

$$W_t^{(n)} = n^{-1/2} \sum_{i=1}^{[nt]} \psi_i^{(n)} \tag{1.111}$$

where $[x]$ denotes the greatest integer not exceeding x. Donsker's theorem states that the process $\{W_t^{(n)}, t \geq 0\}$ converges weakly to the standard Brownian motion W (cf. Billingsley (1968)). Let

$$Z_t^{(n)} = \int_0^t z^{(n)}(t,s)dW_s^{(n)} = \sum_{i=1}^{[nt]} n \left(\int_{(i-1/n)}^{(i/n)} z\left(\frac{[nt]}{n}, s\right) ds \right) (n^{-1/2}\psi_i^{(n)}) \tag{1.112}$$

where $z(t,s)$ is the kernel that transforms standard Brownian into fBm. Note that the function $z^{(n)}(t,.)$ is an approximation to the function $z(t,.)$, namely,

$$z^{(n)}(t,s) = n \int_{s-1/n}^{s} z\left(\frac{[nt]}{n}, u\right) du.$$

Sottinen (2001) proved that the random walk $Z^{(n)}$ converges weakly to standard fBm with index H. For a detailed proof, see Sottinen (2001).

Weak convergence to fBm was also investigated by Beran (1994) and Taqqu (1975). The approximation schemes discussed by them involve Gaussian random variables. Dasgupta (1998) obtained approximations using binary random variables and the representation of fBm due to Mandelbrot and Van Ness (1968). Sottinen's approximation scheme discussed above used i.i.d. random variables with finite variance.

1.8 Simulation of fBm

We mentioned earlier that the increments of fBm with Hurst index H form a sequence called fractional Gaussian noise and this sequence exhibits long-range dependence whenever $\frac{1}{2} < H < 1$. Mandelbrot and Wallis (1969) provided a discrete approximation to fBm and Mandelbrot (1971) suggested a fast algorithm for simulating the fractional Gaussian noise. We now describe a few methods to simulate paths of fBm. Dieker (2004) has given an extensive discussion on this topic comparing different methods of simulation. Our remarks here are based on Dieker (2004).

Willinger, Taqqu, Sherman and Wilson method

The following method for the simulation of fBm paths is due to Willinger *et al*. (1997). Suppose there are S i.i.d. sources transmitting packets of information. Each source s has active and inactive periods modeled by a stationary time series $\{J^{(s)}(t), t \geq 0\}$ where $J^{(s)}(t) = 1$ if the source is sending a packet at time t and $J^{(s)}(t) = 0$ if the source is not sending a packet at time t. Suppose the lengths of the active ('ON') periods are i.i.d. and those of the inactive ('OFF') periods are also i.i.d., and the lengths of ON and OFF periods are independent.

An OFF period follows an ON period and the ON and OFF period lengths may have different distributions. Rescaling time by a factor T, let

$$J_S(Tt) = \int_0^{Tt} \left[\sum_{s=1}^{S} J^{(s)}(u) \right] du$$

be the aggregated cumulative packet counts in the interval $[0, t]$. Suppose the distributions of the ON and OFF periods are Pareto with parameter $1 < \alpha < 2$. Recall that a random variable X has the Pareto distribution with parameter $\alpha > 0$ if $P(X > t) = t^{-\alpha}$ for $t \geq 1$. Note that the ON and OFF periods have infinite variance under this distribution when $1 < \alpha < 2$. Willinger *et al.* (1997) proved that

$$\lim_{T\to\infty} \lim_{S\to\infty} T^{-H} S^{-1/2} \left(J_S(Tt) - \frac{1}{2} T St \right) = \sigma W_H(t)$$

for some $\sigma > 0$ where $H = (3-\alpha/2)$ and $W_H(t)$ denotes fBm with Hurst index H. In other words, the random variable $J_S(Tt)$ closely resembles

$$\frac{1}{2} T St + T^H \sqrt{S} \sigma W_H(t)$$

which is fractional Brownian traffic with mean $M = \frac{1}{2} T S$ and variance coefficient $a = 2\sigma^2 T^{2H-1}$. We say that $A(t) = Mt + \sqrt{aM} W_H(t)$ is fractional Brownian traffic with mean input rate $M > 0$ and variance coefficient a. The process $A(t)$ represents the number of bits (or data packets) that is transmitted in the time interval $[0, t]$.

The method given above can be used for simulation of fBm by aggregating a large number of sources with Pareto ON and OFF periods.

Decreusfond and Lavaud method

Decreusefond and Lavaud (1996) suggested the following method for the simulation of an fBm sample. Recall that fBm $\{W_H(t), t \geq 0\}$ can be represented in the form

$$W_H(t) = \int_0^t K_H(t, s) dW(s)$$

for a suitable function $K_H(t, s)$ where $\{W(t), t \geq 0\}$ is Brownian motion. Suppose that we need an fBm sample in the interval $[0, 1]$. Let $t_j = j/N, j = 0, 1, \ldots, N$. We estimate $W_H(t_j)$ at t_j by the formula

$$W_H(t_j) = \sum_{i=0}^{j} \frac{1}{t_{i+1} - t_i} \left[\int_{t_i}^{t_{i+1}} K_H(t_j, s) ds \right] (W(t_{i+1}) - W(t_i)). \qquad (1.113)$$

Note that the integral

$$\int_{t_i}^{t_{i+1}} K_H(t_j, s)ds$$

cannot be approximated by $K_H(t_j, t_i)$ or $K_H(t_j, t_{i+1})$ since the function $K_H(t_j, t)$ is not continuous with respect to t in $[0, t_j]$.

Dzhaparidze and van Zanten method

Dzhaparidze and van Zanten (2004) obtained a series expansion for fBm. This series involves the positive zeroes $x_1 < x_2 < \ldots$ of the Bessel function J_{-H} of the first kind of order $-H$ and the positive zeroes $y_1 < y_2 < \ldots$ of the Bessel function J_{1-H}. Then

$$W_H(t) = \sum_{n=1}^{\infty} \frac{\sin(x_n t)}{x_n} X_n + \sum_{n=1}^{\infty} \frac{1-\cos(y_n t)}{y_n} Y_n. \tag{1.114}$$

Equality holds in Equation (1.114) in the sense that random processes on both sides have the same finite-dimensional distributions. The random variables $X_i, i \geq 1$, and $Y_i, i \geq 1$, are independent Gaussian random variables with mean zero and with the variances given by

$$Var(X_n) = 2C_H^2 x_n^{-2H} J_{1-H}^{-2}(x_n),$$
$$Var(Y_n) = 2C_H^2 y_n^{-2H} J_{-H}^{-2}(y_n),$$

and

$$C_H^2 = \frac{1}{\pi}\Gamma(1+2H)\sin(\pi H).$$

Furthermore, the series on the right of (1.114) converge absolutely and uniformly in $t \in [0, 1]$ a.s. The series expansion in (1.114) generalizes the result on Karhunen–Loeve-type expansion for Brownian motion. The representation (1.114) can be used for simulating an fBm sample from Gaussian samples. This method is useful as there are efficient algorithms to compute the zeroes of Bessel functions. Furthermore, the zeroes have to be computed only once regardless of the number of samples to be simulated. For computational purposes, the series on the right of (1.114) have to be truncated at some level N. Dzhaparidze and van Zanten (2003) proved that

$$\limsup_{N\to\infty} \frac{N^H}{\sqrt{\log N}} E\left[\sup_{0\leq t\leq 1}\left|\sum_{n>N} \frac{\sin(x_n t)}{x_n} X_n + \sum_{n>N} \frac{1-\cos(y_n t)}{y_n} Y_n\right|\right] < \infty.$$

Kuhn and Linde (2002) proved the rate $N^{-H}\sqrt{\log N}$ is the 'best' possible.

1.9 Remarks on application of modeling by fBm in mathematical finance

Geometric Brownian motion is used for modeling stock prices in the theory of mathematical finance but it was empirically observed that the model may not be suitable due to several reasons, including the fact that the log-share prices may follow long-range dependence. In recent years, fBm has been suggested as a replacement for Brownian motion as the driving force in modeling various real-world phenomena, including the modeling of stock prices. Absence of arbitrage, that is, the impossibility of receiving a risk-less gain by trading in a market, is a basic assumption or a condition that underlies all modeling in financial mathematics. For, if there is a strategy that is feasible for investors and promises a risk-less gain, then the investors would like to buy this strategy and will not sell. By the law of demand and supply, the price of this strategy would increase immediately, indicating that the market is not in equilibrium. Hence the absence of arbitrage is a basic requirement of any useful pricing model. See Rogers (1997) and Bender *et al*. (2006).

The first fundamental theorem of asset pricing (Delbaen and Schachermayer (1994)) links the no-arbitrage property to the martingale property of the discounted stock price process under a suitable pricing measure. Since fBm is not a semimartingale, except when $H = \frac{1}{2}$ (the Brownian motion case), the stock price process driven by fBm cannot be transformed into a martingale in general by an equivalent change of measure. Hence the fundamental theorem rules out these models as sensible pricing models.

Hu and Oksendal (2003) and Elliott and van der Hock (2003) suggested a fractional Black–Scholes model as an improvement over the classical Black–Scholes model using the notion of a Wick integral. Necula (2002) studied option pricing in a fractional Brownian environment using the Wick integral. Common to these fractional Black–Scholes models is that the driving force is fBm but the stochastic integral used is interpreted as the Wick integral. It was shown by these authors that the fractional Black–Scholes models are arbitrage free, contradicting earlier studies that the fractional Black–Scholes models do admit arbitrage. Bjork and Hult (2005) have, however, pointed out that the notion of self-financing trading strategies and the definition of value used by Hu and Oksendal (2003) and others, using the Wick integral, do not have a reasonable economic interpretation.

1.10 Pathwise integration with respect to fBm

Zahle (1998) developed the theory of pathwise integration with respect to fBm when the Hurst index $H > \frac{1}{2}$ using the methods of fractional calculus.

We will now briefly discuss these results. Zahle (1998) extends the classical Lebesgue–Stieltjes integral

$$\int_a^b f(x)\, dg(x)$$

for real or complex-valued functions on a finite interval (a, b) to a large class of integrands f and integrators g of unbounded variation. The techniques used are composition formulas and integration by parts rules for fractional integrals and fractional derivatives (cf. Samko *et al*. (1993)).

Note that if f or g is a smooth function on a finite interval (a, b), the Lebesgue–Stieltjes integral can be written in the form

$$\int_a^b f(x)\, dg(x) = \int_a^b f(x)g'(x)dx$$

or

$$\int_a^b f(x)\, dg(x) = -\int_a^b f'(x)g(x)dx + f(b-)g(b-) - f(a+)g(a+).$$

Here $f(a+) = \lim_{\delta \swarrow 0} f(a+\delta)$ and $g(b-) = \lim_{\delta \nearrow 0} f(b-\delta)$ whenever the limits exist. The main idea of Zahle's approach is to replace the ordinary derivatives by the fractional derivatives. Let

$$f_{a+}(x) = (f(x) - f(a+))1_{(a,b)}(x)$$

and

$$g_{b-}(x) = (g(x) - g(b-))1_{(a,b)}(x)$$

where $1_{(a,b)}(x) = 1$ if $x \in (a, b)$ and $1_{(a,b)}(x) = 0$ otherwise. For a function $f \in L_1(R)$ and $\alpha > 0$, define

$$I_{a+}^{\alpha} f(x) = \frac{1}{\Gamma(\alpha)} \int_a^x (x-y)^{\alpha-1} f(y)dy$$

and

$$I_{b-}^{\alpha} f(x) = \frac{1}{\Gamma(\alpha)} \int_x^b (y-x)^{\alpha-1} f(y)dy$$

where $\Gamma(.)$ is the gamma function. For $p \geq 1$, let $I_{a+}^{\alpha}(L_p)$ be the class of functions f which may be represented as I_{a+}^{α}-integral for some function ϕ in $L_p(R)$. Similarly, let $I_{b-}^{\alpha}(L_p)$ be the class of functions f which may be represented as I_{b-}^{α}-integral for some function ϕ in $L_p(R)$. If $p > 1$, then $f \in I_{a+}^{\alpha}(L_p)$ if and only if $f \in L_p(R)$ and the integrals

$$\int_a^{x-\epsilon} \frac{f(x) - f(y)}{(x-y)^{\alpha+1}} dy$$

converge in $L_p(R)$ as a function of x as $\epsilon \downarrow 0$ defining $f(y) = 0$ if x is not in (a, b). Similarly, $f \in I^{\alpha}_{b-}(L_p)$ if and only if $f \in L_p(R)$ and the integrals

$$\int_{x+\epsilon}^{b} \frac{f(x) - f(y)}{(y - x)^{\alpha+1}} dy$$

converge in $L_p(R)$ as a function of x as $\epsilon \downarrow 0$ defining $f(y) = 0$ if x is not in $[a, b]$ (cf. Samko *et al*. (1993)).

Suppose $f_{a+} \in I^{\alpha}_{a+}(L_p)$ and $g_{b-} \in I^{1-\alpha}_{b-}(L_q)$ for some $1/p + 1/q \leq 1$ and $0 \leq \alpha \leq 1$. Define the integral

$$\int_{a}^{b} f(x)\, dg(x) = (-1)^{\alpha} \int_{a}^{b} D^{\alpha}_{a+} f_{a+}(x) D^{1-\alpha}_{b-} g_{b-}(x) dx + f(a+)(g(b-) - g(a+)) \tag{1.115}$$

for some $0 \leq \alpha \leq 1$ where

$$D^{\alpha}_{a+} f(x) = \frac{1}{\Gamma(1-\alpha)} \left[\frac{f(x)}{(x-a)^{\alpha}} + \alpha \int_{0}^{x} \frac{f(x) - f(y)}{(x-y)^{\alpha+1}} dy \right] 1_{(a,b)}(x)$$

and

$$D^{\alpha}_{b-} f(x) = \frac{(-1)^{\alpha}}{\Gamma(1-\alpha)} \left[\frac{f(x)}{(b-x)^{\alpha}} + \alpha \int_{x}^{b} \frac{f(x) - f(y)}{(y-x)^{\alpha+1}} dy \right] 1_{(a,b)}(x)$$

and the convergence of the integrals above at the singularity $y = x$ holds point wise for almost all x if $p = 1$ and in the $L_p(R)$-sense if $p > 1$. It can be shown that the integral defined by Equation (1.115) is independent of the choice of α (cf. Zahle (1998)). Furthermore, for $\alpha p < 1$, the function $f_{a+} \in I^{\alpha}_{a+}(L_p)$ if and only if $f \in I^{\alpha}_{a+}(L_p)$ and $f(a+)$ exists.

It was pointed out earlier that fBm W^H with Hurst index H has a version with sample paths of Holder exponent H, that is, of Holder continuity of all orders $\lambda < H$, in any finite interval $[0, T]$ with probability one. Holder continuity implies existence path wise of the integral (1.115) and hence of the integral

$$\int_{0}^{t} f(s) dW^H_s, 0 \leq t \leq T$$

with probability one for any function f defined in the interval $[0, T]$ such that $f_{0+} \in I^{\alpha}_{0+}(L_1(0, T))$ with probability one for some $\alpha > 1-H$. Note that there is no condition where the function f has to be adapted with respect to the filtration generated by the process W^H. Let $H^{\lambda}(0, T)$ be the family of functions which are Holder continuous of order λ in the interval $[0, T]$. Suppose $\lambda > 1-H$. Then we can interpret the integral as a Riemann–Stieltjes integral and use the change-of-variable formula given below. In particular, we may define the stochastic integral with respect to W^H for functions of the form $f(t) = \sigma(t, X(t))$ for some real-valued Lipschitz function $\sigma(.,.)$ and any stochastic process with sample paths in

$H^{\lambda}(0,T)$ with probability one for some $\lambda > 1-H$. Since $H > \frac{1}{2}$, it is possible to study SDEs of the type

$$dX(t) = aX(t)dt + bX(t)dW_t^H, t \geq 0 \tag{1.116}$$

or equivalently

$$X(t) = X(0) + a\int_0^t X(s)ds + b\int_0^t X(s)dW_s^H, t \geq 0. \tag{1.117}$$

It can be shown that the solution of the above SDE is

$$X(t) = X(0)\exp[at + bW^H(t)], t \geq 0. \tag{1.118}$$

Change-of-variable formula

It is known that the chain rule

$$dF(f(x)) = F'(f(x))df(x)$$

does not hold for functions f of Holder exponent $\frac{1}{2}$ arising as sample paths of stochastic processes which are semimartingales. However, for functions of Holder exponent greater than $\frac{1}{2}$, the classical formula remains valid in the sense of Riemann–Stieltjes integration. The following change-of-variable formula can be proved (cf. Zahle (1998)).

Theorem 1.27: If $f \in H^{\lambda}(a,b)$ and $F \in C^1(R)$ is a real-valued function such that $F'(f(.)) \in H^{\mu}(a,b)$ for some $\lambda + \mu > 1$, then, for any $y \in (a,b)$,

$$F(f(y)) - F(f(a)) = \int_a^y F'(f(x))df(x). \tag{1.119}$$

Remarks: The conditions in Theorem 1.27 will be satisfied if $f \in H^{\lambda}(a,b)$ for some $\lambda > \frac{1}{2}$ and $F \in C^1(R)$ with Lipschitzian derivative.

Theorem 1.27 can be extended to a more general version in the following way.

Theorem 1.28: Suppose $f \in H^{\lambda}(a,b)$, $F \in C^1(R \times (a,b))$ and $F_1'(f(.),.) \in H^{\mu}(a,b)$, $\lambda + \mu > 1$. Then

$$F(f(y),y) - F(f(a),a) = \int_a^y F_1'(f(x),x)df(x) + \int_a^y F_2'(f(x),x)df(x) \tag{1.120}$$

where F_1' and F_2' are the partial derivatives of f with respect to the first and second variable respectively.

As an example, suppose that $f \in H^{\lambda}(a, b)$ for some $\lambda > \frac{1}{2}$ and $F(u) = u^2$. Then

$$\int_a^y f(x)df(x) = \frac{1}{2}(f^2(y) - f^2(a)).$$

As an application of the change-of-variable formula, it follows that

$$\int_x^y W_t^H \ dW_t^H = \frac{1}{2}[(W_y^H)^2 - (W_x^H)^2], 0 \le x \le y < \infty$$

with probability one provided $H > \frac{1}{2}$.

2

Parametric estimation for fractional diffusion processes

2.1 Introduction

Statistical inference for diffusion-type processes satisfying SDEs driven by Wiener processes was studied earlier and a comprehensive survey of various methods is given in Prakasa Rao (1999a). There has been some recent interest in studying similar problems for stochastic processes driven by fBm.

2.2 SDEs and local asymptotic normality

One of the basic tools in the study of asymptotic theory of statistical inference is the concept of local asymptotic normality. Several important properties of estimators of parameters involved in such processes follow as a consequence of the local asymptotic normality of the family of probability measures generated by the processes. Consider the SDE

$$dX_t = S(\theta, t, X_t)dt + \epsilon dW_t^H,\ X_0 = x_0, 0 \le t \le T \tag{2.1}$$

where $x_0 \in R,\ \epsilon \in (0, 1),\ S(\theta, t, x) : R^d \times [0, T] \times R \to R$ is a non random function of the drift, $\theta \in \Theta \subset R^d$ is an unknown parameter and $W^H = \{W_t^H, 0 \le t \le T\}$ is fBm with Hurst index $H \in (\frac{1}{2}, 1)$. Equation (2.1) models a dynamical system with small noise which is fBm. We will call such a process a *fractional diffusion process* hereafter. Suppose the process $\{X_t, 0 \le t \le T\}$ is observed over an interval $[0, T]$. The problem of interest is the estimation of the parameter θ

Statistical Inference for Fractional Diffusion Processes B. L. S. Prakasa Rao

based on the observation or the data $\{X_t, 0 \leq t \leq T\}$. The problem of estimation of the Hurst index H is also very important. However, we will assume that H is known in the following discussion. We will discuss the problem of estimation of the Hurst index briefly in Chapter 9. Several properties of estimators such as maximum likelihood or Bayes estimators can be derived as a consequence of the local asymptotic normality of the family of probability measures generated by the processes satisfying the SDE. For the study of statistical inference for diffusion-type processes, that is, for the case $H = \frac{1}{2}$, or for semimartingales, see Prakasa Rao (1999a,b).

We now define the concept of *local asymptotic normality* for a family of probability measures. Let $(\mathcal{X}^\epsilon, \mathcal{F}^\epsilon, P_\theta^\epsilon)$ be a family of probability spaces and let $\Theta \subset R^d$ be open. Let $\mathcal{E}_\epsilon = \{\mathcal{X}^\epsilon, \mathcal{F}^\epsilon, P_\theta^\epsilon, \theta \in \Theta\}$ be a family of statistical experiments and X^ϵ be the corresponding observation. Let

$$\frac{dP_{\theta_2}^\epsilon}{dP_{\theta_1}^\epsilon}$$

be the Radon–Nikodym derivative of the absolutely continuous component of the measure $P_{\theta_2}^\epsilon$ with respect to the measure $P_{\theta_1}^\epsilon$. This is called the *likelihood ratio.*

Definition: A family of probability measures $\{P_\theta^\epsilon, \theta \in \Theta\}$ is called *locally asymptotically normal* at $\theta_0 \in \Theta$ as $\epsilon \to 0$, if

$$Z_{\epsilon,\theta_0}(u) = \frac{dP_{\theta_0+\phi_\epsilon u}^\epsilon}{dP_{\theta_0}^\epsilon} = \exp\left[u'\Delta_{\epsilon,\theta_0} - \frac{1}{2}||u||^2 + \psi_\epsilon(u,\theta_0)\right]$$

and $\mathcal{L}(\Delta_{\epsilon,\theta_0}|P_{\theta_0}^\epsilon) \to N(0, J)$ as $\epsilon \to 0$ for all $u \in R^d$ for some non singular $d \times d$ matrix $\phi_\epsilon = \phi_\epsilon(\theta_0)$, where J is the identity matrix of order $d \times d$ and the function ψ is such that

$$\psi_\epsilon(u,\theta_0) \to 0 \text{ in probability } P_{\theta_0}^\epsilon \text{ as } \epsilon \to 0$$

for every $u \in R^d$.

Here $\mathcal{L}(X|P)$ denotes the probability law of the random vector X under the probability measure P, $N(0, J)$ denotes the multivariate normal distribution with mean zero and the covariance matrix J, and u' denotes the transpose of the vector $u \in R^d$.

Androshchuk (2005) gives sufficient conditions for the local asymptotic normality of the family of probability measures $\{P_\theta^\epsilon, \theta \in \Theta\}$ generated by the solutions of the SDE defined by (2.1) as $\epsilon \to 0$. Observe that P_θ^ϵ is a probability measure on the space $C[0, T]$ equipped with the uniform norm.

We will discuss the linear case in more detail in the next and the following sections.

2.3 Parameter estimation for linear SDEs

Le Breton (1998) studied parameter estimation and filtering in a simple linear model driven by fBm. Kleptsyna and Le Breton (2002a) studied parameter estimation problems for the fractional Ornstein–Uhlenbeck-type process discussed in Chapter 1. This is a fractional analogue of the Ornstein–Uhlenbeck process, that is, a continuous time first-order autoregressive process $X = \{X_t, t \geq 0\}$ which is the solution of a one-dimensional homogeneous linear SDE driven by fBm $W^H = \{W_t^H, t \geq 0\}$ with Hurst parameter $H \in [1/2, 1)$. Such a process is the unique Gaussian process satisfying the linear integral equation

$$X_t = \theta \int_0^t X_s ds + \sigma W_t^H, t \geq 0. \tag{2.2}$$

These authors investigate the problem of estimation of the parameters θ and σ^2 based on the observation $\{X_s, 0 \leq s \leq T\}$ and prove that the maximum likelihood estimator $\hat{\theta}_T$ is strongly consistent as $T \to \infty$. We will discuss this problem in more detail in Chapter 3.

We now discuss more general classes of stochastic processes satisfying linear SDEs driven by fBm and study the asymptotic properties of the maximum likelihood and the Bayes estimators for parameters involved in such processes.

Let us consider the SDE

$$dX(t) = [a(t, X(t)) + \theta\, b(t, X(t))]\, dt + \sigma(t)\, dW_t^H, X(0) = 0, t \geq 0 \tag{2.3}$$

where $\theta \in \Theta \subset R$, $W = \{W_t^H, t \geq 0\}$ is fBm with Hurst parameter H and $\sigma(t)$ is a positive non-vanishing function in $[0, \infty)$. In other words, $X = \{X_t, t \geq 0\}$ is a stochastic process satisfying the stochastic integral equation

$$X(t) = \int_0^t [a(s, X(s)) + \theta\, b(s, X(s))]ds + \int_0^t \sigma(s) dW_s^H, t \geq 0. \tag{2.4}$$

Let

$$C(\theta, t) = a(t, X(t)) + \theta\, b(t, X(t)), t \geq 0 \tag{2.5}$$

and assume that the sample paths of the process $\{C(\theta, t)/\sigma(t), t \geq 0\}$ are smooth enough so that the process

$$Q_{H,\theta}(t) = \frac{d}{dw_t^H} \int_0^t k_H(t, s) \frac{C(\theta, s)}{\sigma(s)} ds, t \geq 0 \tag{2.6}$$

is well defined where w_t^H and $k_H(t, s)$ are as defined in Chapter 1. Suppose the sample paths of the process $\{Q_{H,\theta}, 0 \leq t \leq T\}$ belong a.s. to $L^2([0, T], dw_t^H)$. Define

$$Z_t = \int_0^t \frac{k_H(t, s)}{\sigma(s)} dX_s, t \geq 0. \tag{2.7}$$

Then the process $Z = \{Z_t, t \geq 0\}$ is an $(\mathcal{F}_t)$-semimartingale with the decomposition

$$Z_t = \int_0^t Q_{H,\theta}(s)dw_s^H + M_t^H \tag{2.8}$$

where M^H is the fundamental martingale defined in Chapter 1 and the process X admits the representation

$$X_t = \int_0^t K_H^\sigma(t,s)dZ_s \tag{2.9}$$

where the function K_H^σ is as defined by Equation (1.75) in Chapter 1. Let P_θ^T be the measure induced by the process $\{X_t, 0 \leq t \leq T\}$ when θ is the true parameter. Following Theorem 1.20 in Chapter 1, we find that the Radon–Nikodym derivative of P_θ^T with respect to P_0^T is given by

$$\frac{dP_\theta^T}{dP_0^T} = \exp\left[\int_0^T Q_{H,\theta}(s)dZ_s - \frac{1}{2}\int_0^T Q_{H,\theta}^2(s)dw_s^H\right]. \tag{2.10}$$

2.4 Maximum likelihood estimation

We now consider the problem of estimation of the parameter θ based on the observation of the process $X = \{X_t, 0 \leq t \leq T\}$ and study its asymptotic properties as $T \to \infty$.

Strong consistency

Let $L_T(\theta)$ denote the Radon–Nikodym derivative dP_θ^T/dP_0^T. The maximum likelihood estimator (MLE) $\hat{\theta}_T$ is defined by the relation

$$L_T(\hat{\theta}_T) = \sup_{\theta\in\Theta} L_T(\theta). \tag{2.11}$$

We assume that there exists a measurable maximum likelihood estimator. Sufficient conditions can be given for the existence of such an estimator (cf. Lemma 3.1.2 in Prakasa Rao (1987)).

Note that

$$\begin{aligned} Q_{H,\theta}(t) &= \frac{d}{dw_t^H}\int_0^t k_H(t,s)\frac{C(\theta,s)}{\sigma(s)}ds \\ &= \frac{d}{dw_t^H}\int_0^t k_H(t,s)\frac{a(s,X(s))}{\sigma(s)}ds + \theta\frac{d}{dw_t^H}\int_0^t k_H(t,s)\frac{b(s,X(s))}{\sigma(s)}ds \\ &= J_1(t) + \theta J_2(t) \ \ \text{(say)}. \end{aligned} \tag{2.12}$$

Then

$$\log L_T(\theta) = \int_0^T (J_1(t) + \theta J_2(t))dZ_t - \frac{1}{2}\int_0^T (J_1(t) + \theta J_2(t))^2 dw_t^H \tag{2.13}$$

and the likelihood equation is given by

$$\int_0^T J_2(t)dZ_t - \int_0^T (J_1(t) + \theta J_2(t))J_2(t)dw_t^H = 0. \tag{2.14}$$

Hence the MLE $\hat{\theta}_T$ of θ is given by

$$\hat{\theta}_T = \frac{\int_0^T J_2(t)dZ_t - \int_0^T J_1(t)J_2(t)dw_t^H}{\int_0^T J_2^2(t)dw_t^H}. \tag{2.15}$$

Let θ_0 be the true parameter. Using the fact that

$$dZ_t = (J_1(t) + \theta_0 J_2(t))\, dw_t^H + dM_t^H, \tag{2.16}$$

it can be shown that

$$\frac{dP_\theta^T}{dP_{\theta_0}^T} = \exp\left[(\theta - \theta_0)\int_0^T J_2(t)dM_t^H - \frac{1}{2}(\theta - \theta_0)^2 \int_0^T J_2^2(t)dw_t^H\right]. \tag{2.17}$$

Following this representation of the Radon–Nikodym derivative, we obtain that

$$\hat{\theta}_T - \theta_0 = \frac{\int_0^T J_2(t)dM_t^H}{\int_0^T J_2^2(t)\, dw_t^H}. \tag{2.18}$$

Note that the quadratic variation $\langle Z \rangle$ of the process Z is the same as the quadratic variation $\langle M^H \rangle$ of the martingale M^H which in turn is equal to w^H. This follows from the relations (1.79) and (1.73) in Chapter 1. Hence we obtain that

$$[w_T^H]^{-1} \lim_n \sum [Z_{t_{i+1}^{(n)}} - Z_{t_i^{(n)}}]^2 = 1 \text{ a.s. } [P_{\theta_0}]$$

where

$$0 = t_0^{(n)} < t_1^{(n)} < \cdots < t_{n-1}^{(n)} = T$$

is a partition of the interval $[0, T]$ such that

$$\sup_{0 \le i \le n-1} |t_{i+1}^{(n)} - t_i^{(n)}| \to 0$$

as $n \to \infty$. If the function $\sigma(.)$ is an unknown constant σ, the above property can be used to obtain a strongly consistent estimator of σ^2 based on the continuous observation of the process X over the interval $[0, T]$. Hereafter we assume that the non random function $\sigma(t)$ is known.

We now discuss the problem of estimation of the parameter θ on the basis of the observation of the process X or equivalently of the process Z in the interval $[0, T]$.

Strong consistency

Theorem 2.1: The MLE $\hat{\theta}_T$ is strongly consistent, that is,

$$\hat{\theta}_T \to \theta_0 \text{ a.s. } [P_{\theta_0}] \text{ as } T \to \infty \tag{2.19}$$

provided

$$\int_0^T J_2^2(t)dw_t^H \to \infty \text{ a.s. } [P_{\theta_0}] \text{ as } T \to \infty. \tag{2.20}$$

Proof: This theorem follows by observing that the process

$$R_T \equiv \int_0^T J_2(t)dM_t^H, t \geq 0 \tag{2.21}$$

is a local martingale with the quadratic variation process

$$\langle R \rangle_T = \int_0^T J_2^2(t)dw_t^H \tag{2.22}$$

and applying the strong law of large numbers (cf. Liptser (1980), Prakasa Rao (1999b), p. 61) under the condition (2.20) stated above.

Remarks: For the case of the fractional Ornstein–Uhlenbeck-type process investigated in Kleptsyna and Le Breton (2002a), it can be checked that the condition stated in Equation (2.20) holds and hence the MLE $\hat{\theta}_T$ is strongly consistent as $T \to \infty$.

Limiting distribution

We now discuss the limiting distribution of the MLE $\hat{\theta}_T$ as $T \to \infty$.

Theorem 2.2: Assume that the functions $b(t, s)$ and $\sigma(t)$ are such that the process $\{R_t, t \geq 0\}$ is a local continuous martingale and that there exists a norming function $I_t, t \geq 0$, such that

$$I_T^2 \langle R \rangle_T = I_T^2 \int_0^T J_2^2(t)dw_t^H \to \eta^2 \text{ in probability as } T \to \infty \tag{2.23}$$

where $I_T \to 0$ a.s. as $T \to \infty$ and η is a random variable such that $P(\eta > 0) = 1$. Then

$$(I_T R_T, I_T^2 \langle R \rangle_T) \to (\eta Z, \eta^2) \text{ in law as } T \to \infty \tag{2.24}$$

where the random variable Z has the standard normal distribution and the random variables Z and η are independent.

Proof: This theorem follows as a consequence of the central limit theorem for local martingales (cf. Theorem 1.49 and Remark 1.47 in Prakasa Rao (1999b), p. 65).

Observe that

$$I_T^{-1}(\hat{\theta}_T - \theta_0) = \frac{I_T R_T}{I_T^2 \langle R_T \rangle}. \tag{2.25}$$

Applying Theorem 2.2, we obtain the following result.

Theorem 2.3: Suppose the conditions stated in Theorem 2.2 hold. Then

$$I_T^{-1}(\hat{\theta}_T - \theta_0) \to \frac{Z}{\eta} \text{ in law as } t \to \infty \tag{2.26}$$

where the random variable Z has the standard normal distribution and the random variables Z and η are independent.

Remarks: If the random variable η is a constant with probability one, then the limiting distribution of the MLE is normal with mean zero and variance η^{-2}. Otherwise it is a mixture of the normal distribution with mean zero and variance η^{-2} with the mixing distribution as that of η.

2.5 Bayes estimation

Suppose that the parameter space Θ is open and Λ is a prior probability measure on the parameter space Θ. Further suppose that Λ has the density $\lambda(.)$ with respect to the Lebesgue measure and the density function is continuous and positive in an open neighborhood of θ_0, the true parameter. Let

$$\alpha_T \equiv I_T R_T = I_T \int_0^T J_2(t) dM_t^H \tag{2.27}$$

and

$$\beta_T \equiv I_T^2 \langle R \rangle_T = I_T^2 \int_0^T J_2^2(t) dw_t^H. \tag{2.28}$$

We saw earlier that the MLE satisfies the relation

$$\alpha_T = (\hat{\theta}_T - \theta_0) I_T^{-1} \beta_T. \tag{2.29}$$

The posterior density of θ based on the observation $X^T \equiv \{X_s, 0 \le s \le T\}$ is given by

$$p(\theta|X^T) = \frac{dP_\theta^T}{dP_{\theta_0}^T}\lambda(\theta) \Bigg/ \int_\Theta \frac{dP_\theta^T}{dP_{\theta_0}^T}\lambda(\theta)d\theta. \tag{2.30}$$

Let $t = I_T^{-1}(\theta - \hat{\theta}_T)$ and define

$$p^*(t|X^T) = I_T \; p(\hat{\theta}_T + tI_T|X^T). \tag{2.31}$$

Then the function $p^*(t|X^T)$ is the posterior density of the transformed variable $t = I_T^{-1}(\theta - \hat{\theta}_T)$. Let

$$\nu_T(t) \equiv \frac{dP_{\hat{\theta}_T + tI_T}/dP_{\theta_0}}{dP_{\hat{\theta}_T}/dP_{\theta_0}} = \frac{dP_{\hat{\theta}_T + tI_T}}{dP_{\hat{\theta}_T}} \text{ a.s.} \tag{2.32}$$

and

$$C_T = \int_{-\infty}^{\infty} \nu_T(t)\lambda(\hat{\theta}_T + tI_T)dt. \tag{2.33}$$

It can be checked that

$$p^*(t|X^T) = C_T^{-1}\nu_T(t)\lambda(\hat{\theta}_T + tI_T). \tag{2.34}$$

Furthermore, Equations (2.25) and (2.29)–(2.34) imply that

$$\begin{aligned}
\log \nu_T(t) &= I_T^{-1}\alpha_T[(\hat{\theta}_T + tI_T - \theta_0) - (\hat{\theta}_T - \theta_0)] \\
&\quad - \frac{1}{2} I_T^{-2}\beta_T[(\hat{\theta}_T + tI_T - \theta_0)^2 - (\hat{\theta}_T - \theta_0)^2] \\
&= t\alpha_T - \frac{1}{2}t^2\beta_T - t\beta_T I_T^{-1}(\hat{\theta}_T - \theta_0) \\
&= -\frac{1}{2}\beta_T t^2
\end{aligned} \tag{2.35}$$

in view of Equation (2.29).

Suppose that the convergence of the condition in Equation (2.23) holds a.s. under the measure P_{θ_0} and the limit is a constant $\eta^2 > 0$ with probability one. For convenience, we write $\beta = \eta^2$. Then

$$\beta_T \to \beta \text{ a.s. } [P_{\theta_0}] \text{ as } T \to \infty. \tag{2.36}$$

It is obvious that

$$\lim_{T\to\infty} \nu_T(t) = \exp\left[-\frac{1}{2}\beta t^2\right] \text{ a.s. } [P_{\theta_0}] \tag{2.37}$$

and, for any $0 < \varepsilon < \beta$,

$$\log \nu_T(t) \le -\frac{1}{2}t^2(\beta - \varepsilon) \tag{2.38}$$

for every t for T sufficiently large. Furthermore, for every $\delta > 0$, there exists $\varepsilon' > 0$ such that

$$\sup_{|t| > \delta I_T^{-1}} \nu_T(t) \le \exp\left[-\frac{1}{4}\varepsilon' I_T^{-2}\right] \tag{2.39}$$

for T sufficiently large.

Suppose $H(t)$ is a nonnegative measurable function such that, for some $0 < \varepsilon < \beta$,

$$\int_{-\infty}^{\infty} H(t) \exp\left[-\frac{1}{2}t^2(\beta - \varepsilon)\right] dt < \infty. \tag{2.40}$$

Suppose the MLE $\hat{\theta}_T$ is strongly consistent, that is,

$$\hat{\theta}_T \to \theta_0 \text{ a.s. } \quad [P_{\theta_0}] \text{ as } T \to \infty. \tag{2.41}$$

For any $\delta > 0$, consider

$$\begin{aligned}
&\int_{|t|\le\delta I_T^{-1}} H(t) \left|\nu_T(t)\lambda(\hat{\theta}_T + tI_T) - \lambda(\theta_0)\exp\left(-\frac{1}{2}\beta t^2\right)\right| dt \\
&\le \int_{|t|\le\delta I_T^{-1}} H(t)\lambda(\theta_0) \left|\nu_T(t) - \exp\left(-\frac{1}{2}\beta t^2\right)\right| dt \\
&\quad + \int_{|t|\le\delta I_T^{-1}} H(t)\nu_T(t)|\lambda(\theta_0) - \lambda(\hat{\theta}_T + tI_T)|\, dt \\
&= A_T + B_T \text{ (say).}
\end{aligned} \tag{2.42}$$

It is clear that, for any $\delta > 0$,

$$A_T \to 0 \text{ a.s. } [P_{\theta_0}] \text{ as } T \to \infty \tag{2.43}$$

by the dominated convergence theorem in view of the inequality in (2.38), Equation (2.37) and the condition in Equation (2.40). On the other hand, for T sufficiently large,

$$0 \le B_T \le \sup_{|\theta-\theta_0|\le\delta} |\lambda(\theta) - \lambda(\theta_0)| \int_{|t|\le\delta I_T^{-1}} H(t)\exp\left[-\frac{1}{2}t^2(\beta - \varepsilon)\right] dt \tag{2.44}$$

since $\hat{\theta}_T$ is strongly consistent and $I_T^{-1} \to \infty$ a.s. as $T \to \infty$. The last term on the right of the above inequality can be made smaller than any given $\rho > 0$ by choosing δ sufficiently small in view of the continuity of $\lambda(.)$ at θ_0. Combining these remarks with Equations (2.43) and (2.44), we obtain the following lemma.

Lemma 2.4: Suppose conditions (2.36), (2.40) and (2.41) hold. Then there exists $\delta > 0$ such that

$$\lim_{T\to\infty} \int_{|t|\le \delta I_T^{-1}} H(t) \left| \nu_T(t)\lambda(\hat{\theta}_T + tI_T) - \lambda(\theta_0)\exp\left(-\frac{1}{2}\beta t^2\right)\right| dt = 0. \quad (2.45)$$

For any $\delta > 0$, consider

$$\begin{aligned}
&\int_{|t|> \delta I_T^{-1}} H(t) \left| \nu_T(t)\lambda(\hat{\theta}_T + tI_T) - \lambda(\theta_0)\exp\left(-\frac{1}{2}\beta t^2\right)\right| dt \\
&\quad \le \int_{|t|> \delta I_T^{-1}} H(t)\nu_T(t)\lambda(\hat{\theta}_T + tI_T)\, dt + \int_{|t|> \delta I_T^{-1}} H(t)\lambda(\theta_0)\exp\left(-\frac{1}{2}\beta t^2\right) dt \\
&\quad \le \exp\left(-\frac{1}{4}\varepsilon' I_T^{-2}\right) \int_{|t|> \delta I_T^{-1}} H(t)\lambda(\hat{\theta}_T + tI_T)\, dt \\
&\qquad + \lambda(\theta_0) \int_{|t|> \delta I_T^{-1}} H(t)\exp\left(-\frac{1}{2}\beta t^2\right) dt \\
&\quad = U_T + V_T \text{ (say)}.
\end{aligned} \quad (2.46)$$

Suppose the following condition holds for every $\varepsilon > 0$ and $\delta > 0$:

$$\exp(-\varepsilon I_T^{-2}) \int_{|u|>\delta} H(uI_T^{-1})\lambda(\hat{\theta}_T + u)du \to 0 \text{ a.s. } [P_{\theta_0}] \text{ as } T \to \infty. \quad (2.47)$$

It is clear that, for every $\delta > 0$,

$$V_T \to 0 \text{ a.s. } [P_{\theta_0}] \text{ as } T \to \infty \quad (2.48)$$

in view of the condition stated in (2.40) and the fact that $I_T^{-1} \to \infty$ a.s. $[P_{\theta_0}]$ as $T \to \infty$. The condition stated in (2.47) implies that

$$U_T \to 0 \text{ a.s. } [P_{\theta_0}] \text{ as } T \to \infty \quad (2.49)$$

for every $\delta > 0$. Hence we have the following lemma.

Lemma 2.5: Suppose that conditions (2.36), (2.40) and (2.41) hold. Then for every $\delta > 0$,

$$\lim_{T\to\infty} \int_{|t| > \delta I_T^{-1}} H(t) \left| \nu_T(t)\lambda(\hat{\theta}_T + tI_T) - \lambda(\theta_0)\exp\left(-\frac{1}{2}\beta t^2\right)\right| dt = 0. \quad (2.50)$$

Lemmas 2.4 and 2.5 together prove that

$$\lim_{T\to\infty}\int_{|t|>\delta I_T^{-1}} H(t)\left|\nu_T(t)\lambda(\hat{\theta}_T+tI_T)-\lambda(\theta_0)\exp\left(-\frac{1}{2}\beta t^2\right)\right| dt = 0. \quad (2.51)$$

Let $H(t)\equiv 1$ in (2.51). Note that

$$C_T \equiv \int_{-\infty}^{\infty} \nu_T(t)\lambda(\hat{\theta}_T+tI_T)dt.$$

Relation (2.51) implies that

$$C_T \to \lambda(\theta_0)\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}\beta t^2\right) dt = \lambda(\theta_0)\left(\frac{\beta}{2\pi}\right)^{-1/2} \quad \text{a.s. } [P_{\theta_0}] \quad (2.52)$$

as $T\to\infty$. Furthermore,

$$\begin{aligned}\int_{-\infty}^{\infty} H(t)&\left|p^*(t|X^T)-\left(\frac{\beta}{2\pi}\right)^{1/2}\exp\left(-\frac{1}{2}\beta t^2\right)\right| dt\\ &\leq C_T^{-1}\int_{-\infty}^{\infty} H(t)\left|\nu_T(t)\lambda(\hat{\theta}_T+tI_T)-\lambda(\theta_0)\exp\left(-\frac{1}{2}\beta t^2\right)\right| dt\\ &\quad+\int_{-\infty}^{\infty} H(t)\left|C_T^{-1}\lambda(\theta_0)-\left(\frac{\beta}{2\pi}\right)^{1/2}\right|\exp\left(-\frac{1}{2}\beta t^2\right) dt. \quad (2.53)\end{aligned}$$

The last two terms tend to zero a.s. $[P_{\theta_0}]$ by Equations (2.51) and (2.52). Hence we have the following theorem which is an analogue of the Bernstein–von Mises theorem proved in Prakasa Rao (1981) for a class of processes satisfying a linear SDE driven by the standard Wiener process.

Theorem 2.6: Let the assumptions (2.36), (2.40), (2.41) and (2.47) hold where $\lambda(.)$ is a prior density which is continuous and positive in an open neighborhood of θ_0, the true parameter. Then

$$\lim_{T\to\infty}\int_{-\infty}^{\infty} H(t)\left|p^*(t|X^T)-\left(\frac{\beta}{2\pi}\right)^{1/2}\exp\left(-\frac{1}{2}\beta t^2\right)\right| dt = 0 \text{ a.s. } [P_{\theta_0}]. \quad (2.54)$$

As a consequence of the above theorem, we obtain the following result by choosing $H(t)=|t|^m$, for any integer $m\geq 0$.

Theorem 2.7: Assume that the following conditions hold:

$$\text{(C1)}\quad \hat{\theta}_T \to \theta_0 \text{ a.s. } [P_{\theta_0}] \text{ as } T\to\infty, \quad (2.55)$$

$$\text{(C2)}\quad \beta_T \to \beta > 0 \text{ a.s. } [P_{\theta_0}] \text{ as } T\to\infty. \quad (2.56)$$

Further suppose that (C3) $\lambda(.)$ is a prior probability density on Θ which is continuous and positive in an open neighborhood of θ_0, the true parameter and

$$\text{(C4)} \quad \int_{-\infty}^{\infty} |\theta|^m \lambda(\theta) d\theta < \infty \tag{2.57}$$

for some integer $m \geq 0$. Then

$$\lim_{T\to\infty} \int_{-\infty}^{\infty} |t|^m \left| p^*(t|X^T) - \left(\frac{\beta}{2\pi}\right)^{1/2} \exp\left(-\frac{1}{2}\beta t^2\right) \right| dt = 0 \text{ a.s. } [P_{\theta_0}]. \tag{2.58}$$

In particular, choosing $m = 0$, we obtain that

$$\lim_{T\to\infty} \int_{-\infty}^{\infty} \left| p^*(t|X^T) - \left(\frac{\beta}{2\pi}\right)^{1/2} \exp\left(-\frac{1}{2}\beta t^2\right) \right| dt = 0 \text{ a.s. } [P_{\theta_0}] \tag{2.59}$$

whenever the conditions (C1), (C2) and (C3) hold. This is the analogue of the Bernstein–von Mises theorem for a class of diffusion processes proved in Prakasa Rao (1981) and it shows the asymptotic convergence in L_1-mean of the posterior density to the probability density function of a normal distribution with mean zero and suitable variance.

As a corollary to Theorem 2.7, we also obtain that the conditional expectation, under P_{θ_0}, of $[I_T^{-1}(\hat{\theta}_T - \theta)]^m$ converges to the corresponding mth absolute moment of the normal distribution with mean zero and variance β^{-1}.

We define a *regular Bayes estimator* of θ, corresponding to a prior probability density $\lambda(\theta)$ and the loss function $L(\theta, \phi)$, based on the observation X^T, as an estimator which minimizes the posterior risk

$$B_T(\phi) \equiv \int_{-\infty}^{\infty} L(\theta, \phi) p(\theta|X^T) d\theta \tag{2.60}$$

over all estimators ϕ of θ. Here $L(\theta, \phi)$ is a loss function defined on $\Theta \times \Theta$.

Suppose there exists a measurable regular Bayes estimator $\tilde{\theta}_T$ for the parameter θ (cf. Theorem 3.1.3 in Prakasa Rao (1987)). Suppose that the loss function $L(\theta, \phi)$ satisfies the following conditions:

$$L(\theta, \phi) = \ell(|\theta - \phi|) \geq 0 \tag{2.61}$$

and the function $\ell(t)$ is nondecreasing for $t \geq 0$. An example of such a loss function is $L(\theta, \phi) = |\theta - \phi|$. Suppose there exist nonnegative functions $J(t)$, $K(t)$ and $G(t)$ such that

$$\text{(D1)} \quad J(t)\ell(tI_T) \leq G(t) \text{ for all } T \geq 0, \tag{2.62}$$

$$\text{(D2)} \quad J(t)\ell(tI_T) \to K(t) \;\text{ as }\; T \to \infty \tag{2.63}$$

uniformly on bounded intervals of t. Further suppose that the function

$$\text{(D3)} \quad \int_{-\infty}^{\infty} K(t+h) \exp\left(-\frac{1}{2}\beta t^2\right) dt \tag{2.64}$$

has a strict minimum at $h = 0$, and (D4) the function $G(t)$ satisfies conditions similar to those in (2.40) and (2.47). Then we have the following result giving the asymptotic properties of the Bayes risk of the estimator $\tilde{\theta}_T$.

Theorem 2.8: Suppose the conditions (C1) to (C3) in Theorem 2.7 and the conditions (D1) to (D4) stated above hold. Then

$$I_T^{-1}(\tilde{\theta}_T - \hat{\theta}_T) \to 0 \;\text{ a.s. }\; [P_{\theta_0}] \;\text{ as }\; T \to \infty \tag{2.65}$$

and

$$\begin{aligned}\lim_{T\to\infty} R(T)B_T(\tilde{\theta}_T) &= \lim_{T\to\infty} R(T)B_T(\hat{\theta}_T) \\ &= \left(\frac{\beta}{2\pi}\right)^{1/2} \int_{-\infty}^{\infty} K(t) \exp\left(-\frac{1}{2}\beta t^2\right) dt \;\text{ a.s. }\; [P_{\theta_0}].\end{aligned} \tag{2.66}$$

We omit the proof of this theorem as it is similar to the proof of Theorem 4.1 in Borwanker *et al*. (1971).

We observed earlier that

$$I_T^{-1}(\hat{\theta}_T - \theta_0) \to N(0, \beta^{-1}) \;\text{ in law }\; \text{as }\; T \to \infty. \tag{2.67}$$

As a consequence of Theorem 2.8, we obtain that

$$\tilde{\theta}_T \to \theta_0 \;\text{ a.s. }\; [P_{\theta_0}] \;\text{ as }\; T \to \infty \tag{2.68}$$

and

$$I_T^{-1}(\tilde{\theta}_T - \theta_0) \to N(0, \beta^{-1}) \;\text{ in law }\; \text{as }\; T \to \infty. \tag{2.69}$$

In other words, the Bayes estimator is asymptotically normal and has asymptotically the same distribution as the MLE. The asymptotic Bayes risk of the estimator is given by Theorem 2.8. The results discussed in this section are due to Prakasa Rao (2003).

2.6 Berry–Esseen-type bound for MLE

Hereafter we assume that the random variable η in (2.23) is a positive constant with probability one. Hence

$$I_T^{-1}(\hat{\theta}_T - \theta_0) \to N(0, \eta^{-2}) \text{ in law as } T \to \infty \tag{2.70}$$

where $N(0, \eta^{-2})$ denotes the Gaussian distribution with mean zero and variance η^{-2}.

We will now study the rate of convergence of the asymptotic distribution of the MLE $\hat{\theta}_T$. The result studied here is from Prakasa Rao (2005a).

Suppose there exist non random positive functions α_T decreasing to zero and ε_T decreasing to zero such that

$$\alpha_T^{-1}\varepsilon_T^2 \to \infty \text{ as } T \to \infty, \tag{2.71}$$

and

$$\sup_{\theta \in \Theta} P_\theta^T[|\alpha_T \langle R \rangle_T - 1| \geq \varepsilon_T] = O(\varepsilon_T^{1/2}) \tag{2.72}$$

where the process $\{R_t, t \geq 0\}$ is as defined in (2.21). Note that the process $\{R_t, t \geq 0\}$ is a locally square integrable continuous martingale. From the results on the representation of locally square integrable continuous martingales (cf. Ikeda and Watanabe (1981), Chapter II, Theorem 7.2), it follows that there exists a standard Wiener process $\{\tilde{W}(t), t \geq 0\}$ adapted to $(\mathcal{F}_t)$ such that $R_t = \tilde{W}(\langle R \rangle_t), t \geq 0$. In particular

$$R_T \alpha_T^{1/2} = \tilde{W}(\langle R \rangle_T \alpha_T) \text{ a.s. } [P] \tag{2.73}$$

for all $T \geq 0$.

We use the following lemmas in the sequel.

Lemma 2.9: Let $(\Omega, \mathcal{F}, P)$ be a probability space and f and g be $\mathcal{F}$-measurable functions. Further suppose that $g \geq 0$. Then, for any $\varepsilon > 0$,

$$\sup_x \left| P\left(\omega : \frac{f(\omega)}{g(\omega)} \leq x\right) - \Phi(x)\right| \leq \sup_y |P(\omega : f(\omega) \leq y) - \Phi(y)| + P(\omega : |g(\omega) - 1| > \varepsilon) + \varepsilon \tag{2.74}$$

where $\Phi(.)$ is the distribution function of the standard Gaussian distribution.

Proof: See Michel and Pfanzagl (1971).

Lemma 2.10: Let $\{W(t), t \geq 0\}$ be a standard Wiener process and V be a nonnegative random variable. Then, for every $x \in R$ and $\varepsilon > 0$,

$$|P(W(V) \leq x) - \Phi(x)| \leq (2\varepsilon)^{1/2} + P(|V - 1| > \varepsilon). \tag{2.75}$$

Proof: See Hall and Heyde (1980), p. 85.

Let us fix $\theta \in \Theta$. It is clear from the earlier remarks that

$$R_T = \langle R \rangle_T (\hat{\theta}_T - \theta) \tag{2.76}$$

under P_θ^T measure. Then it follows from Lemmas 2.9 and 2.10 that

$$\begin{aligned}
&|P_\theta^T[\alpha_T^{-1/2}(\hat{\theta}_T - \theta) \le x] - \Phi(x)| \\
&= \left| P_\theta^T \left[\frac{R_T}{\langle R \rangle_T} \alpha_T^{-1/2} \le x \right] - \Phi(x) \right| \\
&= \left| P_\theta^T \left[\frac{R_T/\alpha_T^{-1/2}}{\langle R \rangle_T/\alpha_T^{-1}} \le x \right] - \Phi(x) \right| \\
&\le \sup_x |P_\theta^T[R_T \alpha_T^{1/2} \le x] - \Phi(x)| + P_\theta^T[|\alpha_T \langle R \rangle_T - 1| \ge \varepsilon_T] + \varepsilon_T \\
&= \sup_y |P(\tilde{W}(\langle R \rangle_T \alpha_T) \le y) - \Phi(y)| + P_\theta^T[|\alpha_T \langle R \rangle_T - 1| \ge \varepsilon_T] + \varepsilon_T \\
&\le (2\varepsilon_T)^{1/2} + 2P_\theta^T[|\alpha_T \langle R \rangle_T - 1| \ge \varepsilon_T] + \varepsilon_T.
\end{aligned} \tag{2.77}$$

It can be checked that the bound obtained above is of the order $O(\varepsilon_T^{1/2})$ under the condition (2.72) and it is uniform in $\theta \in \Theta$. Hence we have the following result.

Theorem 2.11: Under conditions (2.71) and (2.72),

$$\begin{aligned}
&\sup_{\theta \in \Theta} \sup_{x \in R} |P_\theta^T[\alpha_T^{-1/2}(\hat{\theta}_T - \theta) \le x] - \Phi(x)| \\
&\quad \le (2\varepsilon_T)^{1/2} + 2P_\theta^T[|\alpha_T \langle R \rangle_T - 1| \ge \varepsilon_T] + \varepsilon_T = O(\varepsilon_T^{1/2}).
\end{aligned} \tag{2.78}$$

As a consequence of this result, we have the following theorem giving the rate of convergence of the MLE $\hat{\theta}_T$.

Theorem 2.12: Suppose conditions (2.71) and (2.72) hold. Then there exists a constant $c > 0$ such that for every $d > 0$,

$$\sup_{\theta \in \Theta} P_\theta^T[|\hat{\theta}_T - \theta| \ge d] \le c\varepsilon_T^{1/2} + 2P_\theta^T[|\alpha_T \langle R \rangle_T - 1| \ge \varepsilon_T] = O(\varepsilon_T^{1/2}). \tag{2.79}$$

Proof: Observe that

$$\begin{aligned}
&\sup_{\theta \in \Theta} P_\theta^T[|\hat{\theta}_T - \theta| \ge d] \\
&\quad \le \sup_{\theta \in \Theta} |P_\theta^T[\alpha_T^{-1/2}(\hat{\theta}_T - \theta) \ge d\alpha_T^{-1/2}] - 2(1 - \Phi(d\alpha_T^{-1/2}))|
\end{aligned}$$

$$+2(1-\Phi(d\alpha_T^{-1/2}))$$
$$\leq (2\varepsilon_T)^{1/2} + 2\sup_{\theta\in\Theta} P_\theta^T[|\alpha_T\langle R\rangle_T - 1| \geq \varepsilon_T] + \varepsilon_T$$
$$+2d^{-1/2}\alpha_T^{1/2}(2\pi)^{-1/2}\exp\left(-\frac{1}{2}\alpha_T^{-1}d^2\right) \tag{2.80}$$

by Theorem 2.11 and the inequality

$$1-\Phi(x) < \frac{1}{x\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right) \tag{2.81}$$

for all $x > 0$ (cf. Feller (1968), p. 175). Since

$$\alpha_T^{-1}\varepsilon_T^2 \to \infty \text{ as } T \to \infty,$$

by condition (2.71), it follows that

$$\sup_{\theta\in\Theta} P_\theta^T[|\hat{\theta}_T - \theta| \geq d] \leq c\varepsilon_T^{1/2} + 2\sup_{\theta\in\Theta} P_\theta^T[|\alpha_T\langle R\rangle_T - 1| \geq \varepsilon_T] \tag{2.82}$$

for some constant $c > 0$ and the last term is of the order $O(\varepsilon_T^{1/2})$ by condition (2.72). This proves Theorem 2.12.

2.7 ϵ-upper and lower functions for MLE

Hereafter we assume that the non random function $\sigma(t)$ in (2.3) is a known constant σ. Without loss of generality, we assume that $\sigma = 1$.

Let

$$R_t \equiv \int_0^T J_2(t)dM_t^H, \; t \geq 0. \tag{2.83}$$

The process $\{R_t, t \geq 0\}$ is a continuous local martingale with the quadratic variation process

$$\langle R\rangle_T = \int_0^T J_2^2(t)dw_t^H \equiv I_T(\theta_0) \text{ (say)}. \tag{2.84}$$

Assume that there exist non-random positive functions $A_T \uparrow \infty$ and $\epsilon_T \downarrow 0$ as $T \to \infty$ such that:

(A_1) $A_T\epsilon_T^2 \to \infty$;

(A_2) $\lim_{T\to\infty} \frac{I_T(\theta_0)}{A_T} = 1$ a.s. $[P_{\theta_0}]$;

(A_3) $P_{\theta_0}^T\left[\left|\frac{I_T(\theta_0)}{A_T}-1\right| \geq \epsilon_T\right] = O(\epsilon_T^{1/2})$;

(A_4) $\int_3^\infty \frac{\log\log A_T}{A_T}\epsilon_T^{1/2}dT < \infty$.

From the results on the representation of locally continuous square integrable martingales (cf. Ikeda and Watanabe (1981), Chapter II, Theorem 7.2), it follows that there exists a standard Wiener process $\{W^*(t), t \geq 0\}$ adapted to the filtration $\{\mathcal{F}_t, t \geq 0\}$ such that

$$R_t = W^*(I_t(\theta_0)),\ t \geq 0 \ \text{ a.s. } [P_{\theta_0}].$$

Observe that

$$(\hat{\theta}_T - \theta_0) = \frac{R_T}{I_T(\theta_0)}$$

and hence

$$A_T^{1/2}(\hat{\theta}_T - \theta_0) = \frac{W_T^*(I_T(\theta_0))/A_T^{1/2}}{I_T(\theta_0)/A_T} \text{ a.s. } [P_{\theta_0}]. \tag{2.85}$$

Theorem 2.13: Suppose conditions (A_1) and(A_3) hold. Then

$$\sup_x |P_{\theta_0}^T(A_T^{1/2}(\hat{\theta}_T - \theta_0) \leq x) - \Phi(x)|$$
$$\leq \sqrt{2\epsilon_T} + 2P_{\theta_0}^T\left[\left|\frac{I_T(\theta_0)}{A_T}-1\right| \geq \epsilon_T\right] + \epsilon_T = O(\epsilon_T^{1/2}) \tag{2.86}$$

where $\Phi(.)$ is the standard normal distribution function.

This theorem is a slight variant of Theorem 2.12 and is due to Prakasa Rao (2005a).

Let $h(t)$ be a nonnegative, nondecreasing function converging to infinity as $t \to \infty$. The function $h(t)$ is said to belong to the ϵ-*upper class* of a stochastic process $\{Y(t), t \geq 0\}$, if

$$P(Y(t) > (1+\epsilon)h(t) \text{ infinitely often as } t \to \infty) = 0.$$

The function $h(t)$ is said to belong to the *lower class* of a stochastic process $\{Y(t), t \geq 0\}$, if

$$P(Y(t) > h(t) \text{ infinitely often as } t \to \infty) = 1.$$

We now study ϵ-upper and lower class functions for a normalized process obtained by taking the difference between the MLE of the trend parameter and the true trend parameter for linear SDEs driven by fBm. The results in this section are due to Mishra and Prakasa Rao (2008).

Theorem 2.14: Suppose conditions (A_1), (A_3) and (A_4) hold. Let $h(.)$ be a nonnegative, nondecreasing function converging to infinity as $t \to \infty$. Then the integrals

$$\int_3^\infty \frac{\log\log A_T}{A_T} P_{\theta_0}^T[A_T^{1/2}(\hat{\theta}_T - \theta_0) > h(A_T)]dT$$

and

$$\int_3^\infty \frac{\log\log A_T}{A_T h(A_T)} \exp(-h^2(A_T)/2)dT$$

are either both convergent or both divergent.

Theorem 2.15: Suppose conditions (A_1), (A_3) and (A_4) hold. Let $h(.)$ be a nonnegative, nondecreasing function converging to infinity as $t \to \infty$. Let $0 < \epsilon < 1$. Then

$$P_{\theta_0}^T[A_T^{1/2}(\hat{\theta}_T - \theta_0) > (1+\epsilon)h(A_T) \text{ infinitely often as } T \to \infty] = 0$$

if the integral

$$K(h) \equiv \int_3^\infty \frac{h(A_T)}{A_T} \exp(-h^2(A_T)/2)dT$$

is convergent and

$$P_{\theta_0}^T[A_T^{1/2}(\hat{\theta}_T - \theta_0) > h(A_T) \text{ infinitely often as } T \to \infty] = 1$$

if the integral $K(h)$ is divergent.

Theorem 2.16: Suppose conditions (A_1), (A_3) and (A_4) hold. Let $h(.)$ be a nonnegative, nondecreasing function converging to infinity as $t \to \infty$. Let $C \geq 0$. Then the function $h(.)$ belongs to the ϵ-upper or lower class of the process $\{A_T^{1/2}(\hat{\theta}_T - \theta_0), T \geq 0\}$ according as the integral

$$\int_3^\infty \frac{\log\log A_T}{A_T h(A_T)} \exp(-h^2(A_T)/2)\left(1 + \frac{C}{\log\log A_T}\right) dT$$

is convergent or divergent.

We use the following lemmas to prove Theorems 2.14 to 2.16.

Lemma 2.17: Let $h(.)$ be a nonnegative, nondecreasing function converging to infinity as $t \to \infty$. Let $C \geq 0$. Then the integrals

$$\int_3^\infty \frac{h(A_T)}{A_T} \exp(-h^2(A_T)/2)dT$$

and

$$\int_3^\infty \frac{\log\log A_T}{A_T h(A_T)} \exp(-h^2(A_T)/2) \left(1 + \frac{C}{\log\log A_T}\right) \; dT$$

are either both convergent or both divergent.

The above lemma is a continuous version of a lemma proved in Davis (1969).

Lemma 2.18: Let J denote a class of functions $h(.)$ which are continuous, nondecreasing and converging to infinity as $t \to \infty$. Let $g(.)$ be a non-increasing function from $[1, \infty)$ to $[0, \infty)$ and $f(.)$ be a measurable function from $[1, \infty)$ to $[0, \infty)$. For $h(.) \in J$, define $F(h) = \int_1^\infty g(h(t))f(t)\, dt$ which may be finite or infinite. Assume that:

(a_1) for every $h \in J$, and $A > 1$,

$$\int_1^A g(h(t))f(t)dt < \infty;$$

(a_2) there exists $h_1 \in J$ and $h_2 \in J$ such that $h_1 \le h_2, F(h_1) = \infty$, $F(h_2) < \infty$ and

$$\lim_{A\to\infty} g(h_1(A)) \int_1^A f(t)dt = \infty.$$

Define $\hat{h} = \min[\max(h, h_1), h_2]$. Then, for $h \in J$,

$$\text{(i) } F(h) < \infty \Rightarrow \hat{h} \le h \text{ near infinity} \quad \text{and } F(\hat{h})\infty$$

and

$$\text{(ii) } F(h) = \infty \Rightarrow F(\hat{h}) = \infty.$$

For the proof of Lemma 2.18, see Lemma 2.3 in Jain *et al*. (1975). Observe that $h_1 \le \hat{h} \le h_2$. This lemma allows us to consider only those functions $h \in J$ which satisfy the condition $h_1 \le h \le h_2$ for proving the ϵ-upper and lower class results in Theorem 2.15.

Remarks: Applying Lemma 2.17 and using Theorem 2.15, we find that the integrals

$$\int_3^\infty \frac{\log\log A_T}{A_T h(A_T)} \exp(-h^2(A_T)/2) dT$$

and

$$\int_3^\infty \frac{\log\log A_T}{A_T h(A_T)} \exp(-h^2(A_T)/2) \left(1 + \frac{C}{\log\log A_T}\right) \; dT$$

are either both convergent or both divergent according as the function $h(T)$ belongs to the ϵ-upper class or lower class of the process $\{A_T^{1/2}(\hat{\theta}_T - \theta_0), T \geq 0\}$.

We now prove Theorems 2.14–2.16.

Proof of Theorem 2.14: Let $h(.)$ be a nonnegative, nondecreasing function converging to infinity as $t \to \infty$. Applying Theorem 2.13, we get

$$|P_{\theta_0}^T(A_T^{1/2}(\hat{\theta}_T - \theta_0) > h(A_T)) - (1 - \Phi(h(A_T))| = O(\epsilon_T^{1/2}). \tag{2.87}$$

In view of condition (A_4),

$$\int_3^\infty (\log\log A_T)A_T^{-1}|P_{\theta_0}^T(A_T^{1/2}(\hat{\theta}_T - \theta_0) > h(A_T)) - (1 - \Phi(h(A_T)))|dT < \infty. \tag{2.88}$$

Applying the inequality

$$\left(\frac{1}{x} - \frac{1}{x^3}\right)\frac{1}{\sqrt{2\pi}}\exp(-x^2/2) \leq 1 - \Phi(x) \leq \frac{1}{x}\frac{1}{\sqrt{2\pi}}\exp(-x^2/2)$$

for $x > 0$ (cf. Feller (1968), p. 175), we get

$$\int_3^\infty (\log\log A_T)A_T^{-1}\left|P_{\theta_0}^T(A_T^{1/2}(\hat{\theta}_T - \theta_0) > h(A_T)) - \frac{1}{\sqrt{2\pi}}h^{-1}(A_T)\exp\left(-\frac{h^2(A_T)}{2}\right)\right| dT < \infty. \tag{2.89}$$

This proves Theorem 2.14.

Proof of Theorem 2.15: We have seen, from Equation (2.85), that

$$A_T^{1/2}(\hat{\theta}_T - \theta_0) = \frac{W^*(I_T(\theta_0))/A_T^{1/2}}{I_T(\theta_0)/A_T} \quad \text{a.s. } [P_{\theta_0}]$$

and hence

$$\frac{A_T^{1/2}(\hat{\theta}_T - \theta_0)}{\sqrt{\log\log A_T}} = \left\{\frac{W^*(I_T(\theta_0))}{\sqrt{2I_T(\theta_0)\log\log I_T(\theta_0)}} \times \frac{\sqrt{2}\sqrt{2I_T(\theta_0)\log\log I_T(\theta_0)}}{\sqrt{2A_T\log\log A_T}}\right\} \Big/ \frac{I_T(\theta_0)}{A_T} \quad \text{a.s. } [P_{\theta_0}]$$

This relation in turn shows that

$$\lim_{T\to\infty} \frac{A_T^{1/2}(\hat{\theta}_T - \theta_0)}{\sqrt{\log\log A_T}} - \frac{\sqrt{2}W^*(A_T)}{\sqrt{2A_T\log\log A_T}} = 0 \quad \text{a.s. } [P_{\theta_0}] \tag{2.90}$$

by using condition (A_2) and from the law of iterated logarithm for the Wiener process which shows that

$$\lim_{T\to\infty} \frac{W^*(I_T(\theta_0))}{\sqrt{2I_T(\theta_0)\log\log I_T(\theta_0)}} = 1 \ \text{ a.s. } [P_{\theta_0}]$$

and

$$\lim_{T\to\infty} \frac{W^*(A_T)}{\sqrt{2A_T \log\log A_T}} = 1 \ \text{ a.s. } [P_{\theta_0}]$$

(cf. Kulinich (1985), p. 564). In view of Equation (2.90), it follows that

$$\left| A_T^{1/2}(\hat{\theta}_T - \theta_0) - \frac{W^*(A_T)}{A_T^{1/2}} \right| < \epsilon\sqrt{\log\log A_T} \ \text{ a.s. } [P_{\theta_0}] \tag{2.91}$$

as $T \to \infty$. Let $h_1(t) = \sqrt{\log\log t}$ and $h_2(t) = 2\sqrt{\log\log t}$. Suppose that

$$h_1(A_T) \le h(A_T) \le h_2(A_T) \tag{2.92}$$

for T sufficiently large. We will prove the theorem under the above assumption at first and then extend the result for any arbitrary nonnegative, nondecreasing function $h(t)$ converging to infinity as $t \to \infty$ using Lemma 2.18. In view of the inequality (2.91), we get that

$$\frac{W^*(A_T)}{A_T^{1/2}} - \epsilon\sqrt{\log\log A_T} < A_T^{1/2}(\hat{\theta}_T - \theta_0) < \frac{W^*(A_T)}{A_T^{1/2}} + \epsilon\sqrt{\log\log A_T} \tag{2.93}$$

a.s. $[P_{\theta_0}]$ as $T \to \infty$. Then

$$\begin{aligned} 0 &\le P_{\theta_0}[A_T^{1/2}(\hat{\theta}_T - \theta_0) > (1+\epsilon)h(A_T) \ \text{ infinitely often as } \ T \to \infty] \\ &\le P_{\theta_0}\left[\frac{W^*(A_T)}{A_T^{1/2}} + \epsilon h_1(A_T) > (1+\epsilon)h(A_T) \ \text{ infinitely often as } \ T \to \infty\right] \\ &= P_{\theta_0}\left[\frac{W^*(A_T)}{A_T^{1/2}} > (1+\epsilon)h(A_T) - \epsilon h_1(A_T) \ \text{ infinitely often as } \ T \to \infty\right] \\ &\le P_{\theta_0}\left[\frac{W^*(A_T)}{A_T^{1/2}} > h(A_T) \ \text{ infinitely often as } \ T \to \infty\right] \\ &= 0 \end{aligned} \tag{2.94}$$

by Kolmogorov's test for the Wiener process (cf. Ito and McKean (1965), p. 163) since the function $h(A_T)$ is nondecreasing in T and since $K(h) < \infty$. Equation (2.94) shows that

$$P_{\theta_0}[A_T^{1/2}(\hat{\theta}_T - \theta_0) > (1+\epsilon)h(A_T) \ \text{ infinitely often as } \ T \to \infty] = 0$$

which proves that the function $h(A_T)$ belongs to the ϵ-upper class of the process $A_T^{1/2}(\hat{\theta}_T - \theta_0)$ for every $0 < \epsilon < 1$. On the other hand, it follows, again from (2.92), that

$$\begin{aligned}
1 &\geq P_{\theta_0}[A_T^{1/2}(\hat{\theta}_T - \theta_0) > h(A_T) \text{ infinitely often as } T \to \infty] \\
&\geq P_{\theta_0}\left[\frac{W^*(A_T)}{A_T^{1/2}} - \epsilon h_1(A_T) > h(A_T) \text{ infinitely often as } T \to \infty\right] \\
&\geq P_{\theta_0}\left[\frac{W^*(A_T)}{A_T^{1/2}} > h(A_T) + \epsilon h_1(A_T) \text{ infinitely often as } T \to \infty\right] \\
&\geq P_{\theta_0}\left[\frac{W^*(A_T)}{A_T^{1/2}} > h(A_T) \text{ infinitely often as } T \to \infty\right]. \qquad (2.95)
\end{aligned}$$

Furthermore, the function $h(A_T)$ is nondecreasing in T. Hence, if $K(h) = \infty$, then

$$P_{\theta_0}\left[\frac{W^*(A_T)}{A_T^{1/2}} > h(A_T) \text{ infinitely often as } T \to \infty\right] = 1$$

which in turn shows that

$$P_{\theta_0}[A_T^{1/2}(\hat{\theta}_T - \theta_0) > h(A_T) \text{ infinitely often as } T \to \infty] = 1.$$

Therefore the function $h(A_T)$ belongs to the lower class of the process $A_T^{1/2}(\hat{\theta}_T - \theta_0)$ for every $0 < \epsilon < 1$. Our arguments are similar to those in Jain *et al*. (1975), p. 130. We now extend the result for arbitrary nondecreasing function $h(A_T)$ converging to infinity as $T \to \infty$. Let $h(A_T)$ be such a function. Define $\hat{h}(A_T) = \min(\max(h(A_t), h_1(A_T)), h_2(A_T))$, where $h_1(t)$ and $h_2(t)$ are functions as defined earlier. In view of Lemma 2.18, it follows that

$$K(h) < \infty \Rightarrow K(\hat{h}) < \infty \text{ and } \hat{h} \leq h \text{ near infinity.}$$

Suppose $K(h) < \infty$. Then $K(\hat{h}) < \infty$ and it follows that

$$P_{\theta_0}[A_T^{1/2}(\hat{\theta}_T - \theta_0) > (1+\epsilon)\hat{h}(A_T) \text{ infinitely often as } T \to \infty] = 0.$$

But $\hat{h}(A_T) \leq h(A_T)$ near infinity as observed above. Hence

$$P_{\theta_0}[A_T^{1/2}(\hat{\theta}_T - \theta_0) > (1+\epsilon)h(A_T) \text{ infinitely often as } T \to \infty] = 0.$$

On the other hand, suppose that $K(h) = \infty$. Then, again by Lemma 2.18, we note that $K(h) = \infty \Rightarrow K(\hat{h}) = \infty$ and

$$P_{\theta_0}[A_T^{1/2}(\hat{\theta}_T - \theta_0) > \hat{h}(A_T) \text{ infinitely often} \quad \text{as } T \to \infty] = 1. \tag{2.96}$$

Since $K(h_2) < \infty$, it follows that

$$P_{\theta_0}[A_T^{1/2}(\hat{\theta}_T - \theta_0) \le (1+\epsilon)h_2(A_T) \text{ for large } T] = 1. \tag{2.97}$$

Relations (2.96) and (2.97) imply that $\hat{h}(A_T) \le (1+\epsilon)h_2(A_T)$ near infinity. Since the above inequality holds for all $\epsilon > 0$, taking the limit as $\epsilon \to 0$, it follows that

$$\hat{h}(A_T) \le h_2(A_T) \text{ near infinity.} \tag{2.98}$$

Again, from the definition of $\hat{h}$ and the inequality (2.98), we note that $\hat{h}(A_T) \ge h(A_T)$ for large T. Therefore

$$P_{\theta_0}[A_T^{1/2}(\hat{\theta}_T - \theta_0) > h(A_T) \text{ infinitely often} \quad \text{as } T \to \infty] = 1. \tag{2.99}$$

This completes the proof of Theorem 2.15 following the techniques in Jain *et al*. (1975).

Proof of Theorem 2.16: Let the function $h(A_T)$ belong to the ϵ-upper class of the process $\{A_T^{1/2}(\hat{\theta}_T - \theta_0), T \ge 0\}$. Then

$$\int_3^\infty \frac{h(A_T)}{A_T} \exp\left(-\frac{h^2(A_T)}{2}\right) dT < \infty$$

by Theorem 2.15 which in turn implies that

$$\int_3^\infty \frac{\log\log A_T}{A_T} P_{\theta_0}^T[A_T^{1/2}(\hat{\theta}_T - \theta_0) > h(A_T)]dT < \infty$$

by Theorem 2.14. Conversely, suppose that

$$\int_3^\infty \frac{\log\log A_T}{A_T} P_{\theta_0}^T[A_T^{1/2}(\hat{\theta}_T - \theta_0) > h(A_T)]dT < \infty.$$

Then, Theorem 2.14 implies that

$$\int_3^\infty \frac{\log\log A_T}{A_T h(A_T)} \exp\left(-\frac{h^2(A_T)}{2}\right) dT < \infty.$$

Hence, by the remarks made after the statement given in Lemma 2.18 and by Lemma 2.17, we get that

$$\int_3^\infty \frac{h(A_T)}{A_T} \exp\left(-\frac{h^2(A_T)}{2}\right) dT < \infty$$

which shows that the function $h(A_T)$ belongs to the ϵ-upper class of the process $\{A_T^{1/2}(\hat{\theta}_T - \theta_0), T \geq 0\}$. Replacing 'convergence' by 'divergence' in the above argument, the result for the lower class can be obtained in a similar manner. This proves Theorem 2.16.

We now discuss an example where the conditions (A_1)–(A_4) of Section 2.3 can be checked, ensuring that the results stated in Theorems 2.13 to 2.16 hold.

Consider the SDE

$$dX(t) = [a(t, X(t)) + \theta]dt + dW_t^H, t \geq 0, X(0) = 0. \tag{2.100}$$

From the results described in Section 2.4, we get

$$\hat{\theta}_T - \theta_0 = \frac{\int_0^T J_2(t) dM_t^H}{\int_0^T J_2^2(t) dw_t^H}. \tag{2.101}$$

where

$$J_2(t) = \frac{d}{dw_t^H} \int_0^t k_H(t, s) ds$$

and $k_H(t, s)$ is as defined in (1.70) in Chapter 1. Let

$$R_T = \int_0^T J_2(t) dM_t^H, \ t \geq 0.$$

Then the process $\{R_T, T \geq 0\}$ is a continuous local martingale with the quadratic variation process

$$\langle R \rangle_T = \int_0^T J_2^2(t) dw_t^H = \int_0^T \left(\frac{d}{dw_t^H} \int_0^t k_H(t, s) ds\right)^2 dw_t^H = I_T \ \text{(say)}.$$

It is known that

$$\int_0^t s^{\frac{1}{2}-H}(t-s)^{\frac{1}{2}-H} ds = \frac{\Gamma^2(\frac{3}{2} - H)}{\Gamma(3 - 2H)} t^{2-2H}$$

and hence

$$\frac{d}{dt}\left[\int_0^t s^{\frac{1}{2}-H}(t-s)^{\frac{1}{2}-H} ds\right] = \frac{\Gamma^2(\frac{3}{2} - H)}{\Gamma(2 - 2H)} t^{1-2H}.$$

Therefore

$$\begin{aligned}\frac{d}{dw_t^H}\left[\int_0^t k_H(t,s)ds\right] &= \frac{1}{k_H}\frac{d}{dt}\left[\int_0^t s^{\frac{1}{2}-H}(t-s)^{\frac{1}{2}-H}ds\right]\frac{dt}{dw_t^H} \\ &= \frac{\Gamma^2(\frac{3}{2}-H)}{\Gamma(2-2H)}t^{1-2H}\frac{dt}{dw_t^H} \\ &= \frac{\Gamma^2(\frac{3}{2}-H)}{\Gamma(2-2H)}t^{1-2H}\left[\frac{1}{\lambda_H}(2-2H)t^{1-2H}\right]^{-1} \\ &= C_{1H} \quad \text{(say)} \end{aligned} \tag{2.102}$$

which implies that $\langle R\rangle_T = I_T = C_{2H}t^{2-2H}$ for some positive constant C_{2H}. Let $A_T = C_{2H}T^{2-2H}, \epsilon_T = T^{-\gamma}$ where $0 < \gamma < 1-H$. It can be checked that the conditions $(A_1)-(A_4)$ hold with these choices of A_T and ϵ_T. Hence the results stated in Theorems 2.14–2.16 hold.

Remarks: Results stated in Theorems 2.14–2.16 are analogous to those in Acharya and Mishra (1994) for diffusion processes. However the proof of Theorem 3.3 in their paper is incorrect. It is easy to check that our results continue to hold in the case $H = \frac{1}{2}$, that is, for diffusion processes.

2.8 Instrumental variable estimation

Let us consider the SDE

$$dX(t) = [a(t, X(t)) + \theta\, b(t, X(t))]dt + \sigma(t)dW_t^H,\ X(0) = 0, t \geq 0 \tag{2.103}$$

where $\theta \in \Theta \subset R$, $W = \{W_t^H, t \geq 0\}$ is fBm with known Hurst parameter H and $\sigma(t)$ is a positive non vanishing, non random function on $[0, \infty)$. In other words, $X = \{X_t, t \geq 0\}$ is a stochastic process satisfying the stochastic integral equation

$$X(t) = \int_0^t [a(s, X(s)) + \theta\, b(s, X(s))]ds + \int_0^t \sigma(s)dW_s^H,\ t \geq 0. \tag{2.104}$$

Let

$$C(\theta, t) = a(t, X(t)) + \theta\, b(t, X(t)),\ t \geq 0 \tag{2.105}$$

and assume that the sample paths of the process $\{C(\theta, t)/\sigma(t), t \geq 0\}$ are smooth enough so that the process

$$Q_{H,\theta}(t) = \frac{d}{dw_t^H}\int_0^t k_H(t,s)\frac{C(\theta,s)}{\sigma(s)}ds,\ t \geq 0 \tag{2.106}$$

is well defined where w_t^H and $k_H(t,s)$ are as defined in (1.72) and (1.70) of Chapter 1 respectively. Suppose the sample paths of the process $\{Q_{H,\theta}, 0 \leq t \leq T\}$ belong a.s. to $L^2([0,T], dw_t^H)$. Define

$$Z_t = \int_0^t \frac{k_H(t,s)}{\sigma(s)} dX_s,\ t \geq 0. \tag{2.107}$$

Then the process $Z = \{Z_t, t \geq 0\}$ is an $(\mathcal{F}_t)$-semimartingale with the decomposition

$$Z_t = \int_0^t Q_{H,\theta}(s) dw_s^H + M_t^H \tag{2.108}$$

where M^H is the fundamental martingale defined by (1.73) in Chapter 1 and the process X admits the representation

$$X_t = \int_0^t K_H^\sigma(t,s) dZ_s \tag{2.109}$$

where the function K_H^σ is as defined by (1.75) in Chapter 1. Let P_θ^T be the measure induced by the process $\{X_t, 0 \leq t \leq T\}$ when θ is the true parameter. Following the discussion in Chapter 1, we get that the Radon–Nikodym derivative of P_θ^T with respect to P_0^T is given by

$$\frac{dP_\theta^T}{dP_0^T} = \exp\left[\int_0^T Q_{H,\theta}(s) dZ_s - \frac{1}{2}\int_0^T Q_{H,\theta}^2(s) dw_s^H\right]. \tag{2.110}$$

We now consider another method of estimation of the parameter θ based on observation of the process $X = \{X_t, 0 \leq t \leq T\}$ and study the asymptotic properties of such estimators as $T \to \infty$.

Let $\{\alpha(t), t \geq 0\}$ be a stochastic process such that the function $\alpha(t)$ is $\mathcal{F}_t$-measurable. An example of such a process is $\alpha(t) = K(t, \tilde{X}(t))$, where $\tilde{X}(t) = \{X(s), 0 \leq s \leq t\}$ and $K(.,.)$ is a real-valued measurable function defined on $R_+ \times R$. Suppose that

$$\int_0^T E[(\alpha(t))^2] dw_T^H < \infty.$$

This condition implies that the stochastic integral

$$\int_0^T \alpha(t) dM_t^H$$

exists as a stochastic integral with respect to the martingale $\{M_t^H, \mathcal{F}_t, t \geq 0\}$. In particular

$$E\left(\int_0^T \alpha(t) dM_t^H\right) = 0.$$

Observing that

$$dZ_t = dM_t^H + Q_{H,\theta}(t)dw_t^H$$

from (2.108), we can rewrite the above equation in the form

$$E\left(\int_0^T \alpha(t)(dZ_t - Q_{H,\theta}(t)dw_t^H)\right) = 0$$

or equivalently

$$E\left(\int_0^T \alpha(t)(dZ_t - (J_1(t) + \theta J_2(t))dw_t^H)\right) = 0 \tag{2.111}$$

where

$$\begin{aligned} Q_{H,\theta}(t) &= \frac{d}{dw_t^H}\int_0^t k_H(t,s)\frac{C(\theta,s)}{\sigma(s)}ds \\ &= \frac{d}{dw_t^H}\int_0^t k_H(t,s)\frac{a(s,X(s))}{\sigma(s)}ds + \theta\frac{d}{dw_t^H}\int_0^t k_H(t,s)\frac{b(s,X(s))}{\sigma(s)}ds \\ &= J_1(t) + \theta J_2(t) \quad \text{(say)}. \end{aligned} \tag{2.112}$$

A sample analogue of Equation (2.111) is

$$\int_0^T \alpha(t)(dZ_t - (J_1(t) + \theta J_2(t))dw_t^H) = 0. \tag{2.113}$$

which motivates an instrumental variable estimator defined below.

Definition: Corresponding to the $\mathcal{F}_t$-adapted instrument process $\{\alpha(t), t \geq 0\}$, the *instrumental variable estimator* (IVE) of θ is defined by

$$\tilde{\theta}_T = \frac{\int_0^T \alpha(t)(dZ_t - J_1(t)dw_t^H)}{\int_0^T \alpha(t)J_2(t)dw_t^H}.$$

Choosing the process $\{\alpha(t), t \geq 0\}$ suitably, we can obtain a class of instrumental variable estimators (IVEs) for θ. In analogy with least squares estimation, we can choose $\alpha(t) = J_2(t)$ as defined above and the corresponding IVE may be called a *least squares estimator* (LSE). In fact, it is also the MLE (cf. Prakasa Rao (2003)). In the following discussion, we will choose $\alpha(t) = K(t, \tilde{X}(t))$ where $K(.,.)$ is a real-valued measurable function defined on $R_+ \times R$.

Suppose θ_0 is the true value of the parameter θ. It is easy to check that

$$\tilde{\theta}_T - \theta_0 = \frac{\int_0^T K(t,\tilde{X}(t))dM_t^H}{\int_0^T K(t,\tilde{X}(t))J_2(t)dw_t^H} \tag{2.114}$$

using the fact that

$$dZ_t = (J_1(t) + \theta_0 J_2(t))dw_t^H + dM_t^H. \tag{2.115}$$

Note that the quadratic variation $\langle Z \rangle$ of the process Z is the same as the quadratic variation $\langle M^H \rangle$ of the martingale M^H which in turn is equal to w^H. This follows from Equations (1.79) and (1.73) in Chapter 1. Hence we obtain that

$$[w_T^H]^{-1} \lim_n \sum [Z_{t_{i+1}^{(n)}} - Z_{t_i^{(n)}}]^2 = 1 \text{ a.s. } [P_{\theta_0}]$$

where $(t_i^{(n)})$ is a partition of the interval $[0, T]$ such that $\sup |t_{i+1}^{(n)} - t_i^{(n)}|$ tends to zero as $n \to \infty$. If the function $\sigma(.)$ is an unknown constant σ, the above property can be used to obtain a strongly consistent estimator of σ^2 based on the continuous observation of the process X over the interval $[0, T]$. Hereafter we assume that the non random function $\sigma(t)$ is known.

We now discuss the problem of instrumental variable estimation of the parameter θ on the basis of observation of the process X or equivalently the process Z in the interval $[0, T]$.

Equation (2.114) can be written in the form

$$\tilde{\theta}_T - \theta_0 = \frac{\int_0^T K(t, \tilde{X}(t))dM_t^H}{\int_0^T K(t, \tilde{X}(t))^2 dw_t^H} \frac{\int_0^T K(t, \tilde{X}(t))^2 dw_t^H}{\int_0^T K(t, \tilde{X}(t)) J_2(t) dw_t^H}. \tag{2.116}$$

Strong consistency

Theorem 2.19: The IVE $\tilde{\theta}_T$ is strongly consistent, that is,

$$\tilde{\theta}_T \to \theta_0 \text{ a.s. } [P_{\theta_0}] \text{ as } T \to \infty \tag{2.117}$$

provided

$$\text{(i)} \int_0^T K(t, \tilde{X}(t))^2 dw_t^H \to \infty \text{ a.s. } [P_{\theta_0}] \text{ as } T \to \infty \tag{2.118}$$

and

$$\text{(ii)} \limsup_{T \to \infty} \left| \frac{\int_0^T K(t, \tilde{X}(t))^2 dw_t^H}{\int_0^T K(t, \tilde{X}(t)) J_2(t) dw_t^H} \right| < \infty \text{ a.s. } [P_{\theta_0}]. \tag{2.119}$$

Proof: This theorem follows by observing that the process

$$R_T \equiv \int_0^T K(t, \tilde{X}(t))dM_t^H, \ t \geq 0 \tag{2.120}$$

is a local martingale with the quadratic variation process

$$\langle R\rangle_T = \int_0^T K(t, \tilde{X}(t))^2(t)dw_t^H \tag{2.121}$$

and applying the strong law of large numbers (cf. Liptser (1980), Prakasa Rao (1999b), p. 61) under conditions (i) and (ii) stated above.

Remarks: For the case of the fractional Ornstein–Uhlenbeck-type process defined by Equation (2.1) investigated in Kleptsyna and Le Breton (2002a), it can be checked that the condition stated in the Equations (2.118) and (2.119) hold when $K(t, \tilde{X}(t)) = J_2(t)$ and hence the MLE which is also the LSE is strongly consistent as $T \to \infty$.

Limiting distribution

We now discuss the limiting distribution of the IVE $\tilde{\theta}_T$ as $T \to \infty$. Let

$$\beta_T = \frac{\int_0^T K(t, \tilde{X}(t))^2 dw_t^H}{\int_0^T K(t, \tilde{X}(t))J_2(t)dw_t^H}. \tag{2.122}$$

It is easy to see that

$$\tilde{\theta}_T - \theta_0 = \frac{R_T}{\langle R\rangle_T}\beta_T. \tag{2.123}$$

Theorem 2.20: Assume that the functions $b(t, s)$ and $\sigma(t)$ are such that the process $\{R_t, t \geq 0\}$ is a local continuous martingale and that there exists a process $\{\gamma_t, t \geq 0\}$ such that γ_t is $\mathcal{F}_t$-adapted and

$$\gamma_T^2\langle R\rangle_T = \gamma_T^2 \int_0^T K(t, \tilde{X}(t))^2(t)dw_t^H \to \eta^2 \text{ in probability as } T \to \infty \tag{2.124}$$

where $\gamma_T^2 \to 0$ a.s. $[P]$ a.s. $T \to \infty$ and η is a random variable such that $P(\eta > 0) = 1$. Then

$$(\gamma_T R_T, \gamma_T^2\langle R\rangle_T) \to (\eta Z, \eta^2) \text{ in law as } T \to \infty \tag{2.125}$$

where the random variable Z has the standard normal distribution and the random variables Z and η are independent.

Proof: This theorem follows as a consequence of the central limit theorem for local martingales (cf. Theorem 1.49 and Remark 1.47 in Prakasa Rao (1999b), p. 65).

Observe that

$$\beta_T^{-1}\gamma_T^{-1}(\tilde{\theta}_T - \theta_0) = \frac{\gamma_T R_T}{\gamma_T^2 \langle R \rangle_T}. \tag{2.126}$$

Applying Theorem 2.20, we obtain the following result.

Theorem 2.21: Suppose the conditions stated in Theorem 2.20 hold. Then

$$(\beta_T \gamma_T)^{-1}(\tilde{\theta}_T - \theta_0) \to \frac{Z}{\eta} \text{ in law as } T \to \infty \tag{2.127}$$

where the random variable Z has the standard normal distribution and the random variables Z and η are independent.

Remarks:

(i) If the random variable η is a constant with probability one, then the limiting distribution of the normalized IVE with random norming is normal with mean zero and variance η^{-2}. Otherwise it is a mixture of the normal distributions with mean zero and variance η^{-2} with the mixing distribution as that of η.

(ii) Note that the IVE is not necessarily asymptotically efficient. It is not asymptotically efficient even when the random variable η is a constant. It is asymptotically efficient in this case if $K(t, \tilde{X}(t)) = J_2(t)$ as defined by Equation (2.112). Observe that the IVE reduces to the MLE in the case $K(t, \tilde{X}(t)) = J_2(t)$.

The results discussed in this section are due to Prakasa Rao (2007).

Berry–Esseen-type bound for IVE

Hereafter we assume that the random variable η in (2.124) is a positive constant with probability one. Hence

$$(\beta_T \gamma_T)^{-1}(\tilde{\theta}_T - \theta_0) \to N(0, \eta^{-2}) \text{ in law as } T \to \infty \tag{2.128}$$

where $N(0, \eta^{-2})$ denotes the Gaussian distribution with mean zero and variance η^{-2}. We will now study the rate of convergence of the asymptotic distribution of the IVE in (2.128).

Suppose there exist non-random positive functions δ_T decreasing to zero and ε_T decreasing to zero such that

$$\delta_T^{-1}\varepsilon_T^2 \to \infty \text{ as } T \to \infty \tag{2.129}$$

and

$$\sup_{\theta\in\Theta} P_\theta^T[|\delta_T\langle R\rangle_T - 1| \geq \varepsilon_T] = O(\varepsilon_T^{1/2}) \tag{2.130}$$

where the process $\{R_t, t \geq 0\}$ is as defined in (2.120). Note that the process $\{R_t, t \geq 0\}$ is a locally square integrable continuous martingale. From the results on the representation of locally square integrable continuous martingales (cf. Ikeda and Watanabe (1981), Chapter II, Theorem 7.2), it follows that there exists a standard Wiener process $\{\tilde{W}(t), t \geq 0\}$ adapted to $(\mathcal{F}_t)$ such that $R_t = \tilde{W}(\langle R\rangle_t), t \geq 0$. In particular

$$R_T\delta_T^{1/2} = \tilde{W}(\langle R\rangle_T\delta_T) \text{ a.s. } [P_{\theta_0}] \tag{2.131}$$

for all $T \geq 0$.

Let us fix $\theta \in \Theta$. It is clear from the earlier remarks that

$$R_T = \langle R\rangle_T\beta_T^{-1}(\tilde{\theta}_T - \theta) \tag{2.132}$$

under the P_θ-measure. Then it follows, from Lemmas 2.9 and 2.10, that

$$\begin{aligned}
&|P_\theta[\delta_T^{-1/2}\beta_T^{-1}(\hat{\theta}_T - \theta_0) \leq x] - \Phi(x)| \\
&\quad = \left|P_\theta\left[\frac{R_T}{\langle R\rangle_T}\delta_T^{-1/2} \leq x\right] - \Phi(x)\right| \\
&\quad = \left|P_\theta\left[\frac{R_T/\delta_T^{-1/2}}{\langle R\rangle_T/\delta_T^{-1}} \leq x\right] - \Phi(x)\right| \\
&\quad \leq \sup_x |P_\theta[R_T\delta_T^{1/2} \leq x] - \Phi(x)| \\
&\qquad + P_\theta[|\delta_T\langle R\rangle_T - 1| \geq \varepsilon_T] + \varepsilon_T \\
&\quad = \sup_y |P(\tilde{W}(\langle R\rangle_T\delta_T) \leq y) - \Phi(y)| + P_\theta[|\delta_T\langle R\rangle_T - 1| \geq \varepsilon_T] + \varepsilon_T \\
&\quad \leq (2\varepsilon_T)^{1/2} + 2P_\theta[|\delta_T\langle R\rangle_T - 1| \geq \varepsilon_T] + \varepsilon_T.
\end{aligned} \tag{2.133}$$

It is clear that the bound obtained above is of order $O(\varepsilon_T^{1/2})$ under the condition (2.130) and it is uniform in $\theta \in \Theta$. Hence we have the following result.

Theorem 2.22: Under conditions (2.129) and (2.130),

$$\begin{aligned}
&\sup_{\theta\in\Theta}\sup_{x\in R} |P_\theta[\delta_T^{-1/2}\beta_T^{-1}(\tilde{\theta}_T - \theta) \leq x] - \Phi(x)| \\
&\quad \leq (2\varepsilon_T)^{1/2} + 2P_\theta[|\delta_T\langle R\rangle_T - 1| \geq \varepsilon_T] + \varepsilon_T = O(\varepsilon_T^{1/2}).
\end{aligned} \tag{2.134}$$

As a consequence of this result, we have the following theorem giving the rate of convergence of the IVE $\tilde{\theta}_T$.

Theorem 2.23: Suppose conditions (2.129) and (2.130) hold. Then there exists a constant $c > 0$ such that for every $d > 0$,

$$\sup_{\theta\in\Theta} P_\theta[\beta_T^{-1}|\tilde{\theta}_T - \theta| \geq d] \leq c\varepsilon_T^{1/2} + 2P_\theta[|\delta_T\langle R\rangle_T - 1| \geq \varepsilon_T]$$
$$= O(\varepsilon_T^{1/2}). \tag{2.135}$$

Proof: Observe that

$$\begin{aligned}
&\sup_{\theta\in\Theta} P_\theta[\beta_T^{-1}|\tilde{\theta}_T - \theta| \geq d] \\
&\quad \leq \sup_{\theta\in\Theta} |P_\theta[\delta_T^{-1/2}\beta_T^{-1}(\tilde{\theta}_T - \theta) \geq d\delta_T^{-1/2}] - 2(1 - \Phi(d\delta_T^{-1/2}))| \\
&\qquad +2(1 - \Phi(d\delta_T^{-1/2})) \\
&\quad \leq (2\varepsilon_T)^{1/2} + 2\sup_{\theta\in\Theta} P_\theta[|\delta_T\langle R\rangle_T - 1| \geq \varepsilon_T] + \varepsilon_T \\
&\qquad +2d^{-1/2}\delta_T^{1/2}(2\pi)^{-1/2}\exp\left(-\frac{1}{2}\delta_T^{-1}d^2\right)
\end{aligned} \tag{2.136}$$

by Theorem 2.22 and the inequality

$$1 - \Phi(x) < \frac{1}{x\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right) \tag{2.137}$$

for all $x > 0$ (cf. Feller (1968), p. 175). Since

$$\delta_T^{-1}\varepsilon^2(T) \to \infty \text{ as } T \to \infty$$

by the condition (2.129), it follows that

$$\sup_{\theta\in\Theta} P_\theta[\beta_T^{-1}|\tilde{\theta}_T - \theta| \geq d] \leq c\varepsilon_T^{1/2} + 2\sup_{\theta\in\Theta} P_\theta[|\delta_T\langle R\rangle_T - 1| \geq \varepsilon_T] \tag{2.138}$$

for some constant $c > 0$ and the last term is of order $O(\varepsilon_T^{1/2})$ by the condition (2.130). This proves Theorem 2.23.

3

Parametric estimation for fractional Ornstein–Uhlenbeck-type process

3.1 Introduction

We studied parametric inference for processes defined by linear SDEs driven by fBm in the previous chapter. We now consider a special case of such processes, namely, the fractional Ornstein–Uhlenbeck-type process studied in Chapter 1. Le Breton (1998) studied parameter estimation and filtering in a simple linear model driven by fBm. Kleptsyna and Le Breton (2002a) studied parameter estimation problems for a fractional Ornstein–Uhlenbeck-type process. Such processes play a potentially important role in the modeling of financial time series. The fractional Ornstein–Uhlenbeck process is a fractional analogue of the Ornstein–Uhlenbeck-process, that is, a continuous time first-order autoregressive process $X = \{X_t, t \geq 0\}$ which is the solution of a one-dimensional homogeneous linear SDE driven by fBm $W^H = \{W_t^H, t \geq 0\}$ with Hurst parameter $H \in (1/2, 1)$. Such a process is the unique Gaussian process satisfying the linear integral equation

$$X_t = X_0 + \theta \int_0^t X_s ds + \sigma W_t^H,\ t \geq 0.$$

We now investigate the problem of estimation of the parameters θ and σ^2 based on the observation $\{X_s, 0 \leq s \leq T\}$ and prove that the MLE $\hat{\theta}_T$ is strongly consistent as $T \to \infty$. We follow the notation given in Chapter 1. For convenience, we recall the earlier notation.

Statistical Inference for Fractional Diffusion Processes B. L. S. Prakasa Rao

3.2 Preliminaries

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ be a stochastic basis satisfying the usual conditions and the processes discussed in the following are $(\mathcal{F}_t)$-adapted. Further, the natural filtration of a process is understood as the P-completion of the filtration generated by this process.

Let $W^H = \{W_t^H, t \geq 0\}$ be standard fBm with Hurst parameter $H \in (1/2, 1)$, that is, a Gaussian process with continuous sample paths such that $W_0^H = 0$, $E(W_t^H) = 0$ and

$$E(W_s^H W_t^H) = \frac{1}{2}[s^{2H} + t^{2H} - |s-t|^{2H}],\ t \geq 0, s \geq 0. \tag{3.1}$$

Let us consider a stochastic process $\{X_t, t \geq 0\}$ defined by the stochastic integral equation

$$X_t = x_0 + \theta \int_0^t X(s)ds + \sigma W_t^H,\ 0 \leq t \leq T \tag{3.2}$$

where θ is an unknown drift parameter. For convenience, we write the above integral equation in the form of a SDE

$$dX_t = \theta X(t)dt + \sigma dW_t^H,\ X_0 = x_0, 0 \leq t \leq T, \tag{3.3}$$

driven by fBm W^H. For a discussion on the equivalence of (3.2) and (3.3), see Cheridito *et al*. (2003). Even though the process X is not a semimartingale, one can associate a semimartingale $Z = \{Z_t, t \geq 0\}$ which is called a *fundamental semimartingale* such that the natural filtration $(\mathcal{Z}_t)$ of the process Z coincides with the natural filtration $(\mathcal{X}_t)$ of the process X (Kleptsyna *et al*. (2000a)). Define, for $0 < s < t$,

$$k_H = 2H\Gamma\left(\frac{3}{2} - H\right)\Gamma\left(H + \frac{1}{2}\right), \tag{3.4}$$

$$k_H(t,s) = k_H^{-1} s^{\frac{1}{2}-H}(t-s)^{\frac{1}{2}-H}, \tag{3.5}$$

$$\lambda_H = \frac{2H\ \Gamma(3-2H)\Gamma(H+\frac{1}{2})}{\Gamma(\frac{3}{2}-H)}, \tag{3.6}$$

$$w_t^H = \lambda_H^{-1} t^{2-2H}, \tag{3.7}$$

and

$$M_t^H = \int_0^t k_H(t,s)dW_s^H,\ t \geq 0. \tag{3.8}$$

The process M^H is a Gaussian martingale, called the *fundamental martingale* (cf. Norros *et al*. (1999)), and its quadratic variance $\langle M^H \rangle_t = w_t^H$. Furthermore, the natural filtration of the martingale M^H coincides with the natural filtration of fBm W^H. Let

$$K_H(t,s) = H(2H-1)\frac{d}{ds}\int_s^t r^{H-\frac{1}{2}}(r-s)^{H-\frac{3}{2}}dr,\ 0 \leq s \leq t. \tag{3.9}$$

The sample paths of the process $\{X_t, t \geq 0\}$ are smooth enough so that the process Q defined by

$$Q(t) = \frac{d}{dw_t^H}\int_0^t k_H(t,s)X_s ds,\ t \in [0,T] \tag{3.10}$$

is well defined where w^H and k_H are as defined in (3.7) and (3.5) respectively and the derivative is understood in the sense of absolute continuity with respect to the measure generated by w^H. Moreover, the sample paths of the process Q belong to $L^2([0,T], dw^H)$ a.s. $[P]$. The following theorem due to Kleptsyna *et al*. (2000a) associates a fundamental semimartingale Z associated with the process X such that the natural filtration $(\mathcal{Z}_t)$ coincides with the natural filtration $(\mathcal{X}_t)$ of X.

Theorem 3.1: Let the process $Z = (Z_t, t \in [0,T])$ be defined by

$$Z_t = \int_0^t k_H(t,s)dX_s \tag{3.11}$$

where the function $k_H(t,s)$ is as defined in (3.5). Then the following results hold:

(i) The process Z is an $(\mathcal{F}_t)$-semimartingale with the decomposition

$$Z_t = \theta\int_0^t Q(s)dw_s^H + \varepsilon M_t^H \tag{3.12}$$

where M^H is the Gaussian martingale defined by (3.8);

(ii) the process X admits the representation

$$X_t = \int_0^t K_H(t,s)dZ_s \tag{3.13}$$

where the function K_H is as defined in (3.9); and

(iii) the natural filtrations of $(\mathcal{Z}_t)$ and $(\mathcal{X}_t)$ coincide.

3.3 Maximum likelihood estimation

Theorem 3.1 implies that the information available from the observation $\{X_s, 0 \leq s \leq T\}$ is equivalent to that given by the process $\{Z_s, 0 \leq s \leq T\}$ for every

$T \geq 0$. We note that the parameter σ^2 can be estimated with probability one on any finite time interval. This follows from the representation (3.12) and noting that the quadratic variation of the process Z in the interval $[0, T]$ is given by $\langle Z \rangle_t = \sigma^2 w_t^H$ a.s. Hence the parameter σ^2 is obtained by noting that

$$(w_T^H)^{-1} \lim_{n \to \infty} \sum [Z_{t_{i+1}^{(n)}} - Z_{t_i^{(n)}}]^2 = \sigma^2 \text{ a.s.}$$

where $\{t_i^{(n)}\}$ is a suitable partition of the interval $[0, T]$ such that $\sup_i |t_{i+1}^{(n)} - t_i^{(n)}| \to 0$ as $n \to \infty$. In view of this observation, we assume hereafter that σ^2 is known and, without loss of generality, we assume that $\sigma^2 = 1$ in the following discussion.

Let P_θ^T be the probability measure generated by the process $\{X_s, 0 \leq s \leq T\}$ or equivalently by $\{Z_s, 0 \leq s \leq T\}$. Applying the Girsanov-type formula derived in Kleptsyna *et al*. (2000a) as discussed in Chapter 1, it follows that the measures P_θ^T and P_0^T are absolutely continuous with respect to each other and

$$\log \frac{dP_\theta^{(T)}}{dP_0^{(T)}} = \theta \int_0^T Q(s) dZ_s - \frac{\theta^2}{2} \int_0^T Q^2(s) dw_s^H. \tag{3.14}$$

It is easy to check that the MLE of θ, based on the observation $\{X_s, 0 \leq s \leq T\}$, is given by

$$\hat{\theta}_T = \frac{\int_0^T Q(s) dZ_s}{\int_0^T Q^2(s) dw_s^H}. \tag{3.15}$$

Proposition 3.2: The estimator $\hat{\theta}_T$ is strongly consistent, that is, $\lim_{T \to \infty} \hat{\theta}_t = \theta$ a.s. for every $\theta \in R$.

Proof: From the representation of Z given by (3.12), it can be shown that

$$\hat{\theta}_T - \theta = \frac{\int_0^T Q(s) dM_s^H}{\int_0^T Q^2(s) dw_s^H} \tag{3.16}$$

where M^H is the fundamental Gaussian martingale defined in (3.8) with the quadratic variation w^H. Since

$$\int_0^T Q^2(s) dw_s^H$$

is the quadratic variation of the local martingale

$$\int_0^T Q(s) dM_s^H,$$

it follows that $\hat{\theta}_T - \theta$ converges to zero a.s. as $T \to \infty$ provided

$$\lim_{T\to\infty} \int_0^T Q^2(s) dw_s^H = \infty \text{ a.s.}$$

This follows from the strong law of large numbers for local martingales due to Liptser (1980) (cf. Prakasa Rao (1999b)). Let

$$\psi_t^H(\theta; a) = E_\theta \left[\exp\left(-a \int_0^T Q^2(s) dw_s^H \right) \right], \quad a > 0.$$

Kleptsyna and Le Breton (2002a) obtained an explicit formula for the function $\Psi_T^H(\theta, a)$ for $a > 0$ and proved that

$$\lim_{T\to\infty} \Psi_T^H(\theta, a) = 0.$$

This in turn proves that

$$\lim_{T\to\infty} \int_0^T Q^2(s) dw_s^H = \infty \text{ a.s.}$$

completing the proof of this result.

Let

$$B(\theta, T) = E_\theta[\hat{\theta}_T - \theta]$$

and

$$V(\theta, T) = E_\theta[(\hat{\theta}_T - \theta)^2].$$

Then $B(\theta, T)$ is the bias and $V(\theta, T)$ is the mean square error of the estimator $\hat{\theta}_T$ when θ is the true parameter. The following result is due to Kleptsyna and Le Breton (2002a). We do not give details here. Recall that $\frac{1}{2} < H < 1$. We say that $f(t) \simeq g(t)$ if $f(t)/g(t) \to 1$ as $t \to \infty$.

Proposition 3.3: The following properties hold:

(i) If $\theta < 0$, then, as $T \to \infty$,

$$B(\theta, T) \simeq 2T^{-1} \text{ and } V(\theta, T) \simeq 2|\theta| T^{-1}.$$

(ii) If $\theta = 0$, then, for all T,

$$B(0, T) = B(0, 1)\, T^{-1} \text{ and } V(\theta, T) = V(0, 1) T^{-2}.$$

(iii) If $\theta < 0$, then, as $T \to \infty$,

$$B(\theta, T) \simeq -2\sqrt{\pi \sin \pi H}\theta^{3/2} e^{-\theta T} T^{1/2} \text{ and}$$

$$V(\theta, T) \simeq 2\sqrt{\pi \sin \pi H}\theta^{5/2} e^{-\theta T} T^{1/2}.$$

Remarks: Observe that the rates of convergence to zero, for the bias and mean square error, do not depend on the Hurst index H. Furthermore, if $\theta < 0$, then the process is asymptotically stationary. Limiting properties of the MLE $\hat{\theta}_T$ are discussed in Chapter 2.

Following the notation introduced by Kleptsyna and Le Breton (2002a), we now discuss the rate of convergence for the MLE for the fractional Ornstein–Uhlenbeck-type process studied by them. Kleptsyna and Le Breton (2002a) obtained expressions for the bias and the mean square error of the estimator $\hat{\theta}_T$ in terms of the function

$$\psi_T^H(\theta; a) = E_\theta\left[\exp\left(-a\int_0^T Q^2(s)dw_s^H\right)\right], \ a > 0. \tag{3.17}$$

where E_θ denotes the expectation when θ is the true parameter. They obtained a closed form expression for this function involving modified Bessel functions of the first kind (cf. Watson (1995)) and analyzed the asymptotic behavior as $T \to \infty$ for different values of θ. It follows that

$$E_\theta\left(\int_0^T Q^2(s)dw_s^H\right) = -\lim_{a\to 0+}\frac{d\psi_T^H(\theta; a)}{da} \tag{3.18}$$

from (3.17). Let

$$\mathcal{L}_T^H(\theta; \rho) = E_\theta\left[\exp\left(-\rho\int_0^T Q(s)dZ_s\right)\right], \ \rho > 0. \tag{3.19}$$

Kleptsyna and Le Breton (2002a) also obtained explicit expressions for the function $\mathcal{L}_T^H(\theta; \rho)$ again in terms of the modified Bessel functions of the first kind and one can show that

$$E_\theta\left(\int_0^T Q(s)dZ_s\right) = -\lim_{\rho\to 0+}\frac{d\mathcal{L}_T^H(\theta; \rho)}{d\rho}. \tag{3.20}$$

It seems to be difficult to obtain an explicit functional form for the expectations defined in (3.18) and (3.20).

Remarks: One can approach the above problem by computing the joint characteristic function of the vector

$$\left(\int_0^T Q(s)dZ_s, \int_0^T Q^2(s)dw_s^H\right)$$

explicitly by using the results in Kleptsyna and Le Breton (2002a) and then following the technique in Bose (1986) using Esseen's lemma. However, this approach does not seem to be helpful in view of the complex nature of the above characteristic function involving the modified Bessel functions of first kind.

Suppose there exist functions α_T decreasing to zero as $T \to \infty$ and ε_T decreasing to zero as $T \to \infty$ such that

$$\sup_{\theta \in \Theta} P_\theta^T \left[\left| \alpha_T \int_0^T Q^2(t) dw_t^H - 1 \right| \geq \varepsilon_T \right] = O(\varepsilon_T^{1/2}). \tag{3.21}$$

Then it follows that

$$\sup_{\theta \in \Theta} \sup_x |P_\theta^T [\alpha_T^{-1/2}(\hat{\theta}_T - \theta) \leq x] - \Phi(x)| = O(\varepsilon_T^{1/2}) \tag{3.22}$$

and

$$\sup_{\theta \in \Theta} \sup_d P_\theta^T [|\alpha_T^{-1/2}(\hat{\theta}_T - \theta)| \geq d] = O(\varepsilon_T^{1/2}) \tag{3.23}$$

from Theorems 2.11 and 2.12 discussed in Chapter 2.

3.4 Bayes estimation

Suppose that the parameter space Θ is open and Λ is a prior probability measure on the parameter space $\Theta \subset R$. Further suppose that the probability measure Λ has a density $\lambda(.)$ with respect to the Lebesgue measure and the density function is continuous and positive in an open neighborhood of θ_0, the true parameter. The posterior density of θ, given the observation $X^T \equiv \{X_s, 0 \leq s \leq T\}$, is given by

$$p(\theta | X^T) = \frac{(dP_\theta^T / dP_{\theta_0}^T)\lambda(\theta)}{\int_\Theta (dP_\theta^T / dP_{\theta_0}^T)\lambda(\theta)\, d\theta}. \tag{3.24}$$

We define the Bayes estimate (BE) $\tilde{\theta}_T$ of the parameter θ, based on the path X^T and the prior density $\lambda(\theta)$, to be the minimizer of the function

$$B_T(\phi) = \int_\Theta L(\theta, \phi)\; p(\theta | X^T) d\theta,\; \phi \in \Theta$$

where $L(\theta, \phi)$ is a given loss function defined on $\Theta \times \Theta$. In particular, for the quadratic loss function $|\theta - \phi|^2$, the Bayes estimator is the posterior mean given by

$$\tilde{\theta}_T = \frac{\int_\Theta u p(u | X^T) du}{\int_\Theta p(v | X^T) dv}.$$

Suppose the loss function $L(\theta, \phi) : \Theta \times \Theta \to R$ satisfies the following conditions:

D(i) $L(\theta, \phi) = \ell(|\theta - \phi|)$;

D(ii) $\ell(\theta)$ is nonnegative and continuous on R;

D(iii) $\ell(.)$ is symmetric;

D(iv) the sets $\{\theta : \ell(\theta) < c\}$ are convex sets and are bounded for all $c > 0$; and

D(v) there exist numbers $\gamma > 0,\ H_0 \geq 0$ such that, for $H \geq H_0$,

$$\sup\left\{\ell(\theta) : |\theta| \leq H^{\gamma}\right\} \leq \inf\{\ell(\theta) : |\theta| \geq H\}.$$

Obviously, the loss function of the form $L(\theta, \phi) = |\theta - \phi|^2$ satisfies the conditions D(i)–D(v).

Asymptotic properties of the Bayes estimator $\tilde{\theta}_T$ follow from the results discussed in Chapter 2. We do not deal with them again here.

3.5 Probabilities of large deviations of MLE and BE

Let

$$\psi_T^H(\theta; a) = E_\theta\left[\exp\left(-a\int_0^T Q_H^2(t)dw^H(t)\right)\right]$$

for $a > 0$. Kleptsyna and Le Breton (2002a) proved that

$$\psi_T^H(\theta; a) = \left(\frac{4(\sin \pi H)\sqrt{\theta^2 + 2a}e^{-\theta T}}{\pi T D_T^H(\theta; \sqrt{\theta^2 + 2a})}\right)^{1/2}$$

where

$$D_T^H(\theta; \beta) = \left[\beta \cosh\frac{\beta}{2}T - \theta \sinh\frac{\beta}{2}T\right]^2 I_{-H}\left(\frac{\beta}{2}T\right) I_{H-1}\left(\frac{\beta}{2}T\right)$$
$$-\left[\beta \sinh\frac{\beta}{2}T - \theta \cosh\frac{\beta}{2}T\right]^2 I_{1-H}\left(\frac{\beta}{2}T\right) I_H\left(\frac{\beta}{2}T\right)$$

where I_ν is the Bessel function of the first kind and order ν (cf. Watson (1995)). It was also proved in Kleptsyna and Le Breton (2002a) that

$$\lim_{T\to\infty} \psi_T^H(\theta; a) = 0$$

and hence

$$\lim_{T\to\infty} \int_0^T Q_H^2(t)dw^H(t) = +\infty \text{ a.s. } P_\theta.$$

In view of these observations, we make the following assumption.

Condition (A): Fix $\theta \in \Theta$. Suppose that there exists a function α_T tending to zero as $T \to \infty$ such that

$$\alpha_T \int_0^T Q_H^2(t)dw^H(t) \to c > 0 \text{ a.s. } [P_\theta] \text{ as } T \to \infty$$

and there exists a neighborhood N_θ of θ such that

$$\sup_{\phi \in N_\theta} E_\phi \left[\alpha_T \int_0^T Q_H^2(t) dw^H(t) \right] = O(1)$$

as $T \to \infty$.

We now prove the following theorems giving the large-deviation probabilities for the MLE and BE discussed in Section 3.3.

Theorem 3.4: Under the condition (A) stated above, there exist positive constants C_1 and C_2, depending on θ and T, such that for every $\gamma > 0$,

$$P_\theta^T\{|\alpha_T^{-1/2}(\hat{\theta}_T - \theta)| > \gamma\} \le C_1 e^{-C_2\gamma^2}$$

where $\hat{\theta}_T$ is the MLE of the parameter θ.

Theorem 3.5: Under the condition (A) stated above, there exist positive constants C_3 and C_4, depending on θ and T, such that for every $\gamma > 0$,

$$P_\theta^T\{|\alpha_T^{-1/2}(\tilde{\theta}_T - \theta)| > \gamma\} \le C_3 e^{-C_4\gamma^2}$$

where $\tilde{\theta}_T$ is the BE of the parameter θ with respect to the prior $\lambda(.)$ and the loss function $L(., .)$ satisfying the conditions D(i)–D(v).

Let E_θ^T denote the expectation with respect to the probability measure P_θ^T. Fix $\theta \in \Theta$. For proofs of the theorems stated above, we need the following lemmas. Define

$$Z_T(u) = \frac{dP^T_{\theta + u\alpha_T{}^{1/2}}}{dP_\theta^T}.$$

Lemma 3.6: Under the conditions stated above, there exist positive constants c_1 and d_1 such that

$$E_\theta^T[Z_T^{\frac{1}{2}}(u)] \le d_1 e^{-c_1 u^2}$$

for $-\infty < u < \infty$.

Lemma 3.7: Under the conditions stated above, there exists a positive constant c_2 such that

$$E_\theta^T \left\{ Z_T^{\frac{1}{2}}(u_1) - Z_T^{\frac{1}{2}}(u_2) \right\}^2 \le c_2(u_1 - u_2)^2$$

for $-\infty < u_1, u_2 < \infty$.

Lemma 3.8: Let $\xi(x)$ be a real-valued random function defined on a closed subset F of the Euclidean space R^k. Assume that random process $\xi(x)$ is measurable and separable. Assume that the following conditions are fulfilled: there exist numbers $m \geq r > k$ and a positive continuous function on $G(x) : R^k \to R$ bounded on the compact sets such that for all $x, h \in F, x + h \in F$,

$$E|\xi(x)|^m \leq G(x), \;\; E|\xi(x+h) - \xi(x)|^m \leq G(x)\|h\|^r.$$

Then, with probability one, the realizations of $\xi(t)$ are continuous functions on F. Moreover, let

$$\omega(\delta, \xi, L) = \sup |\xi(x) - \xi(y)|$$

where the upper bound is taken over $x, y \in F$ with $\|x-y\| \leq h, \; \|x\| \leq L, \; \|y\| \leq L$; then

$$E(\omega(h, \xi, L)) \leq B_0 \left(\sup_{\|x\| \leq L} G(x) \right)^{1/m} L^{k/m} h^{(r-k)/m} \log(h^{-1})$$

where the constant B_0 depends on m, r and k.

We will use this lemma with $\xi(u) = Z_T^{1/2}(u), m = 2, r = 2, k = 1, G(x) = e^{-cx^2}$ and $L = H + r + 1$. For a proof of this lemma, see Ibragimov and Khasminskii (1981) (correction, Kallianpur and Selukar (1993)).

Proof of Lemma 3.6: We know that

$$\begin{aligned}
&E_\theta^T(Z_T^{1/2}(u)) \\
&= E_\theta^T \left(\frac{dP_{\theta+u\sqrt{\alpha_T}}^T}{dP_\theta^T} \right)^{1/2} \\
&= E_\theta^T \left\{ \exp\left[\frac{u\sqrt{\alpha_T}}{2} \int_0^T Q_H(t) dM^H(t) - \frac{1}{4} u^2 \alpha_T \int_0^T Q_H^2(t) dw^H(t) \right] \right\} \\
&= E_\theta^T \left\{ \exp\left[\frac{u\sqrt{\alpha_T}}{2} \int_0^T Q_H(t) dM^H(t) - \frac{u^2 \alpha_T}{6} \int_0^T Q_H^2(t) dw^H(t) \right] \right. \\
&\quad \left. \times \exp\left[-\frac{u^2\alpha_T}{12} \int_0^T Q_H^2(t) dw^H(t) \right] \right\} \\
&\leq \left\{ E_\theta^T \left[\exp\left(\frac{1}{2} u\sqrt{\alpha_T} \int_0^T Q_H(t) dM^H(t) \right. \right. \right. \\
&\quad \left. \left. \left. -(1/6) u^2 \alpha_T \int_0^T Q_H^2(t) dw^H(t) \right) \right]^{4/3} \right\}^{3/4}
\end{aligned}$$

$$\times \left\{ E_\theta^T \left[\exp\left(-(1/12)u^2\alpha_T \int_0^T Q_H^2(t)dw^H(t) \right) \right]^4 \right\}^{1/4}$$

(by Holder's inequality)

$$= \left\{ E_\theta^T \left[\exp\left(\frac{2}{3}u\sqrt{\alpha_T} \int_0^T Q_H(t)dM^H(t) \right.\right.\right.$$
$$\left.\left.\left. -(2/9)u^2\alpha_T \int_0^T Q_H^2(t)dw^H(t) \right) \right] \right\}^{3/4}$$
$$\times \left\{ E_\theta^T \exp\left[-(1/3)u^2\alpha_T Q_H^2(t)dt \right] \right\}^{1/4}$$
$$\leq \left\{ E_\theta^T \left[\exp\left(-\frac{1}{3}u^2\alpha_T \int_0^T Q_H^2(t)dw^H(t) \right) \right] \right\}^{1/4}$$

(since the first term is less than or equal to 1 (cf. Gikhman and Skorokhod (1972)). The last term is bounded by $e^{-c_1 u^2}$ for some positive constant c_1 depending on θ and T by assumption (A) which completes the proof of Lemma 3.6.

We now prove Lemma 3.7.

Proof of Lemma 3.7: Note that

$$E_\theta^T \left[Z_T^{\frac{1}{2}}(u_1) - Z_T^{\frac{1}{2}}(u_2) \right]^2 = E_\theta^T \left[Z_T(u_1) + Z_T(u_2) \right] - 2E_\theta^T \left[Z_T^{\frac{1}{2}}(u_1) Z_T^{\frac{1}{2}}(u_2) \right]$$
$$= 2\left\{ 1 - E_\theta^T \left[Z_T^{\frac{1}{2}}(u_1) Z_T^{\frac{1}{2}}(u_2) \right] \right\}$$

since

$$E_\theta^T Z_T(u) = E_\theta^T \left\{ \exp\left[u\sqrt{\alpha_T} \int_0^T Q_H(t)dM^H(t) \right.\right.$$
$$\left.\left. -\frac{1}{2}u^2\alpha_T \int_0^T Q_H^2(t)dw^H(t) \right] \right\} = 1.$$

Denote

$$V_T = \left(\frac{dP_{\theta_2}^T}{dP_{\theta_1}^T} \right)^{1/2}$$

where $\theta_1 = \theta + u_1\sqrt{\alpha_T}$ and $\theta_2 = \theta + u_2\sqrt{\alpha_T}$. Then

$$V_T = \exp\left[\frac{1}{2}(u_2 - u_1)\sqrt{\alpha_T} \int_0^T Q_H(t)dM^H(t) \right.$$
$$\left. -\frac{1}{4}(u_2 - u_1)^2\alpha_T \int_0^T Q_H^2(t)dw^H(t) \right].$$

Now

$$
\begin{aligned}
&E_\theta^T\left[Z_T^{\frac{1}{2}}(u_1)Z_T^{\frac{1}{2}}(u_2)\right]\\
&\quad = E_\theta^T\left[\left(\frac{dP^T_{\theta+u_1\sqrt{\alpha_T}}}{dP_\theta^T}\right)^{1/2}\left(\frac{dP^T_{\theta+u_2\sqrt{\alpha_T}}}{dP_\theta^T}\right)^{1/2}\right]\\
&\quad = \int\left(\frac{dP^T_{\theta_1}}{dP^T_\theta}\right)^{1/2}\left(\frac{dP^T_{\theta_2}}{dP^T_\theta}\right)^{1/2} dP^T_\theta\\
&\quad = \int\left(\frac{dP^T_{\theta_2}}{dP^T_{\theta_1}}\right)^{1/2} dP^T_{\theta_1} = E^T_{\theta_1}(V_T)\\
&\quad = E^T_{\theta_1}\left\{\exp\left[\frac{1}{2}(u_2-u_1)\sqrt{\alpha_T}\int_0^T Q_H(t)dM^H(t)\right.\right.\\
&\qquad \left.\left.-\frac{1}{4}(u_2-u_1)^2\alpha_T\int_0^T Q_H^2(t)dw^H(t)\right]\right\}.
\end{aligned}
$$

Thus

$$
\begin{aligned}
&2\left\{1-E_\theta^T\left[Z_T^{\frac{1}{2}}(u_1)Z_T^{\frac{1}{2}}(u_2)\right]\right\}\\
&\quad = 2\left\{1-E^T_{\theta_1}\left[\exp\left(\frac{1}{2}(u_2-u_1)\sqrt{\alpha_T}\int_0^T Q_H(t)dM^H(t)\right.\right.\right.\\
&\qquad \left.\left.\left.-\frac{1}{4}(u_2-u_1)^2\alpha_T\int_0^T Q_H^2(t)dw^H(t)\right)\right]\right\}\\
&\quad \le 2\left\{1-\exp\left[E^T_{\theta_1}\left(\frac{1}{2}(u_2-u_1)\sqrt{\alpha_T}\int_0^T Q_H(t)dM^H(t)\right.\right.\right.\\
&\qquad \left.\left.\left.-\frac{1}{4}(u_2-u_1)^2\alpha_T\int_0^T Q_H^2(t)dw^H(t)\right)\right]\right\} \text{ (by Jensen's inequality)}\\
&\quad = 2\left\{1-\exp\left[-\frac{(u_2-u_1)^2}{4}\alpha_T E^T_{\theta_1}\int_0^T Q_H^2(t)dw^H(t)\right]\right\}\\
&\quad \le 2\left[\frac{(u_2-u_1)^2}{4}\alpha_T E^T_{\theta_1}\int_0^T Q_H^2(t)dw^H(t)\right] \text{ (since } 1-e^{-x}\le x, x\ge 0)\\
&\quad = c_2(u_2-u_1)^2
\end{aligned}
$$

for some positive constant c_2 depending on θ and T.

Proof of Theorem 3.4: Denote $U = \{u : \theta + u \in \Theta\}$. Let Γ_r be the set of u such that $L + r \leq |u| \leq L + r + 1$. We use the following inequality to prove the theorem:

$$P_\theta^T\{\sup_{\Gamma_r} Z_T(u) \geq 1\} \leq c_3(1 + L + r)^{\frac{1}{2}} e^{-\frac{1}{4}(L+r)^2} \tag{3.25}$$

for some positive constant c_3. Observe that

$$\begin{aligned}
P_\theta^T\{|\alpha_T^{-1/2}(\hat{\theta}_T - \theta)| > L\} &\leq P_\theta^T\{\sup_{|u| > L, u \in U} Z_T(u) \geq Z_T(0)\} \\
&\leq \sum_{r=0}^{\infty} P_\theta^T\{\sup_{\Gamma_r} Z_T(u) \geq 1\} \\
&\leq c_4 \sum_{r=0}^{\infty} e^{-c_5(L+r)^2} \\
&\leq c_6 e^{-c_7 L^2}.
\end{aligned}$$

This proves Theorem 3.4. We now prove the inequality (3.25). We divide the set Γ_r into N subsets $\{\Gamma_r^{(j)}, 1 \leq j \leq N\}$ each with length at most h. The number of such subsets $N \leq [1/h] + 1$. Choose $u_j \in \Gamma_r^{(j)}$, $1 \leq j \leq N$. Then

$$\begin{aligned}
P_\theta^T\{\sup_{u \in \Gamma_r} Z_T(u) \geq 1\} \leq {} & \sum_{j=1}^{N} P_\theta^T\left\{Z_T(u_j) \geq \frac{1}{2}\right\} \\
& + P_\theta^T\left\{\sup_{|u-v| \leq h, |u|, |v| \leq L+r+1} |Z_T^{\frac{1}{2}}(u) - Z_T^{\frac{1}{2}}(v)| \geq \frac{1}{2}\right\}.
\end{aligned} \tag{3.26}$$

From Chebyshev's inequality and in view of Lemma 3.6, it follows that

$$P_\theta^T\left\{Z_T^{\frac{1}{2}}(u_j) \geq \frac{1}{2}\right\} \leq c_8 e^{-(L+r)^2},\ 1 \leq j \leq N$$

for some positive constant c_8. Applying Lemma 3.8 with $\xi(u) = Z_T^{1/2}(u)$, and using Lemma 3.7, we obtain that

$$E_\theta^T\left[\sup_{\substack{|u-v| \leq h \\ |u|, |v| \leq (L+r+1)}} |Z_T^{1/2}(u) - Z_T^{1/2}(v)|\right] \leq c_9(L + r + 1)^{\frac{1}{2}} h^{1/2} \log(h^{-1})$$

for some positive constant c_9. Hence

$$P_\theta^T\{\sup_{u \in \Gamma_r} Z_T(u) \geq 1\} \leq c_{10}\left\{\frac{1}{h} e^{-(L+r)^2} + (L + r + 1)^{\frac{1}{2}} h^{1/2} \log(h^{-1})\right\}$$

for some positive constant c_{10} depending on θ and T by using (3.26). Choosing $h = e^{-(L+r)^2/2}$, we prove the inequality in Theorem 3.4.

Proof of Theorem 3.5: Observe that conditions (1) and (2) in Theorem 5.2 in Chapter I of Ibragimov and Khasminskii (1981) are satisfied by Lemmas 3.6 and 3.7. In view of the conditions D(i)–D(v) on the loss function mentioned in Section 3.4, we can prove Theorem 3.5 by using Theorem 5.2 in Chapter I of Ibragimov and Khasminskii (1981) with $\alpha = 2$ and $g(u) = u^2$. We omit the details.

Remarks: Bahadur (1960) suggested measuring the asymptotic efficiency of an estimator δ_T of a parameter θ by the magnitude of concentration of the estimator over the interval of a fixed length (independent of T), that is, by the magnitude of the probability $P_\theta\{|\delta_T - \theta| < \gamma\}$. From the result obtained in Theorem 3.4 proved above, we note that the probability $P_\theta\{|\hat{\theta}_T - \theta| < \gamma\}$ is bounded above by $C_1 e^{-C_2\gamma^2\alpha_T^{-1}}$, $C_1 > 0, C_2 > 0$ for the MLE $\hat{\theta}_T$. This bound in turn decreases exponentially to zero as $T \to \infty$ for any fixed $\gamma > 0$. Following the techniques in Theorem 9.3 of Chapter I in Ibragimov and Khasminskii (1981), it can be shown that the MLE is Bahadur efficient under some additional conditions. A similar result follows for the Bayes estimator $\tilde{\theta}_T$ following Theorem 3.5. The norming factor α_T can be chosen to be T^{-1} if $\theta < 0$, T^{-2} if $\theta = 0$ and $e^{-\theta T}T^{1/2}$ in case $\theta > 0$. This can be seen from Proposition 2.3 of Kleptsyna and Le Breton (2002a). Observe that the norming factor α_T tends to zero as $T \to \infty$.

The results discussed here are due to Mishra and Prakasa Rao (2006).

Sharp large deviations for MLE

Bercu *et al*. (2008) have recently obtained results on sharp large deviations for fractional Ornstein–Uhlenbeck-type processes. We will briefly discuss their results.

Consider the fractional Ornstein–Uhlenbeck-type process governed by the SDE

$$dX_t = \theta X_t dt + dW_t^H, \; X_0 = 0 \tag{3.27}$$

where $\theta < 0$ and $H > \frac{1}{2}$. Let

$$k_H(t,s) = k_H^{-1} s^{\frac{1}{2}-H}(t-s)^{\frac{1}{2}-H}, \; 0 < s < t$$

where k_H is as defined by (3.4),

$$Z_t = \int_0^t k_H(t,s)dX_s, \quad M_t = \int_0^t k_H(t,s)dW_s^H$$

with

$$\langle M \rangle_t = w_t^H = \lambda_H^{-1} t^{2-2H}$$

and λ_H as given in (3.6). From the earlier discussions, we know that

$$Z_t = \theta \int_0^t k_H(t,s) X_s ds + M_t$$

and

$$Z_t = \theta \int_0^t Q_s d\langle M \rangle_s + M_t$$

where

$$Q_t = \frac{\ell_H}{2} \left(t^{2H-1} Z_t + \int_0^t s^{2H-1} dZ_s \right), \quad \ell_H = \lambda_H / [2(1-H)].$$

This follows from results in Kleptsyna and Le Breton (2002a). The score function, which is the derivative of the log-likelihood function based on the observations over the interval $[0, T]$, is given by

$$\Sigma_T(\theta) = \int_0^T Q_t dZ_t - \theta \int_0^T Q_t^2 d\langle M \rangle_t.$$

Let

$$S_t = \int_0^T Q_t^2 d\langle M \rangle_t.$$

Bishwal (2008) studied the large-deviation properties of the log-likelihood ratio

$$(\theta - \theta_1) \int_0^T Q_t dZ_t - \frac{\theta^2 - \theta_1^2}{2} \int_0^T Q_t^2 d\langle M \rangle_t$$

for $\theta \neq \theta_1$. We will not discuss these properties.

Definition: A family of random variables $\{V_t, t \geq 0\}$ is said to satisfy the *large-deviation principle* (LDP) with *rate function* $I(.)$ if the function $I(.)$ is lower semi continuous from R to $[0, \infty]$ such that, for any closed set $F \subset R$,

$$\limsup_{T \to \infty} \frac{1}{t} \log P(V_t \in F) \leq - \inf_{x \in F} I(x).$$

and, for any open set $G \subset R$,

$$- \inf_{x \in G} I(x) \leq \liminf_{T \to \infty} P(V_t \in G)$$

The function $I(.)$ is called a *good rate function* if the level sets, that is, sets of the form $\{x : I(x) \leq c\}$, are compact sets in R (cf. Dembo and Zeitouni (1998)).

For any a, b in R, let

$$L_T(a,b) = \frac{1}{T} \log E[e^{R_T(a,b)}]$$

where

$$R_T(a,b) = a \int_0^T Q_t dZ_t + b \int_0^T Q_t^2 d\langle M \rangle_t.$$

Bercu *et al*. (2008) proved the following theorem.

Theorem 3.9: The following properties hold:

(i) $\lim_{T\to\infty} S_T/T = -1/2\theta$ a.s.;

(ii) $\frac{1}{\sqrt{T}}\left(S_T + \frac{T}{2\theta}\right) \xrightarrow{\mathcal{L}} N\left(0, -\frac{1}{2\theta^3}\right)$ as $T \to \infty$; and

(iii) the process $\{S_T/T, T \geq 0\}$ satisfies the LDP with good rate function $I(.)$ defined by

$$\begin{aligned} I(c) &= \frac{(2\theta c+1)^2}{8c} && \text{if } 0 < c \leq -\frac{1}{2\theta\delta_H} \\ &= \frac{c\theta^2}{2}(1-\delta_H^2) + \frac{\theta}{2}(1-\delta_H) && \text{if } c \geq -\frac{1}{2\theta\delta_H} \\ &= \infty && \text{otherwise} \end{aligned} \tag{3.28}$$

where

$$\delta_H = \frac{1-\sin(\pi H)}{1+\sin(\pi H)}. \tag{3.29}$$

Note that $\delta_H = 0$ if $H = \frac{1}{2}$. Bercu *et al*. (2008) derived the sharp LDP for the process $\{S_T/T\}$ following the ideas of Bahadur and Ranga Rao (1960).

We saw earlier that the MLE is the solution of the equation $\Sigma_T(\theta) = 0$ and it is given by

$$\hat{\theta}_T = \frac{\int_0^T Q_t dZ_t}{\int_0^T Q_t^2 d\langle M \rangle_t}.$$

It was mentioned earlier that the MLE $\hat{\theta}_T$ is strongly consistent as $T \to \infty$. It is easy to check that

$$P(\sqrt{T}(\hat{\theta}_T - \theta) \leq c) = P(V_T(c) \leq 0)$$

where

$$V_T(c) = \frac{1}{\sqrt{T}} \int_0^T Q_s dZ_s - \left(\frac{c}{\sqrt{T}} + \theta \right) \frac{S_T}{\sqrt{T}}. \tag{3.30}$$

Bercu *et al*. (2008) proved that

$$V_T(c) \xrightarrow{\mathcal{L}} V(c) \text{ as } T \to \infty$$

where the random variable $V(c)$ has the Gaussian distribution with mean $c/2\theta$ and variance $-1/2\theta$. Note that

$$P(V(c) \le 0) = -\frac{1}{4\pi\theta} \int_{-\infty}^{c} e^{x^2/2\theta} dx.$$

Hence

$$\sqrt{T}(\hat{\theta}_T - \theta) \xrightarrow{\mathcal{L}} N(0, -2\theta) \text{ as } T \to \infty.$$

Bercu *et al*. (2008) proved that the MLE $\hat{\theta}_T$ obeys LDP with a good rate function.

Theorem 3.10: The MLE $\hat{\theta}_T$ satisfies the LDP with the good rate function

$$\begin{aligned} I(c) &= -\frac{(c-\theta)^2}{4c} \quad \text{if } c < \frac{\theta}{3} \\ &= 2c - \theta \quad \text{if } c \ge \frac{\theta}{3}. \end{aligned} \tag{3.31}$$

Observe that the rate function is independent of the Hurst index H.

3.6 Minimum L_1-norm estimation

In spite of the fact that MLEs are consistent and asymptotically normal, and also asymptotically efficient in general, they have some shortcomings at the same time. Their calculation is often cumbersome as the expression for MLE involves stochastic integrals which need good approximations for computational purposes. Furthermore, the MLE is not robust in the sense that a slight perturbation in the noise component will change the properties of the MLE substantially. In order to circumvent such problems, the minimum distance approach is proposed. Properties of the minimum distance estimators (MDEs) were discussed in Millar (1984) in a general framework. We now obtain the minimum L_1-norm estimates of the drift parameter of a fractional Ornstein–Uhlenbeck-type process and investigate the asymptotic properties of such estimators following the work of Kutoyants and Pilibossian (1994).

We now consider the problem of estimation of the parameter θ based on the observation of a fractional Ornstein–Uhlenbeck-type process $X = \{X_t, 0 \leq t \leq T\}$ satisfying the SDE

$$dX_t = \theta X(t)dt + \varepsilon dW_t^H,\ X_0 = x_0,\ 0 \leq t \leq T \tag{3.32}$$

for a fixed time T where $\theta \in \Theta \subset R$ and study its asymptotic properties as $\varepsilon \to 0$.

Let $x_t(\theta)$ be the solution of the above differential equation with $\varepsilon = 0$. It is obvious that

$$x_t(\theta) = x_0 e^{\theta t},\ 0 \leq t \leq T. \tag{3.33}$$

Let

$$S_T(\theta) = \int_0^T |X_t - x_t(\theta)|dt. \tag{3.34}$$

We define θ_ε^* to be a *minimum L_1-norm estimator* if there exists a measurable selection θ_ε^* such that

$$S_T(\theta_\varepsilon^*) = \inf_{\theta \in \Theta} S_T(\theta). \tag{3.35}$$

Conditions for the existence of a measurable selection are given in Lemma 3.1.2 in Prakasa Rao (1987). We assume that there exists a measurable selection θ_ε^* satisfying the above equation. An alternate way of defining the estimator θ_ε^* is by the relation

$$\theta_\varepsilon^* = \arg \inf_{\theta \in \Theta} \int_0^T |X_t - x_t(\theta)|dt. \tag{3.36}$$

Consistency

Let θ_0 denote the true parameter. For any $\delta > 0$, define

$$g(\delta) = \inf_{|\theta - \theta_0| > \delta} \int_0^T |X_t(\theta) - x_t(\theta_0)|dt. \tag{3.37}$$

Note that $g(\delta) > 0$ for any $\delta > 0$.

Theorem 3.11: For every $p > 0$, there exists a constant $K(p, H)$ such that for every $\delta > 0$,

$$\begin{aligned} P_{\theta_0}^{(\varepsilon)}\{|\theta_\varepsilon^* - \theta_0| > \delta\} &\leq 2^p T^{pH+p} K(p, H) e^{|\theta_0|Tp} (g(\delta))^{-p} \varepsilon^p \\ &= O((g(\delta))^{-p} \varepsilon^p). \end{aligned} \tag{3.38}$$

Proof: Let $||.||$ denote the L_1-norm. Then

$$\begin{aligned}
P_{\theta_0}^{(\varepsilon)}\{|\theta_\varepsilon^* - \theta_0| > \delta\} &= P_{\theta_0}^{(\varepsilon)}\{\inf_{|\theta-\theta_0|\le\delta} ||X - x(\theta)|| > \inf_{|\theta-\theta_0|>\delta} ||X - x(\theta)||\} \\
&\le P_{\theta_0}^{(\varepsilon)}\{\inf_{|\theta-\theta_0|\le\delta}(||X - x(\theta_0)|| + ||x(\theta) - x(\theta_0)||) \\
&\quad > \inf_{|\theta-\theta_0|>\delta}(||x(\theta) - x(\theta_0)|| - ||X - x(\theta_0)||)\} \\
&= P_{\theta_0}^{(\varepsilon)}\{2||X - x(\theta_0)|| > \inf_{|\theta-\theta_0|>\delta} ||x(\theta) - x(\theta_0)||\} \\
&= P_{\theta_0}^{(\varepsilon)}\{||X - x(\theta_0)|| > \frac{1}{2}g(\delta)\}. \qquad (3.39)
\end{aligned}$$

Since the process X_t satisfies the SDE (3.32), it follows that

$$\begin{aligned}
X_t - x_t(\theta_0) &= x_0 + \theta_0 \int_0^t X_s ds + \varepsilon W_t^H - x_t(\theta_0) \\
&= \theta_0 \int_0^t (X_s - x_s(\theta_0))ds + \varepsilon W_t^H \qquad (3.40)
\end{aligned}$$

since $x_t(\theta) = x_0 e^{\theta t}$. Let $U_t = X_t - x_t(\theta_0)$. Then it follows from the above equation that

$$U_t = \theta_0 \int_0^t U_s \; ds + \varepsilon W_t^H. \qquad (3.41)$$

Let $V_t = |U_t| = |X_t - x_t(\theta_0)|$. The above relation implies that

$$V_t = |X_t - x_t(\theta_0)| \le |\theta_0| \int_0^t V_s ds + \varepsilon |W_t^H|. \qquad (3.42)$$

Applying the Gronwall–Bellman lemma (cf. Kutoyants (1994), Lemma 1.11, p. 25), we obtain

$$\sup_{0\le t\le T} |V_t| \le \varepsilon e^{|\theta_0 T|} \sup_{0\le t\le T} |W_t^H|. \qquad (3.43)$$

Hence

$$\begin{aligned}
P_{\theta_0}^{(\varepsilon)}\left\{||X - x(\theta_0)|| > \frac{1}{2}g(\delta)\right\} &\le P\left\{\sup_{0\le t\le T} |W_t^H| > \frac{e^{-|\theta_0 T|}g(\delta)}{2\varepsilon T}\right\} \\
&= P\left\{W_T^{H*} > \frac{e^{-|\theta_0 T|}g(\delta)}{2\varepsilon T}\right\}. \qquad (3.44)
\end{aligned}$$

Applying Proposition 1.9 of Chapter 1 (cf. Novikov and Valkeila (1999)) to the upper bound obtained above, we get

$$P_{\theta_0}^{(\varepsilon)}\{|\theta_\varepsilon^* - \theta_0| > \delta\} \leq 2^p T^{pH+p} K(p,H) e^{|\theta_0 T|p} (g(\delta))^{-p} \varepsilon^p$$
$$= O((g(\delta))^{-p} \varepsilon^p). \tag{3.45}$$

Remarks: As a consequence of the above theorem, we obtain that θ_ε^* converges in probability to θ_0 under $P_{\theta_0}^{(\varepsilon)}$-measure as $\varepsilon \to 0$. Furthermore, the rate of convergence is of the order $(O(\varepsilon^p))$ for every $p > 0$.

Asymptotic distribution

We will now study the asymptotic distribution of the estimator θ_ε^* after suitable scaling. It can be checked that

$$X_t = e^{\theta_0 t}\left\{x_0 + \int_0^t e^{-\theta_0 s} \varepsilon dW_s^H\right\} \tag{3.46}$$

or equivalently

$$X_t - x_t(\theta_0) = \varepsilon e^{\theta_0 t} \int_0^t e^{-\theta_0 s} dW_s^H. \tag{3.47}$$

Let

$$Y_t = e^{\theta_0 t} \int_0^t e^{-\theta_0 s} dW_s^H. \tag{3.48}$$

Note that $\{Y_t, 0 \leq t \leq T\}$ is a Gaussian process and can be interpreted as the 'derivative' of the process $\{X_t, 0 \leq t \leq T\}$ with respect to ε. Applying Theorem 1.18 in Chapter 1, we obtain that, P-a.s.,

$$Y_t e^{-\theta_0 t} = \int_0^t e^{-\theta_0 s} dW_s^H = \int_0^t K_H^f(t,s) dM_s^H,\ t \in [0,T] \tag{3.49}$$

where $f(s) = e^{-\theta_0 s}$, $s \in [0,T]$, and M^H is the fundamental Gaussian martingale associated with fBm W^H. In particular it follows that the random variable $Y_t e^{-\theta_0 t}$ and hence Y_t have a normal distribution with mean zero and, furthermore, for any $h \geq 0$,

$$cov(Y_t, Y_{t+h}) = e^{2\theta_0 t + \theta_0 h} E\left(\int_0^t e^{-\theta_0 u} dW_u^H \int_0^{t+h} e^{-\theta_0 v} dW_v^H\right)$$
$$= e^{2\theta_0 t + \theta_0 h} H(2H-1) \int_0^t \int_0^{t+h} e^{-\theta_0(u+v)} |u-v|^{2H-2} du dv$$
$$= e^{2\theta_0 t + \theta_0 h} \gamma_H(t, t+h)\ \text{(say)}. \tag{3.50}$$

In particular

$$Var(Y_t) = e^{2\theta_0 t}\gamma_H(t,t). \tag{3.51}$$

Hence $\{Y_t, 0 \leq t \leq T\}$ is a zero-mean Gaussian process with $Cov(Y_t, Y_s) = e^{\theta_0(t+s)}\gamma_H(t,s)$ for $s \geq t$.

Let

$$\zeta = \arg \inf_{-\infty<u<\infty} \int_0^T |Y_t - utx_0 e^{\theta_0 t}|dt. \tag{3.52}$$

Theorem 3.12: As $\varepsilon \to 0$, the random variable $\varepsilon^{-1}(\theta_\varepsilon^* - \theta_0)$ converges in probability to a random variable whose probability distribution is the same as that of ζ under P_{θ_0}.

Proof: Let $x_t'(\theta) = x_0 t e^{\theta t}$ and let

$$Z_\varepsilon(u) = ||Y - \varepsilon^{-1}(x(\theta_0 + \varepsilon u) - x(\theta_0))|| \tag{3.53}$$

and

$$Z_0(u) = ||Y - ux'(\theta_0)||. \tag{3.54}$$

Furthermore, let

$$A_\varepsilon = \{\omega : |\theta_\varepsilon^* - \theta_0| < \delta_\varepsilon\},\ \delta_\varepsilon = \varepsilon^\tau,\ \tau \in (1/2, 1),\ L_\varepsilon = \varepsilon^{\tau-1}. \tag{3.55}$$

Observe that the random variable $u_\varepsilon^* = \varepsilon^{-1}(\theta_\varepsilon^* - \theta_0)$ satisfies the equation

$$Z_\varepsilon(u_\varepsilon^*) = \inf_{|u|<L_\varepsilon} Z_\varepsilon(u),\ \omega \in A_\varepsilon. \tag{3.56}$$

Define

$$\zeta_\varepsilon = \arg \inf_{|u|<L_\varepsilon} Z_0(u). \tag{3.57}$$

Observe that, with probability one,

$$\begin{aligned}\sup_{|u|<L_\varepsilon} |Z_\varepsilon(u) - Z_0(u)| &= \left|\left|\left|Y - ux'(\theta_0) - \frac{1}{2}\varepsilon u^2 x''(\tilde{\theta})\right|\right| - ||Y - ux'(\theta_0)||\right| \\ &\leq \frac{\varepsilon}{2} L_\varepsilon^2 \sup_{|\theta-\theta_0<\delta_\varepsilon} \int_0^T |x''(\theta)|dt \\ &\leq C\varepsilon^{2\tau-1}. \end{aligned} \tag{3.58}$$

Here $\tilde{\theta} = \theta_0 + \alpha(\theta - \theta_0)$ for some $\alpha \in (0, 1)$. Note that the last term in the above inequality tends to zero as $\varepsilon \to 0$. Furthermore, the process $\{Z_0(u), -\infty < u <$

$\infty\}$ has a unique minimum u^* with probability one. This follows from the arguments given in Theorem 2 of Kutoyants and Pilibossian (1994). In addition, we can choose the interval $[-L, L]$ such that

$$P_{\theta_0}^{(\varepsilon)}\{u_\varepsilon^* \in (-L, L)\} \geq 1 - \beta g(L)^{-p} \tag{3.59}$$

and

$$P\{u^* \in (-L, L)\} \geq 1 - \beta g(L)^{-p} \tag{3.60}$$

where $\beta > 0$. Note that $g(L)$ increases as L increases. The processes $Z_\varepsilon(u), u \in [-L, L]$, and $Z_0(u), u \in [-L, L]$, satisfy the Lipschitz conditions and $Z_\varepsilon(u)$ converges uniformly to $Z_0(u)$ over $u \in [-L, L]$. Hence the minimizer of $Z_\varepsilon(.)$ converges to the minimizer of $Z_0(u)$. This completes the proof.

Remarks:

(i) We saw earlier that the process $\{Y_t, 0 \leq t \leq T\}$ is a zero-mean Gaussian process with the covariance function

$$Cov(Y_t, Y_s) = e^{\theta_0(t+s)}\gamma_H(t, s)$$

for $s \geq t$. Recall that

$$\zeta = \arg \inf_{-\infty<u<\infty} \int_0^T |Y_t - utx_0e^{\theta_0 t}|dt. \tag{3.61}$$

It is not clear what the distribution of ζ is. Observe that for every u, the integrand in the above integral is the absolute value of a Gaussian process $\{J_t, 0 \leq t \leq T\}$ with the mean function $E(J_t) = -utx_0e^{\theta_0 t}$ and the covariance function

$$Cov(J_t, J_s) = e^{\theta_0(t+s)}\gamma_H(t, s)$$

for $s \geq t$. It would be interesting to say something about the distribution of ζ through simulation studies even if an explicit computation of the distribution seems to be difficult. The results studied here are due to Prakasa Rao (2004a).

(ii) Hu and Nualart (2009) proposed a type of LSE for the parameter θ and studied its properties. This estimator is motivated by the following heuristic argument. The LSE is obtained by minimizing

$$\int_0^T |x_t' - \theta x_t|^2 dt$$

which leads to the solution

$$\hat{\theta}_T = \frac{\int_0^T X_t dX_t}{\int_0^T X_t^2 dt}.$$

Hu and Nualart (2009) call this estimator the LSE and interpret the integral $\int_0^T X_t dW_t^H$ as a divergence-type integral which we have not used or discussed in this book. We do not go into the details here. As this estimator is difficult to compute, they suggest another estimator

$$\tilde{\theta}_T = \left[\frac{1}{H\Gamma(2H)T}\int_0^T X_t^2 dt\right]^{-1/2H}.$$

It was shown that this estimator is strongly consistent and asymptotically normal as $T \to \infty$. In fact

$$\sqrt{T}(\tilde{\theta}_T - \theta) \xrightarrow{\mathcal{L}} N\left(0, -\frac{\theta}{4H^2}\sigma_H^2\right)$$

as $T \to \infty$ where

$$\sigma_H^2 = (4H-1)\left[1 + \frac{\Gamma(3-4H)\Gamma(4H-1)}{\Gamma(2-2H)\Gamma(2H)}\right].$$

(iii) It is well known that the sequential estimation methods might lead to efficient estimators from the process observed, possibly over a shorter expected period of observation time as compared to observation over a fixed observation time. We have investigated the conditions for such a phenomenon for estimating the drift parameter of a fractional Ornstein–Uhlenbeck-type process in Prakasa Rao (2004b). Novikov (1972) investigated the asymptotic properties of a sequential MLE for the drift parameter in the Ornstein–Uhlenbeck process.

4

Sequential inference for some processes driven by fBm

4.1 Introduction

Parametric estimation for classes of stochastic processes, satisfying linear SDEs driven by fBm and observed over a fixed period of time T, was studied in Chapter 2. It is well known that the sequential estimation methods might lead to efficient estimators from a process observed possibly over a shorter expected period of observation time as compared to estimators based on predetermined fixed observation time. We now investigate the conditions for such a phenomenon. Novikov (1972) investigated the asymptotic properties of a sequential MLE for the drift parameter in the Ornstein–Uhlenbeck process. We now discuss analogous results for the fractional Ornstein–Uhlenbeck-type process. We further study the problem of sequential testing for a simple hypothesis that the observable process is noise modeled by fBm against the simple hypothesis that the process contains an unobservable signal along with noise. We follow the notation used in Chapter 1.

4.2 Sequential maximum likelihood estimation

Consider the fractional Ornstein–Uhlenbeck-type process defined by the SDE

$$dX_t = \theta X_t dt + dW_t^H, \quad t \geq 0 \tag{4.1}$$

Statistical Inference for Fractional Diffusion Processes B. L. S. Prakasa Rao

where θ is an unknown parameter. We now consider the problem of sequential maximum likelihood estimation of the parameter θ. Let h be a nonnegative number. Define the stopping rule $\tau(h)$ by the rule

$$\tau(h) = \inf\left\{t : \int_0^t Q^2(s)dw_s^H \geq h\right\} \tag{4.2}$$

where the process $\{Q(t), t \geq 0\}$ is as defined in (3.10) of Chapter 3. Kleptsyna and Le Breton (2002a) have shown that

$$\lim_{t\to\infty} \int_0^t Q^2(s)dw_s^H = +\infty \text{ a.s. } [P_\theta] \tag{4.3}$$

for every $\theta \in R$. Then it can be shown that $P_\theta(\tau(h) < \infty) = 1$. If the process is observed up to a previously determined time T , we have observed that the MLE is given by

$$\hat{\theta}_T = \left[\int_0^T Q^2(s)dw_s^H\right]^{-1} \int_0^T Q(s)dZ_s. \tag{4.4}$$

The estimator

$$\begin{aligned}\hat{\theta}(h) &\equiv \hat{\theta}_{\tau(h)}\\ &= \left[\int_0^{\tau(h)} Q^2(s)dw_s^H\right]^{-1} \int_0^{\tau(h)} Q(s)dZ_s\\ &= h^{-1}\int_0^{\tau(h)} Q(s)dZ_s \end{aligned}\tag{4.5}$$

is called the *sequential maximum likelihood estimator* of θ. We now study the asymptotic properties of the estimator $\hat{\theta}(h)$.

We will first prove a lemma which is an analogue of the Cramér–Rao inequality for sequential plans $(\tau(X), \hat{\theta}_\tau(X))$ for estimating the parameter θ satisfying the property

$$E_\theta\{\hat{\theta}_\tau(X)\} = \theta \tag{4.6}$$

for all θ.

Lemma 4.1: Suppose that differentiation under the integral sign with respect to θ on the left of Equation (4.6) is permissible. Further suppose that

$$E_\theta\left\{\int_0^{\tau(X)} Q^2(s)dw_s^H\right\} < \infty \tag{4.7}$$

for all θ. Then

$$Var_\theta\{\hat{\theta}_\tau(X)\} \geq \left(E_\theta\left\{\int_0^{\tau(X)} Q^2(s)dw_s^H\right\}\right)^{-1} \tag{4.8}$$

for all θ.

Proof: Let P_θ be the measure generated by the process $\{X(t), 0 \leq t \leq \tau(X)\}$ for given θ. Following Theorem 1.20 in Chapter 1, it can be shown that

$$\frac{dP_\theta}{dP_{\theta_0}} = \exp\left[(\theta-\theta_0)\int_0^{\tau(X)} Q(s)dZ_s - \frac{1}{2}(\theta^2-\theta_0^2)\int_0^{\tau(X)} Q^2(s)dw_s^H\right] \quad \text{a.s. } [P_{\theta_0}] \tag{4.9}$$

by using Sudakov's lemma (cf. Basawa and Prakasa Rao (1980), p. 352). Differentiating the functions on both sides of Equation (4.6) with respect to θ under the integral sign, we get that

$$E_\theta\left[\hat{\theta}_\tau(X)\left\{\int_0^{\tau(X)} Q(s)dZ_s - \theta\int_0^{\tau(X)} Q^2(s)dw_s^H\right\}\right] = 1. \tag{4.10}$$

Theorem 1.19 from Chapter 1 implies that

$$dZ_s = \theta Q(s)dw_s^H + dM_s^H \tag{4.11}$$

and hence

$$\int_0^T Q(s)dZ_s = \theta\int_0^T Q^2(s)dw_s^H + \int_0^T Q(s)dM_s^H. \tag{4.12}$$

The above relation in turn implies that

$$E_\theta\left\{\int_0^{\tau(X)} Q(s)dZ_s - \theta\int_0^{\tau(X)} Q^2(s)dw_s^H\right\} = 0 \tag{4.13}$$

and

$$E_\theta\left\{\int_0^{\tau(X)} Q(s)dZ_s - \theta\int_0^{\tau(X)} Q^2(s)dw_s^H\right\}^2 = E_\theta\left\{\int_0^{\tau(X)} Q^2(s)dw_s^H\right\} \tag{4.14}$$

from the properties of the fundamental martingale M^H and the fact that the quadratic variation $\langle M^H\rangle_t$ of the process M^H is w_t^H. Applying the Cauchy–Schwartz inequality to the left of Equation (4.14), we obtain that

$$Var_\theta\{\hat{\theta}_\tau(X)\} \geq \left(E_\theta\left\{\int_0^{\tau(X)} Q^2(s)dw_s^H\right\}\right)^{-1} \tag{4.15}$$

for all θ.

Definition: A sequential plan $(\tau(X), \hat{\theta}_\tau(X))$ is said to be *efficient* if there is equality in (4.8) for all θ.

We now prove the main result indicating the efficiency of the sequential plan defined above.

Theorem 4.2: Consider the fractional Ornstein–Uhlenbeck-type process governed by the SDE (4.1) driven by fBm W^H with $H \in [\frac{1}{2}, 1)$. Then the sequential plan $(\tau(h), \hat{\theta}(h))$ defined by Equations (4.2) and (4.5) has the following properties for all θ:

(i) $\hat{\theta}(h) \equiv \hat{\theta}_{\tau(h)}$ is normally distributed with $E_\theta(\hat{\theta}(h)) = \theta$ and $Var_\theta(\hat{\theta}(h)) = h^{-1}$;

(ii) the plan is efficient; and

(iii) the plan is closed, that is, $P_\theta(\tau(h) < \infty) = 1$.

Proof: Let

$$J_T = \int_0^T Q(s) dM_s^H. \tag{4.16}$$

From the results in Kartazas and Shreve (1988), Revuz and Yor (1991) and Ikeda and Watanabe (1981), it follows that there exists a standard Wiener process W such that

$$J_T = W(\langle J \rangle_T) \text{ a.s.} \tag{4.17}$$

with respect to the filtration $\{\mathcal{F}_{\tau_t}, t \geq 0\}$ under P where $\tau_t = \inf\{s : \langle J \rangle_s > t\}$. Hence the process

$$\int_0^{\tau(h)} Q(s) dM_s^H \tag{4.18}$$

is a standard Wiener process. Observe that

$$\begin{aligned}
\hat{\theta}(h) &= h^{-1} \int_0^{\tau(h)} Q(s) dZ_s \\
&= h^{-1} \left[\theta \int_0^{\tau(h)} Q^2(s) dw_s^H + \int_0^{\tau(h)} Q(s) dM_s^H \right] \\
&= \theta + h^{-1} \int_0^{\tau(h)} Q(s) dM_s^H \\
&= \theta + h^{-1} J_{\tau(h)} \\
&= \theta + h^{-1} W(\langle J \rangle_{\tau(h)})
\end{aligned} \tag{4.19}$$

which proves that the estimator $\hat{\theta}(h)$ is normally distributed with mean θ and variance h^{-1}. Since

$$E_\theta \left\{ \int_0^{\tau(h)} Q^2(s) dw_s^H \right\} = h, \tag{4.20}$$

it follows that the plan is efficient by Lemma 4.1. Since

$$P_\theta(\tau(h) \geq T) = P_\theta \left(\int_0^T Q^2(s) dw_s^H < h \right) \tag{4.21}$$

for every $T \geq 0$, it follows that $P_\theta(\tau(h) < \infty) = 1$ from the observation

$$P_\theta \left(\int_0^\infty Q^2(s) dw_s^H = \infty \right) = 1. \tag{4.22}$$

The results in this section are from Prakasa Rao (2004b).

4.3 Sequential testing for simple hypothesis

We now study the sequential testing problem for a simple null hypothesis that an observable process is a noise modeled by fBm against the simple alternate hypothesis that the process also contains an unobservable signal along with the noise. The motivation for the present study comes from the observation that the problem can be looked at as modeling in the branch of signal processing. Suppose we surmise that a signal (which is unobserved) is possibly transmitted over a channel corrupted by fBm. We are interested in testing the simple hypothesis that there is no transmitted signal but only noise modeled by fBm that is transmitted through the channel against the hypothesis that a signal is transmitted corrupted by noise modeled by fBm. We prove the existence of an optimal sequential testing procedure for such a problem. The results obtained are analogues of similar results for diffusion processes derived in Liptser and Shiryayev (2001) and are due to Prakasa Rao (2005b).

Suppose that $\theta = \{\theta_t, t \geq 0\}$ is an unobservable $\mathcal{F}_t$-adapted process independent of fBm $W = \{W_t^H, t \geq 0\}$. Suppose that one of the following two hypotheses holds for the $\mathcal{F}_t$-adapted observable process $\psi = \{\psi_t, t \geq 0\}$:

$$H_0 : d\psi_t = dW_t^H, \quad \psi_0 = 0, t \geq 0; \tag{4.23}$$

and

$$H_1 : d\psi_t = \theta_t dt + dW_t^H, \quad \psi_0 = 0, t \geq 0. \tag{4.24}$$

If we interpret the process θ as a signal and fBm W^H as the noise, then we are interested in testing the simple hypothesis H_1 indicating the presence of the signal in the observation of the process ψ against the simple hypothesis H_0 that the signal θ is absent. Assume that the sample paths of the process $\{\theta_t, t \geq 0\}$ are smooth enough so that the process

$$Q(t) = \frac{d}{dw_t^H} \int_0^t k_H(t,s)\theta_s ds, \quad t \geq 0 \tag{4.25}$$

is well defined almost everywhere where w_t^H and $k_H(t,s)$ are as defined in (1.72) and (1.70) in Chapter 1 respectively. Suppose the sample paths of the process $\{Q(t), 0 \leq t \leq T\}$ belong a.s. to $L^2([0,T], dw_t^H)$ for every $T \geq 0$. Define

$$Z_t = \int_0^t k_H(t,s) d\psi_s, \quad t \geq 0. \tag{4.26}$$

Then the process $Z = \{Z_t, t \geq 0\}$ is an $(\mathcal{F}_t)$-semimartingale with the decomposition

$$Z_t = \int_0^t Q(s) dw_s^H + M_t^H \tag{4.27}$$

where M^H is the fundamental martingale defined by (1.73) of Chapter 1 and the process ψ admits the representation

$$\psi_t = \int_0^t K_H(t,s) dZ_s. \tag{4.28}$$

Here the function $K_H(.,.)$ is given by (1.75) in Chapter 1 with $f \equiv 1$. We denote the probability measure of the process ψ under H_i as P_i for $i = 0, 1$. Let E denote the expectation under the probability measure P and E_i denote the expectation under the hypothesis $H_i, i = 0, 1$. Let P_i^T be the measure induced by the process $\{\psi_t, 0 \leq t \leq T\}$ under the hypothesis H_i. Following Theorem 1.20 in Chapter 1, we get that the Radon–Nikodym derivative of P_1^T with respect to P_0^T is given by

$$\frac{dP_1^T}{dP_0^T} = \exp\left[\int_0^T Q(s) dZ_s - \frac{1}{2}\int_0^T Q^2(s) dw_s^H\right]. \tag{4.29}$$

Let us consider the sequential plan $\Delta = \Delta(\tau, \delta)$ for testing H_0 versus H_1 characterized by the stopping time τ and the decision function δ. We assume that τ is a stopping time with respect to the family of σ-algebras $\mathcal{B}_t = \sigma\{x : x_s, s \leq t\}$ where $x = \{x_t, t \geq 0\}$ are continuous functions with $x_0 = 0$. The decision function $\delta = \delta(x)$ is $\mathcal{B}_\tau$-measurable and takes the values 0 and 1. Suppose x is the observed sample path. If $\delta(x)$ takes the value 0, then it amounts to the decision that the hypothesis H_0 is accepted; and if $\delta(x)$ takes the value 1, then it will indicate the acceptance of the hypothesis H_1. For any sequential plan $\Delta = \Delta(\tau, \delta)$, define

$$\alpha(\Delta) = P_1(\delta(\psi) = 0), \;\; \beta(\Delta) = P_0(\delta(\psi) = 1).$$

Observe that $\alpha(\Delta)$ and $\beta(\Delta)$ are the first and second kinds of errors respectively. Let $\Delta_{\alpha,\beta}$ be the class of sequential plans for which

$$\alpha(\Delta) \leq \alpha, \quad \beta(\Delta) \leq \beta$$

where $0 < \alpha + \beta < 1$, and

$$E_i\left(\int_0^{\tau(\psi)} m_t^2(\psi)dw_t^H\right) < \infty, \quad i = 0, 1. \tag{4.30}$$

where $m_t(\psi) = E_1(Q(t)|\mathcal{F}_t^{\psi})$ and $\{\mathcal{F}_t^{\psi}\}$ is the filtration generated by the process ψ. We now state the main theorem giving the *optimum sequential plan* subject to the conditions stated above.

Theorem 4.3: Suppose the process $Q = \{Q(t), \mathcal{F}_t, t \geq 0\}$ defined above satisfies the condition

$$E|Q(t)| < \infty, \quad 0 \leq t < \infty. \tag{4.31}$$

Let

$$m_t(\psi) = E_1(Q(t)|\mathcal{F}_t^{\psi}). \tag{4.32}$$

Suppose that

$$P_i\left(\int_0^{\infty} m_t^2(\psi)dw_t^H = \infty\right) = 1, \quad i = 0, 1. \tag{4.33}$$

Then there exists a sequential plan $\tilde{\Delta} = \Delta(\tilde{\tau}, \tilde{\delta})$ in the class $\Delta_{\alpha,\beta}$ which is *optimal* in the sense that for any other sequential plan $\Delta = \Delta(\tau, \delta)$ in $\Delta_{\alpha,\beta}$,

$$E_i\left(\int_0^{\tilde{\tau}(\psi)} m_t^2(\psi)dw_t^H\right) \leq E_i\left(\int_0^{\tau(\psi)} m_t^2(\psi)dw_t^H\right), \; i = 0, 1. \tag{4.34}$$

The sequential plan $\tilde{\Delta} = \Delta(\tilde{\tau}, \tilde{\delta})$ is defined by the relations

$$\tilde{\tau}(\psi) = \inf\{t : \lambda_t(\psi) \geq B \;\text{ or }\; \lambda_t(\psi) \leq A\}$$

and

$$\begin{aligned}\tilde{\delta}(\psi) &= 1 \;\text{ if }\; \lambda_{\tilde{\tau}(\psi)} \geq B,\\ &= 0 \;\text{ if }\; \lambda_{\tilde{\tau}(\psi)} \leq A,\end{aligned}$$

where

$$\lambda_t(\psi) = \int_0^t m_s(\psi)dZ_s - \frac{1}{2}\int_0^t m_s^2(\psi)dw_s^H$$

and

$$A = \log \frac{\alpha}{1-\beta}, \quad B = \log \frac{1-\alpha}{\beta}.$$

Furthermore,

$$E_0 \left(\int_0^{\tilde{\tau}(\psi)} m_t^2(\psi) dw_t^H \right) = 2\ V(\beta, \alpha), \tag{4.35}$$

and

$$E_1 \left(\int_0^{\tilde{\tau}(\psi)} m_t^2(\psi) dw_t^H \right) = 2\ V(\alpha, \beta), \tag{4.36}$$

where

$$V(x, y) = (1-x) \log \frac{1-x}{y} + x \log \frac{x}{1-y}. \tag{4.37}$$

We first derive three lemmas which will be used to prove Theorem 4.3.

Lemma 4.4: The sequential plan $\tilde{\Delta} = \Delta(\tilde{\tau}, \tilde{\delta})$ satisfies the properties

$$P_i(\tilde{\tau}(\psi) < \infty) = 1, \quad i = 0, 1. \tag{4.38}$$

Proof: Note that

$$P_0(\tilde{\tau}(\psi) < \infty) = P(\tilde{\tau}(W^H) < \infty)$$

since $\psi_t = W_t^H$ under H_0. Let

$$\sigma_n(W^H) = \inf \left\{ t : \int_0^t m_s^2(W^H) dw_s^H \geq n \right\}.$$

Then

$$\lambda_{\tilde{\tau}(W^H) \wedge \sigma_n(W^H)}(W^H) = \int_0^{\tilde{\tau}(W^H) \wedge \sigma_n(W^H)} m_s(W^H) dM_s^H - \frac{1}{2} \int_0^{\tilde{\tau}(W^H) \wedge \sigma_n(W^H)} m_s^2(W^H) dw_s^H$$

and

$$A \leq \lambda_{\tilde{\tau}(W^H) \wedge \sigma_n(W^H)}(W^H) \leq B.$$

Hence

$$A \leq E(\lambda_{\tilde{\tau}(W^H) \wedge \sigma_n(W^H)}(W^H)) \leq B$$

which implies that

$$E\left(\int_0^{\tilde{\tau}(W^H)\wedge\sigma_n(W^H)} m_s^2(W^H)dw_s^H\right) \leq 2(B-A) < \infty$$

since $0 < \alpha + \beta < 1$. In particular, we have

$$E\left(\int_0^{\tilde{\tau}(W^H)} m_s^2(W^H)dw_s^H\right) \leq 2(B-A) < \infty. \tag{4.39}$$

Since

$$E\left(\int_0^{\tilde{\tau}(W^H)} m_s^2(W^H)dw_s^H\right) \geq E\left(I_{\{\tilde{\tau}(W^H)=\infty\}}\int_0^{\infty} m_s^2(W^H)dw_s^H\right),$$

it follows that $P(\tilde{\tau}(W^H) < \infty) = 1$ from Equation (4.33). Applying an analogous argument, we can prove that $P_1(\tilde{\tau}(\psi) < \infty) = 1$. This completes the proof.

Let

$$\nu_t = Z_t - \int_0^t m_s(\psi)dw_s^H. \tag{4.40}$$

Then

$$dZ_t = m_s(\psi)dw_s^H + d\nu_t, \quad t \geq 0 \tag{4.41}$$

where $\{\nu_t, \mathcal{F}_t^{\psi}, t \geq 0\}$ is a continuous Gaussian martingale with $\langle\nu\rangle_t = m_t^H$. Furthermore, under the hypothesis H_1,

$$\lambda_t(\psi) = \int_0^t m_s(\psi)d\nu_s + \frac{1}{2}\int_0^t m_s^2(\psi)dw_s^H. \tag{4.42}$$

This can be seen from Theorem 1.20 in Chapter 1 (cf. Kleptsyna *et al.* (2000a), Theorem 2).

Remarks: Observe that the random variable $\lambda_{\tilde{\tau}(\psi)}$ takes the values A and B only a.s. under the probability measures P_0 as well as P_1.

Lemma 4.5: The sequential plan $\tilde{\Delta} = \Delta(\tilde{\tau}, \tilde{\delta})$ defined in Theorem 4.3. has the property

$$\alpha(\tilde{\Delta}) = \alpha, \quad \beta(\tilde{\Delta}) = \beta.$$

Proof: Since

$$\alpha(\tilde{\Delta}) = P_1(\tilde{\delta}(\psi) = 0) = P_1(\lambda_{\tilde{\tau}(\psi)}(\psi) = A)$$

and

$$\beta(\tilde{\Delta}) = P_0(\tilde{\delta}(\psi) = 1) = P_1(\lambda_{\tilde{\tau}(\psi)}(\psi) = B),$$

it is sufficient to prove that

$$P_1(\lambda_{\tilde{\tau}(\psi)}(\psi) = A) = \alpha, \quad P_0(\lambda_{\tilde{\tau}(\psi)}(\psi) = A) = \beta. \tag{4.43}$$

Following the techniques in Liptser and Shiryayev (2001), p. 251, let $a(x)$ and $b(x)$, $A \leq x \leq B$, be the solutions of the differential equations

$$a''(x) + a'(x) = 0, \quad a(A) = 1, a(B) = 0 \tag{4.44}$$

and

$$b''(x) + b'(x) = 0, \quad b(A) = 0, b(B) = 1. \tag{4.45}$$

It can be checked that

$$a(x) = \frac{e^A(e^{B-x} - 1)}{e^B - e^A}, \quad b(x) = \frac{e^x - e^A}{e^B - e^A} \tag{4.46}$$

and

$$a(0) = \alpha, \quad b(0) = \beta. \tag{4.47}$$

We will first prove that

$$P_1(\lambda_{\tilde{\tau}(\psi)}(\psi) = A) = \alpha. \tag{4.48}$$

Let

$$\sigma_n(\psi) = \inf\left\{t : \int_0^t m_s^2(\psi) dw_s^H \geq n\right\}.$$

Applying the generalized Ito–Ventzell formula for continuous local martingales (cf. Prakasa Rao (1999b), p. 76), we obtain that

$$\begin{aligned} a(\lambda_{\tilde{\tau}(\psi)\wedge\sigma_n(\psi)}(\psi)) &= a(0) + \int_0^{\tilde{\tau}(\psi)\wedge\sigma_n(\psi)} a'(\lambda_s(\psi)) m_s(\psi) d\nu_s \\ &\quad + \frac{1}{2}\int_0^{\tilde{\tau}(\psi)\wedge\sigma_n(\psi)} [a'(\lambda_s(\psi)) + a''(\lambda_s(\psi))] m_s^2(\psi) dw_s^H \\ &= \alpha + \int_0^{\tilde{\tau}(\psi)\wedge\sigma_n(\psi)} a'(\lambda_s(\psi)) m_s(\psi) d\nu_s. \end{aligned} \tag{4.49}$$

But

$$E_1 \int_0^{\tilde{\tau}(\psi)\wedge\sigma_n(\psi)} [a'(\lambda_s(\psi))m_s(\psi)]^2 dw_s^H \leq$$
$$\sup_{A\leq x\leq B} [a'(x)]^2 \; E_1\left(\int_0^{\tilde{\tau}(\psi)\wedge\sigma_n(\psi)} m_s^2(\psi)dw_s^H\right) \leq n \sup_{A\leq x\leq B} [a'(x)]^2 < \infty.$$

Hence

$$E_1\left(\int_0^{\tilde{\tau}(\psi)\wedge\sigma_n(\psi)} a'(\lambda_s(\psi))m_s(\psi)d\nu_s\right) = 0.$$

Taking the expectation under the probability measure P_1 on both sides of (4.49), we get that

$$E_1(a(\lambda_{\tilde{\tau}(\psi)\wedge\sigma_n(\psi)}(\psi))) = \alpha$$

Observe that the function $a(x)$ is bounded on the interval $[A, B]$ and $\sigma_n(\psi) \to \infty$ a.s. under P_1 as $n \to \infty$. An application of the dominated convergence theorem proves that

$$E_1[a(\lambda_{\tilde{\tau}(\psi)\wedge\sigma_n(\psi)}(\psi))] = \alpha. \tag{4.50}$$

Applying Lemma 4.4, noting that $\lambda_{\tilde{\tau}(\psi)}$ takes only the values A and B a.s. under the probability measure P_1, and observing that $a(A) = 1$ and $a(B) = 0$, we find that

$$\begin{aligned}\alpha &= E_1[a(\lambda_{\tilde{\tau}(\psi)})]\\ &= 1.P_1(\lambda_{\tilde{\tau}(\psi)} = A) + 0.P_1(\lambda_{\tilde{\tau}(\psi)} = B)\\ &= P_1(\lambda_{\tilde{\tau}(\psi)} = A).\end{aligned} \tag{4.51}$$

Similar arguments show that

$$P_0(\lambda_{\tilde{\tau}(\psi)} = B) = \beta. \tag{4.52}$$

Lemma 4.6: The relations (4.35) and (4.36) hold for the sequential plan $\tilde{\Delta} = \Delta(\tilde{\tau}, \tilde{\delta})$.

Proof: Proof of this lemma is analogous to the proof of Lemma 17.9 in Liptser and Shiryayev (2001) and the result can be proved as an application of the generalized Ito–Ventzell formula for continuous local martingales (cf. Prakasa Rao (1999b)). Let $g_i(x)$, $A \leq x \leq B$, $i = 0, 1$, be the solutions of the differential equations

$$g_i''(x) + (-1)^{i+1} g_i'(x) = -2, \quad g_i(A) = g_i(B) = 0, i = 0, 1.$$

It can be checked that

$$g_0(x) = 2\left[\frac{(e^B - e^{A+B-x})(B-A)}{e^B - e^A} + A - x\right],$$

$$g_1(x) = 2\left[\frac{(e^B - e^x)(B-A)}{e^B - e^A} - B + x\right]$$

and

$$g_0(0) = -2\ V(\beta, \alpha),\ g_1(0) = 2\ V(\alpha, \beta).$$

Suppose the hypothesis H_0 holds. Define

$$\sigma_n(W^H) = \inf\left\{t : \int_0^t m_s^2(W^H) dw_s^H \geq n\right\}, \quad n \geq 1.$$

Applying the generalized Ito–Ventzell formula to $g_0(\lambda_t(W^H))$, we obtain that

$$\begin{aligned}
&g_0(\lambda_{\tilde{\tau}(W^H)\wedge\sigma_n(W^H)}(W^H)) \\
&\quad = g_0(0) + \int_0^{\tilde{\tau}(W^H)\wedge\sigma_n(W^H)} g_0'(\lambda_s(W^H)) m_s(W^H) dM_s^H \\
&\qquad - \frac{1}{2}\int_0^{\tilde{\tau}(W^H)\wedge\sigma_n(W^H)} [g_0'(\lambda_s(W_s^H)) \quad - g_0''(\lambda_s(W_s^H))] m_s^2(W_s^H) dw_s^H \\
&\quad = g_0(0) + \int_0^{\tilde{\tau}(W^H)\wedge\sigma_n(W^H)} g_0'(\lambda_s(W_s^H)) m_s(W_s^H) dM_s^H \\
&\qquad + \int_0^{\tilde{\tau}(W^H)\wedge\sigma_n(W^H)} m_s^2(W_s^H) dw_s^H.
\end{aligned} \tag{4.53}$$

Since

$$E_0\left(\int_0^{\tilde{\tau}(W^H)\wedge\sigma_n(W^H)} g_0'(\lambda_s(W_s^H)) m_s(W_s^H) dM_s^H\right) = 0,$$

taking expectations with respect to the probability measure P_0 on both sides of Equation (4.53), we have

$$E_0\left(\int_0^{\tilde{\tau}(W^H)\wedge\sigma_n(W^H)} m_s^2(W_s^H) dw_s^H\right) = -g_0(0) + E_0(g_0(\lambda_{\tilde{\tau}(W^H)\wedge\sigma_n(W^H)}(W^H))).$$

Taking the limit as $n \to \infty$, we obtain that

$$E_0\left(\int_0^{\tilde{\tau}(\psi)} m_t^2(\psi) dw_t^H\right) = -g_0(0) = 2V(\beta, \alpha). \tag{4.54}$$

Similarly we can prove that

$$E_1\left(\int_0^{\tilde{\tau}(\psi)} m_t^2(\psi)dw_t^H\right) = -g_1(0) = 2V(\alpha, \beta). \tag{4.55}$$

This completes the proof of Lemma 4.6.

We now prove Theorem 4.3.

Proof of Theorem 4.3: Let $\Delta = \Delta(\tau, \delta)$ be any sequential plan in the class $\Delta_{\alpha,\beta}$. Let P_i^τ be the restriction of the probability measure P_i restricted to the σ-algebra $\mathcal{B}_\tau$ for $i = 0, 1$. In view of conditions (4.30), (4.31), (4.33) and the representation (4.42), it follows that the probability measures $P_i^\tau, i = 0, 1$, are equivalent by Theorem 7.10 in Liptser and Shiryayev (2001). Furthermore,

$$\log \frac{dP_1^\tau}{dP_0^\tau}(\tau, W^H) = \int_0^{\tau(W^H)} m_s(W^H)dM_s^H - \frac{1}{2}\int_0^{\tau(W^H)} m_s^2(W^H)dw_s^H$$

and

$$\log \frac{dP_1^\tau}{dP_0^\tau}(\tau, \psi) = \int_0^{\tau(\psi)} m_s(\psi)dZ_s - \frac{1}{2}\int_0^{\tau(\psi)} m_s^2(\psi)dw_s^H.$$

Therefore

$$\begin{aligned} E_0\left(\log \frac{dP_0^\tau}{dP_1^\tau}(\tau, \psi)\right) &= \frac{1}{2}E_0\left(\int_0^{\tau(\psi)} m_s^2(\psi)dw_s^H\right) \\ &= \frac{1}{2}E_0\left(\int_0^{\tau(W^H)} m_s^2(W^H)dw_s^H\right) \end{aligned} \tag{4.56}$$

and

$$E_1\left(\log \frac{dP_1^\tau}{dP_0^\tau}(\tau, \psi)\right) = \frac{1}{2}E_1\left(\int_0^{\tau(\psi)} m_s^2(\psi)dw_s^H\right). \tag{4.57}$$

Applying Jensen's inequality and following the arguments similar to those in Liptser and Shiryayev (2001), pp. 254–255, it can be shown that

$$\begin{aligned} \frac{1}{2}E_1\left(\int_0^{\tau(\psi)} m_s^2(\psi)dw_s^H\right) &\geq (1-\alpha)\log\frac{1-\alpha}{\beta} + \alpha\log\frac{\alpha}{1-\beta} \\ &= \frac{1}{2}E_1\left(\int_0^{\tilde{\tau}(\psi)} m_s^2(\psi)dw_s^H\right) \end{aligned} \tag{4.58}$$

by using Lemma 4.6. Hence

$$E_1\left(\int_0^{\tilde{\tau}(\psi)} m_s^2(\psi)dw_s^H\right) \leq E_1\left(\int_0^{\tau(\psi)} m_s^2(\psi)dw_s^H\right). \tag{4.59}$$

Similarly we can prove that

$$E_0\left(\int_0^{\tilde{\tau}(\psi)} m_s^2(\psi)dw_s^H\right) \leq E_0\left(\int_0^{\tau(\psi)} m_s^2(\psi)dw_s^H\right). \tag{4.60}$$

This completes the proof of Theorem 4.3

Remarks: As a special case of the above result, suppose that $\theta_t = h(t)$ where $h(t)$ is a non-random but differentiable function such that

$$\int_0^\infty h^2(t)dt = \infty, \quad h(t)h'(t) \geq 0, t \geq 0. \tag{4.61}$$

Let α, β be given such that $0 < \alpha + \beta < 1$.

Let $\Delta_{\alpha,\beta}$ be the class of sequential plans as discussed earlier for given α, β with $0 < \alpha + \beta < 1$. Consider the plan $\Delta_T = (T, \delta_T)$ having the fixed observation time T for $0 < T < \infty$ and belonging to the class $\Delta_{\alpha,\beta}$. Then the optimal sequential plan $\tilde{\Delta} = (\tilde{\tau}, \tilde{\delta}) \in \Delta_{\alpha,\beta}$ has the properties

$$E_i(\tilde{\tau}) \leq T, \quad i = 0, 1. \tag{4.62}$$

This can be seen by checking that, for $i = 0, 1$,

$$\begin{aligned} E_i\left(\int_0^{\tilde{\tau}(\psi)} h^2(t)dt\right) &\leq E_i\left(\int_0^T h^2(t)dt\right) \\ &= \int_0^T h^2(t)dt = \Phi(T) \text{ (say)} \end{aligned} \tag{4.63}$$

which in turn implies that

$$\begin{aligned} \Phi(T) &\geq E_i\left(\int_0^{\tilde{\tau}(\psi)} h^2(t)dt\right) \\ &= E_i(\Phi(\tilde{\tau}(\psi))) \\ &\geq \Phi(E_i(\tilde{\tau}(\psi))) \end{aligned} \tag{4.64}$$

by observing that the function $\Phi(.)$ is convex and by applying Jensen's inequality. The inequality derived above in turn proves that

$$E_i(\tilde{\tau}(\psi)) \leq T, \; i = 0, 1. \tag{4.65}$$

5

Nonparametric inference for processes driven by fBm

5.1 Introduction

We have discussed parametric inference for a class of processes driven by fBm when the trend parameter θ is finite dimensional. We now consider the problem of estimation when it is infinite dimensional, in particular when the parameter θ is an unknown function. We now discuss the problem of nonparametric estimation or identification of the 'drift' or 'trend' function $\theta(.)$ for a class of stochastic processes satisfying a SDE

$$dX_t = \theta(t)X_t dt + dW_t^H,\ X_0 = \tau, t \geq 0, \tag{5.1}$$

where τ is a Gaussian random variable and $\{W_t^H\}$ is fBm, and the problem of estimation of the 'drift' or 'trend' function $S(.)$ for SDEs of the type

$$dX_t = S(X_t)dt + \epsilon\, dW_t^H, X_0 = x_0,\ 0 \leq t \leq T \tag{5.2}$$

5.2 Identification for linear stochastic systems

Consider the SDE

$$dX_t = \theta(t)X_t dt + dW_t^H,\ X_0 = \tau, t \geq 0 \tag{5.3}$$

where τ is a Gaussian random variable and $\{W_t^H\}$ is fBm. We use the method of sieves and study the asymptotic properties of the estimator. Identification of non-stationary diffusion models by the method of sieves is studied in Nguyen and Pham (1982). The results discussed here are from Prakasa Rao (2004c).

Statistical Inference for Fractional Diffusion Processes B. L. S. Prakasa Rao

Estimation by the method of sieves

Let us consider the linear stochastic system

$$dX(t) = \theta(t)X(t)dt + dW_t^H,\ X(0) = \tau,\ \ 0 \le t \le T \tag{5.4}$$

where $\theta(t) \in L^2([0,T], dt)$, $W = \{W_t^H, t \ge 0\}$ is fBm with Hurst parameter H and τ is a Gaussian random variable independent of fBm W. In other words, $X = \{X_t, t \ge 0\}$ is a stochastic process satisfying the stochastic integral equation

$$X(t) = \tau + \int_0^t \theta(s)X(s)ds + W_t^H,\ 0 \le t \le T. \tag{5.5}$$

Let

$$C_\theta(t) = \theta(t)\ X(t),\ 0 \le t \le T \tag{5.6}$$

and assume that the sample paths of the process $\{C_\theta(t), 0 \le t \le T\}$ are smooth enough so that the process

$$Q_{H,\theta}(t) = \frac{d}{dw_t^H}\int_0^t k_H(t,s)C_\theta(s)ds, 0 \le t \le T \tag{5.7}$$

is well defined where w_t^H and $k_H(t,s)$ are as defined in (1.72) and (1.70) in Chapter 1 respectively. Suppose the sample paths of the process $\{Q_{H,\theta}(t), 0 \le t \le T\}$ belong a.s. to $L^2([0,T], dw_t^H)$. Define

$$Z_t = \int_0^t k_H(t,s)dX_s,\ 0 \le t \le T. \tag{5.8}$$

Then the process $Z = \{Z_t, 0 \le t \le T\}$ is an $(\mathcal{F}_t)$-semimartingale with the decomposition

$$Z_t = \int_0^t Q_{H,\theta}(s)dw_s^H + M_t^H \tag{5.9}$$

where M^H is the fundamental martingale defined by (1.73) in Chapter 1 and the process X admits the representation

$$X_t = X_0 + \int_0^t K_H(t,s)dZ_s \tag{5.10}$$

where the function K_H is as defined by (1.75) in Chapter 1 with $f \equiv 1$. Let P_θ^T be the measure induced by the process $\{X_t, 0 \le t \le T\}$ when $\theta(.)$ is the true 'drift' function. Following Theorem 1.20 in Chapter 1, we get that the Radon–Nikodym derivative of P_θ^T with respect to P_0^T is given by

$$\frac{dP_\theta^T}{dP_0^T} = \exp\left[\int_0^T Q_{H,\theta}(s)dZ_s - \frac{1}{2}\int_0^T Q_{H,\theta}^2(s)dw_s^H\right]. \tag{5.11}$$

Suppose the process X is observable on $[0, T]$ and $X_i, 1 \le i \le n$, is a random sample of n independent observations of the process X on $[0, T]$. Following the representation of the Radon–Nikodym derivative of P_θ^T with respect to P_0^T given above, it follows that the log-likelihood function corresponding to the observations $\{X_i, 1 \le i \le n\}$ is given by

$$L_n(X_1, \ldots, X_n; \theta) \equiv L_n(\theta) = \sum_{i=1}^{n} \left(\int_0^T Q_{H,\theta}^{(i)}(s) dZ_i(s) - \frac{1}{2} \int_0^T [Q_{H,\theta}^{(i)}]^2(s) dw_s^H \right) \tag{5.12}$$

where the process $Q_{H,\theta}^{(i)}$ is as defined by the relation (5.7) for the process X_i. For convenience in notation, we write $Q_{i,\theta}(s)$ hereafter for $Q_{H,\theta}^{(i)}(s)$.

Let $\{V_n, n \ge 1\}$ be an increasing sequence of subspaces of finite dimensions $\{d_n\}$ such that $\cup_{n \ge 1} V_n$ is dense in $L^2([0, T], dt)$. The method of sieves consists of maximizing $L_n(\theta)$ on the subspace V_n. Let $\{e_i\}$ be a set of linearly independent vectors in $L^2([0, T], dt)$ such that the set of vectors $\{e_1, \ldots, e_{d_n}\}$ is a basis for the subspace V_n for every $n \ge 1$. For $\theta \in V_n$, $\theta(.) = \sum_{j=1}^{d_n} \theta_j e_j(.)$, we have

$$\begin{aligned} Q_{i,\theta}(t) &= \frac{d}{dw_t^H} \int_0^t k_H(t, s) \theta(s) X_i(s) ds \\ &= \frac{d}{dw_t^H} \int_0^t k_H(t, s) \left[\sum_{j=1}^{d_n} \theta_j e_j(s) \right] X_i(s) ds \\ &= \sum_{j=1}^{d_n} \theta_j \frac{d}{dw_t^H} \int_0^t k_H(t, s) e_j(s) X_i(s) ds \\ &= \sum_{j=1}^{d_n} \theta_j \Gamma_{i,j}(t) (\text{say}). \end{aligned} \tag{5.13}$$

Furthermore,

$$\begin{aligned} \int_0^T Q_{i,\theta}(t) dZ_i(t) &= \int_0^T \left[\sum_{j=1}^{d_n} \theta_j \Gamma_{i,j}(t) \right] dZ_i(t) \\ &= \sum_{j=1}^{d_n} \theta_j \int_0^T \Gamma_{i,j}(t) dZ_i(t) \\ &= \sum_{j=1}^{d_n} \theta_j R_{i,j} \quad (\text{say}) \end{aligned} \tag{5.14}$$

and

$$\begin{aligned}\int_0^T Q_{i,\theta}^2(t)dw_t^H &= \int_0^T \left[\sum_{j=1}^{d_n} \theta_j \Gamma_{i,j}(t)\right]^2 dw_t^H \\ &= \sum_{j=1}^{d_n}\sum_{k=1}^{d_n} \theta_j\theta_k \int_0^T \Gamma_{i,j}(t)\Gamma_{i,k}(t)dw_t^H \\ &= \sum_{j=1}^{d_n}\sum_{k=1}^{d_n} \theta_j\theta_k \langle R_{i,j}, R_{i,k}\rangle \end{aligned}\tag{5.15}$$

where $\langle ., .\rangle$ denotes the quadratic covariation. Therefore the log-likelihood function corresponding to the observations $\{X_i, 1 \le i \le n\}$ is given by

$$\begin{aligned}L_n(\theta) &= \sum_{i=1}^n \left(\int_0^T Q_{i,\theta}(t)dZ_i(t) - \frac{1}{2}\int_0^T Q_{i,\theta}^2(t)dw_t^H\right) \\ &= \sum_{i=1}^n \left[\sum_{j=1}^{d_n} \theta_j R_{i,j} - \frac{1}{2}\sum_{j=1}^{d_n}\sum_{k=1}^{d_n} \theta_j\theta_k\langle R_{i,j}, R_{i,k}\rangle\right] \\ &= n\left[\sum_{j=1}^{d_n} \theta_j B_j^{(n)} - \frac{1}{2}\sum_{j=1}^{d_n}\sum_{k=1}^{d_n}\theta_j\theta_k A_{j,k}^{(n)}\right]\end{aligned}\tag{5.16}$$

where

$$B_j^{(n)} = n^{-1}\sum_{i=1}^n R_{i,j}, \quad 1 \le j \le d_n \tag{5.17}$$

and

$$A_{j,k}^{(n)} = n^{-1}\sum_{i=1}^n \langle R_{i,j}, R_{i,k}\rangle, \quad 1 \le j, k \le d_n. \tag{5.18}$$

Let $\theta^{(n)}$, $B^{(n)}$ and $A^{(n)}$ be the vectors and the matrix with elements θ_j, $j = 1, \ldots, d_n$, $B_j^{(n)}$, $j = 1, \ldots, d_n$, and $A_{j,k}^{(n)}$, $j, k = 1, \ldots, d_n$, as defined above. Then the log-likelihood function can be written in the form

$$L_n(\theta) = n\left[B^{(n)}\theta^{(n)} - \frac{1}{2}\theta^{(n)\prime}A^{(n)}\theta^{(n)}\right]. \tag{5.19}$$

Here α' denotes the transpose of the vector α. The restricted MLE $\hat{\theta}^{(n)}(.)$ of $\theta(.)$ is given by

$$\hat{\theta}^{(n)}(.) = \sum_{j=1}^{d_n} \hat{\theta}_j^{(n)} e_j(.) \tag{5.20}$$

where

$$\hat{\theta}^{(n)} = (\hat{\theta}_1^{(n)}, \ldots, \hat{\theta}_{d_n}^{(n)}) \tag{5.21}$$

is the solution of the equation

$$A^{(n)}\hat{\theta}^{(n)} = B^{(n)}. \tag{5.22}$$

Assuming that $A^{(n)}$ is invertible, we get that

$$\hat{\theta}^{(n)} = (A^{(n)})^{-1} B^{(n)}. \tag{5.23}$$

We now construct an orthonormal basis for V_n with respect to a suitable inner product so that the matrix $A^{(n)}$ is transformed into an identity matrix as $n \to \infty$. Note that

$$\begin{aligned} A_{j,k}^{(n)} \to \int_0^T E\Bigg[\left(\frac{d}{dw_t^H}\int_0^t k_H(t,s)e_j(s)X(s)ds\right) \\ \times \left(\frac{d}{dw_t^H}\int_0^t k_H(t,s)e_k(s)X(s)ds\right)\Bigg] dw_t^H \end{aligned} \tag{5.24}$$

a.s. as $n \to \infty$ by the strong law of large numbers. We now consider a sequence $\psi_j, j \geq 1$, such that $\psi_j, 1 \leq j \leq d_n$, is an orthonormal basis of V_n in the sense of the inner product

$$\begin{aligned} \langle h, g\rangle = \int_0^T E\Bigg[\left(\frac{d}{dw_t^H}\int_0^t k_H(t,s)h(s)X(s)ds\right) \\ \times \left(\frac{d}{dw_t^H}\int_0^t k_H(t,s)g(s)X(s)ds\right)\Bigg] dw_t^H. \end{aligned} \tag{5.25}$$

Let $\hat{\eta}_1^{(n)}, \hat{\eta}_2^{(n)}, \ldots, \hat{\eta}_{d_n}^{(n)}$ be the coordinates of $\hat{\theta}^{(n)}(.)$ in the new basis $\psi_j, 1 \leq j \leq d_n$. Then the vector

$$\hat{\eta}^{(n)} = (\hat{\eta}_1^{(n)}, \hat{\eta}_2^{(n)}, \ldots, \hat{\eta}_{d_n}^{(n)}) \tag{5.26}$$

is the solution of the equation

$$a^{(n)}\hat{\eta}^{(n)} = b^{(n)} \tag{5.27}$$

where $a^{(n)}$ and $b^{(n)}$ are the matrix and the vector with general elements

$$\begin{aligned} a_{j,k}^{(n)} = n^{-1}\sum_{i=1}^n \int_0^T \Bigg\{\frac{d}{dw_t^H}\left[\int_0^t k_H(t,s)\psi_j(s)X_i(s)ds\right] \\ \times \frac{d}{dw_t^H}\left[\int_0^t k_H(t,s)\psi_k(s)X_i(s)ds)\right]\Bigg\} dw_t^H, \end{aligned} \tag{5.28}$$

and

$$b_j^{(n)} = n^{-1} \sum_{i=1}^{n} \int_0^T \frac{d}{dw_t^H} \left[\int_0^t k_H(t,s)\psi_j(s)X_i(s)ds \right] dZ_i(t). \tag{5.29}$$

Let $\theta^{(n)}(.) = \sum_{k=1}^{d_n} \eta_i \psi_i(.)$ be the orthogonal projection of $\theta(.)$ onto V_n in the sense of the inner product $\langle ., . \rangle$ defined above. Observe that

$$\begin{aligned}
& b_j^{(n)} - \sum_{k=1}^{d_n} a_{j,k}^{(n)} \eta_k \\
&= n^{-1} \sum_{i=1}^{n} \int_0^T Q_{i,\psi_j}(t) dZ_i(t) - \sum_{k=1}^{d_n} a_{j,k}^{(n)} \eta_k \\
&= n^{-1} \sum_{i=1}^{n} \int_0^T Q_{i,\psi_j}(t) [Q_{i,\theta}(t) dw_t^H + dM_{it}^H] - \sum_{k=1}^{d_n} a_{j,k}^{(n)} \eta_k \\
&= n^{-1} \sum_{i=1}^{n} \int_0^T Q_{i,\psi_j}(t) Q_{i,\theta}(t) dw_t^H \\
&\quad + n^{-1} \sum_{i=1}^{n} \int_0^T Q_{i,\psi_j}(t) dM_{it}^H - \sum_{k=1}^{d_n} a_{j,k}^{(n)} \eta_k \\
&= n^{-1} \sum_{i=1}^{n} \int_0^T Q_{i,\psi_j}(t) \left(\sum_{r=1}^{\infty} \eta_r Q_{i,\psi_r}(t) \right) dw_t^H \\
&\quad + n^{-1} \sum_{i=1}^{n} \int_0^T Q_{i,\psi_j}(t) dM_{it}^H - \sum_{k=1}^{d_n} a_{j,k}^{(n)} \eta_k \\
&= n^{-1} \sum_{i=1}^{n} \int_0^T Q_{i,\psi_j}(t) \left(\sum_{r=1}^{d_n} \eta_r Q_{i,\psi_r}(t) + \sum_{r=d_n}^{\infty} \eta_r Q_{i,\psi_r}(t) \right) dw_t^H \\
&\quad + n^{-1} \sum_{i=1}^{n} \int_0^T Q_{i,\psi_j}(t) dM_{it}^H - n^{-1} \sum_{k=1}^{d_n} \eta_k \int_0^T Q_{i,\psi_j}(t) Q_{i,\psi_k}(t) dw_t^H \\
&= n^{-1} \sum_{i=1}^{n} \int_0^T Q_{i,\psi_j}(t) Q_{i,\theta-\theta^{(n)}}(t) dw_t^H + n^{-1} \sum_{i=1}^{n} \int_0^T Q_{i,\psi_j}(t) dM_{it}^H
\end{aligned}$$

$$= n^{-1}\sum_{i=1}^{n}\int_0^T [Q_{i,\psi_j}(t)Q_{i,\theta-\theta^{(n)}}(t) - E(Q_{i,\psi_j}(t)Q_{i,\theta-\theta^{(n)}}(t))]dw_t^H$$

$$+n^{-1}\sum_{i=1}^{n}\int_0^T Q_{i,\psi_j}(t)dM_{it}^H \tag{5.30}$$

since

$$\langle\theta - \theta^{(n)}, \psi_j\rangle = 0 \tag{5.31}$$

for $1 \le j \le d_n$ by the orthogonality of the basis $\{\psi_k, k \ge 1\}$ and the fact that

$$\langle\theta - \theta^{(n)}, \psi_j\rangle = E\left[\int_0^T Q_{i,\psi_j}(t)Q_{i,\theta-\theta^{(n)}}(t)dw_t^H\right]. \tag{5.32}$$

Hence

$$a^{(n)}(\hat{\eta}^{(n)} - \eta^{(n)}) = c^{(n)} \tag{5.33}$$

where $\eta^{(n)}$ and $c^{(n)}$ are vectors with components $\eta_j, 1 \le j \le d_n$, and

$$c_j^{(n)} = n^{-1}\sum_{i=1}^{n}\int_0^T [Q_{i,\psi_j}(t)Q_{i,\theta-\theta^{(n)}}(t) - E(Q_{i,\psi_j}(t)Q_{i,\theta-\theta^{(n)}}(t))]\, dw_t^H$$

$$+\, n^{-1}\sum_{i=1}^{n}\int_0^T Q_{i,\psi_j}(t)dM_{it}^H. \tag{5.34}$$

Let $\delta_{jk} = 0$ if $j \ne k$ and $\delta_{jk} = 1$ if $j = k$. In view of the orthonormality of the basis $\{\psi_j, j \ge 1\}$, it follows that

$$a_{j,k}^{(n)} - \delta_{j,k} = n^{-1}\sum_{i=1}^{n}\int_0^T (Q_{i,\psi_j}(t)Q_{i,\psi_k}(t) - E(Q_{i,\psi_j}(t)Q_{i,\psi_k}(t)))dw_t^H$$

$$= n^{-1}\zeta_{ijk} \ \ \text{(say)} \tag{5.35}$$

and

$$c_j^{(n)} = n^{-1}\sum_{i=1}^{n}\int_0^T [Q_{i,\psi_j}(t)Q_{i,\theta-\theta^{(n)}}(t) - E(Q_{i,\psi_j}(t)Q_{i,\theta-\theta^{(n)}}(t))]dw_t^H$$

$$+n^{-1}\sum_{i=1}^{n}\int_0^T Q_{i,\psi_j}(t)dM_{it}^H$$

$$= n^{-1}\sum_{i=1}^{n}\zeta_{ij}^{(n)} + n^{-1}\sum_{i=1}^{n}\tilde{\zeta}_{ij} \ \ \text{(say)}. \tag{5.36}$$

Note that $E[a_{j,k}^{(n)}] = \delta_{jk}$ and $E(\zeta_{ijk}) = 0$. Hence

$$\begin{aligned}
&E[a_{j,k}^{(n)} - \delta_{jk}]^2 \\
&\quad = Var(a_{j,k}^{(n)}) \\
&\quad = n^{-1} Var(\zeta_{1jk}) \ \text{(since } X_i, 1 \le i \le n, \text{ are i.i.d.)} \\
&\quad = n^{-1} E(\zeta_{1jk}^2) \\
&\quad = n^{-1} E\left[\int_0^T (Q_{i,\psi_j}(t) Q_{i,\psi_k}(t) - E[Q_{i,\psi_j}(t) Q_{i,\psi_k}(t)]) dw_t^H\right]^2 \\
&\quad \le n^{-1} E\left[\int_0^T (Q_{i,\psi_j}(t) Q_{i,\psi_k}(t) - E[Q_{i,\psi_j}(t) Q_{i,\psi_k}(t)])^2 dw_t^H \ \ w_T^H\right] \\
&\qquad\qquad \text{(by the Cauchy–Schwartz inequality)} \\
&\quad = n^{-1} \left\{\int_0^T E[Q_{i,\psi_j}(t) Q_{i,\psi_k}(t) - E[Q_{i,\psi_j}(t) Q_{i,\psi_k}(t)]^2] dw_t^H\right\} \ \ w_T^H \\
&\quad \le n^{-1} w_T^H \int_0^T E[Q_{i,\psi_j}(t) Q_{i,\psi_k}(t)]^2 dw_t^H. \qquad (5.37)
\end{aligned}$$

Note that the process $\{Q_{H,\theta}(t), t \ge 0\}$ defined by Equation (5.7) is a Gaussian process and the fundamental martingale M^H is a Gaussian martingale. We now state a lemma to get an upper bound for the expression on the right of Equation (5.37). The proof of this lemma is easy.

Lemma 5.1: Let $X_i, i = 1, 2$, be Gaussian random variables. Then there exists a positive constant C such that

$$E[X_1^2 X_2^2] \le C E(X_1^2) E(X_2^2). \qquad (5.38)$$

Applying Lemma 5.1 on the right of Equation (5.37), we get

$$\begin{aligned}
E[a_{j,k}^{(n)} - \delta_{jk}]^2 &\le n^{-1} w_T^H \int_0^T E[Q_{i,\psi_j}(t) Q_{i,\psi_k}(t)]^2 dw_t^H \\
&\le C n^{-1} w_T^H \int_0^T E[Q_{i,\psi_j}(t)^2] E[Q_{i,\psi_k}(t)]^2 dw_t^H \\
&= C n^{-1} w_T^H \sup_{0 \le t \le T} E[Q_{i,\psi_j}(t)^2] \int_0^T E[Q_{i,\psi_k}(t)]^2 dw_t^H \\
&= C n^{-1} w_T^H \sup_{0 \le t \le T} E[Q_{i,\psi_j}(t)^2] \qquad (5.39)
\end{aligned}$$

since $\int_0^T E(Q_{i,\psi_k}(t))^2 dw_t^H = 1$ by the choice of the orthonormal basis $\psi_j, j \ge 1$.

Observe that $E(\tilde{\zeta}_{ij}) = 0$ and $E(\zeta_{ij}^{(n)}) = 0$. Furthermore,

$$\begin{aligned} E(\tilde{\zeta}_{ij}^2) &= E\left[\int_0^T Q_{i,\psi_j}(t) dM_{it}^H\right]^2 \\ &= \int_0^T E[Q_{i,\psi_j}^2(t)] dw_t^H \\ &= 1 \end{aligned} \tag{5.40}$$

and it follows by the arguments given earlier and Lemma 5.1 that

$$E((\zeta_{ij}^{(n)})^2) \leq C w_T^H \sup_{0 \leq t \leq T} E[Q_{i,\psi_j}(t)^2] ||\theta - \theta^{(n)}||^2. \tag{5.41}$$

We will now estimate $E(c_j^{(n)})^2$. Note that $E(c_j^{(n)}) = 0$. Hence

$$\begin{aligned} E(c_j^{(n)})^2 &= Var(c_j^{(n)}) \\ &= n^{-1} Var(\zeta_{1j}^{(n)} + \tilde{\zeta}_{1j}) \\ &\leq n^{-1} E(\zeta_{1j}^{(n)} + \tilde{\zeta}_{1j})^2 \\ &\leq 2n^{-1}[E(\zeta_{1j}^{(n)})^2 + E(\tilde{\zeta}_{1j})^2] \\ &\leq 2n^{-1}[1 + C w_T^H \sup_{0 \leq t \leq T} E[Q_{1,\psi_j}(t)^2] ||\theta - \theta^{(n)}||^2]. \end{aligned} \tag{5.42}$$

Lemma 5.2: Let $||A|| = \sup\{||Ax||, ||x|| \leq 1\}$ be the operator norm of a matrix $A = ((a_{jk}))$. Then $||A||^2 \leq \sum a_{jk}^2$ and

$$||A^{-1}|| \leq \left\{1 + \left[\sum_{j,k} (a_{jk} - \delta_{jk})^2\right]^{-1/2}\right\}^{-1} \tag{5.43}$$

provided that

$$\sum_{j,k} (a_{jk} - \delta_{jk})^2 < 1.$$

Proof: See Lemma 3 of Nguyen and Pham (1982).

We now have the following result.

Theorem 5.3: Suppose V_n is an increasing sequence of subspaces of $L^2([0, T], dt)$ of dimension d_n defined by the inner product given by (5.25) such that $d_n \to \infty$ and let $\{\psi_j, 1 \le j \le d_n\}$ be an orthonormal basis of V_n. Suppose that

$$\frac{d_n^2 \gamma_n^2}{n} \to 0$$

and

$$\frac{\gamma_n d_n}{n} \to 0$$

as $n \to \infty$ where

$$\gamma_n = \sup_{0 \le t \le T} \sup_{1 \le j \le d_n} E\left[\frac{d}{dw_t^H} \int_0^t k_H(t,s)\psi_j(s)X(s)ds\right]^2. \tag{5.44}$$

Then

$$||\hat{\eta}^{(n)} - \eta^{(n)}|| \to 0 \tag{5.45}$$

in probability as $n \to \infty$.

Proof: Observe that

$$\hat{\eta}^{(n)} - \eta^{(n)} = a^{(n)^{-1}} c^{(n)} \tag{5.46}$$

from Equation (5.33). Applying Lemma 5.2, we get

$$||\hat{\eta}^{(n)} - \eta^{(n)}|| \le \left\{1 - \left[\sum_{j=1}^{d_n} \sum_{k=1}^{d_n} (a_{j,k}^{(n)} - \delta_{jk})^2\right]^{1/2}\right\}^{-1} ||c^{(n)}||. \tag{5.47}$$

Applying the bounds obtained in (5.41) and (5.42), we get

$$E\left[\sum_{j=1}^{d_n} \sum_{k=1}^{d_n} (a_{j,k}^{(n)} - \delta_{jk})^2\right] \le Cn^{-1} d_n^2 \gamma_n^2 \tag{5.48}$$

and the last term tends to zero as $n \to \infty$. Similarly

$$E||c^{(n)}||^2 \le C\gamma_n[n^{-1}d_n + n^{-1}d_n\gamma_n||\theta - \theta^{(n)}||^2] \tag{5.49}$$

and the last term tends to zero as $n \to \infty$. Hence

$$||\hat{\eta}^{(n)} - \eta^{(n)}|| \to 0 \tag{5.50}$$

in probability as $n \to \infty$.

Corollary 5.4: Under the conditions stated in Theorem 5.3,

$$\lim_{n\to\infty} \frac{d}{dw_t^H} \int_0^t k_H(t,s)(\hat{\theta}^{(n)}(s) - \theta^{(n)}(s))X(s)ds = 0$$

in probability.

Proof: Observe that

$$||\hat{\theta}^{(n)} - \theta^{(n)}||^2 = \int_0^T E\left[\frac{d}{dw_t^H}\int_0^t k_H(t,s)(\hat{\theta}^{(n)}(s) - \theta^{(n)}(s))X(s)ds\right]^2 dw_t^H$$

which can also be written in the form

$$\sum_{j=1}^{d_n} |\hat{\eta}_j^{(n)} - \eta_j|^2 + \sum_{j=d_n+1}^{\infty} \eta_j^2.$$

The first term in the above sum tends to zero by Theorem 5.3. Since the set $\cup_{n\geq 1} V_n$ is dense in $L^2([0,T])$, it is also dense in the metric generated by the norm corresponding to the inner product $\langle .,. \rangle$. The result on convergence in probability follows as a consequence of standard arguments.

Lemma 5.5: Let $\lambda^{(n)} = (\lambda_1^{(n)}, \lambda_2^{(n)}, \ldots, \lambda_{d_n}^{(n)})$ be such that

$$\sum_{j=1}^{d_n} (\lambda_j^{(n)})^2 \to \lambda^2 \text{ as } n \to \infty. \tag{5.51}$$

Then the random variable $\sqrt{n}\sum_{j=1}^{d_n} \lambda_j^{(n)} c_j^{(n)}$ is asymptotically normal with mean zero and variance λ^2.

Proof: In view of Equation (5.36), it follows that

$$\sqrt{n}\sum_{j=1}^{d_n} \lambda_j^{(n)} c_j^{(n)} = n^{-1/2}\sum_{i=1}^{n}\left[\sum_{j=1}^{d_n}\lambda_j^{(n)}\zeta_{ij}^{(n)} + \sum_{j=1}^{d_n}\lambda_j^{(n)}\tilde{\zeta}_{ij}\right]. \tag{5.52}$$

Using the arguments used to derive the inequality (5.41), it can be checked that

$$E\left[\sum_{j=1}^{d_n}\lambda_j^{(n)}\zeta_{ij}^{(n)}\right]^2 \leq C w_T^H n^{-1}\gamma_n \sum_{j=1}^{d_n}(\lambda_j^{(n)})^2||\theta - \theta^{(n)}||^2. \tag{5.53}$$

Note that $E(\zeta_{ij}^{(n)}) = 0$ and

$$
\begin{aligned}
& E\left(\sqrt{n}\sum_{j=1}^{d_n}\lambda_j^{(n)}c_j^{(n)}\right)^2 \\
= \quad & Var\left(\sqrt{n}\sum_{j=1}^{d_n}\lambda_j^{(n)}c_j^{(n)}\right) \\
= \quad & Var\left(\sum_{j=1}^{d_n}\lambda_j^{(n)}\zeta_{1j}^{(n)}\right) + Var\left(\sum_{j=1}^{d_n}\lambda_j^{(n)}\tilde{\zeta}_{1j}\right) \\
& +2Cov\left(\sum_{j=1}^{d_n}\lambda_j^{(n)}\zeta_{1j}^{(n)}, \sum_{j=1}^{d_n}\lambda_j^{(n)}\tilde{\zeta}_{1j}\right) \\
& \text{(since } X_i, 1 \le i \le n, \text{ are i.i.d.)} \\
= \quad & E\left[\sum_{j=1}^{d_n}\lambda_j^{(n)}\zeta_{1j}^{(n)}\right]^2 + E\left[\sum_{j=1}^{d_n}\lambda_j^{(n)}\tilde{\zeta}_{1j}\right]^2 \\
& +2Cov\left(\sum_{j=1}^{d_n}\lambda_j^{(n)}\zeta_{1j}^{(n)}, \sum_{j=1}^{d_n}\lambda_j^{(n)}\tilde{\zeta}_{1j}\right) \\
= \quad & O\left(n^{-1}w_T^H\gamma_n\sum_{j=1}^{d_n}(\lambda_j^{(n)})^2||\theta - \theta^{(n)}||^2\right) + n^{-1}\sum_{j=1}^{d_n}(\lambda_j^{(n)})^2 \\
& +2n^{-1}Cov\left(\sum_{j=1}^{d_n}\lambda_j^{(n)}\zeta_{1j}^{(n)}, \sum_{j=1}^{d_n}\lambda_j^{(n)}\tilde{\zeta}_{1j}\right). \qquad (5.54)
\end{aligned}
$$

The first term on the right of the above equation tends to zero since $\gamma_n/n \le \gamma_n d_n^2/n \to 0$, $||\theta - \theta^{(n)}|| \to 0$ and the second term $\sum_{j=1}^{d_n}(\lambda_j^{(n)})^2 \to \lambda^2$ as $n \to \infty$. It is easy to see that the third term tends to zero as $n \to \infty$ by using the Cauchy–Schwartz inequality. In other words,

$$\sum_{i=1}^{n}\left[\sum_{j=1}^{d_n}\lambda_j^{(n)}\zeta_{ij}^{(n)}\right]^2 = o_p(1) \qquad (5.55)$$

and

$$Var\left(\sum_{j=1}^{d_n}\lambda_j^{(n)}\tilde{\zeta}_{ij}\right) = \sum_{j=1}^{d_n}(\lambda_j^{(n)})^2 = \lambda^2 + o(1) \qquad (5.56)$$

as $n \to \infty$. We now obtain the asymptotic normality from central limit theorems for triangular arrays.

As a consequence of the above lemma, the following theorem can be proved.

Theorem 5.6: Let $\lambda^{(n)}$ be as in the Lemma 5.5. Suppose that the conditions stated in Theorem 5.3 hold. In addition, suppose that

$$\frac{d_n^3 \gamma_n^2}{n} \to 0 \text{ as } n \to \infty$$

and

$$\frac{d_n^3 \gamma_n^3}{n} \to 0 \text{ as } n \to \infty.$$

Then

$$\sqrt{n} \sum_{j=1}^{d_n} \lambda_j^{(n)} (\hat{\eta}_i^{(n)} - \eta_i) \tag{5.57}$$

is asymptotically normal with mean zero and variance λ^2.

Proof: Observe that

$$a^{(n)}(\hat{\eta}^{(n)} - \eta^{(n)}) = c^{(n)} \tag{5.58}$$

and hence

$$\hat{\eta}^{(n)} - \eta^{(n)} = (a^{(n)})^{-1} c^{(n)} \tag{5.59}$$

or equivalently

$$\hat{\eta}^{(n)} - \eta^{(n)} - c^{(n)} = (a^{(n)})^{-1}(I - a^{(n)})c^{(n)}. \tag{5.60}$$

Denoting the operator norm and the Euclidean norm by the same symbol $||.||$, we get

$$|\lambda^{(n)'}(\hat{\eta}^{(n)} - \eta^{(n)} - c^{(n)})| \le ||\lambda^{(n)}||\ ||(a^{(n)})^{-1}||\ ||a^{(n)} - I||\ ||c^{(n)}||. \tag{5.61}$$

Relations (5.47) and (5.48) prove that

$$\begin{aligned} E||a^{(n)} - I||^2 &\le E\left[\sum_{j=1}^{d_n} \sum_{k=1}^{d_n} (a_{j,k}^{(n)} - \delta_{jk})^2\right] \\ &\le C\gamma_n n^{-1} d_n^2 \gamma_n^2 \end{aligned} \tag{5.62}$$

and

$$nE||c^{(n)}||^2 \leq C\gamma_n[d_n + d_n\gamma_n||\theta - \theta^{(n)}||^2]. \tag{5.63}$$

Therefore

$$\begin{aligned}(E[\sqrt{n}||a^{(n)} - I||||c^{(n)}||])^2 \\ \leq nE||c^{(n)}||^2 E||a^{(n)} - I||^2 \\ \leq C\gamma_n([d_n + d_n\gamma_n||\theta - \theta^{(n)}||^2])(n^{-1}d_n^2\gamma_n)\end{aligned} \tag{5.64}$$

and the last term tends to zero since $d_n^3\gamma_n^2/n \to 0$ as $n \to \infty$ and $d_n^3\gamma_n^3/n \to 0$ as $n \to \infty$.Therefore

$$\sqrt{n}||a^{(n)} - I||||c^{(n)}|| \to 0 \tag{5.65}$$

in probability as $n \to \infty$. We observed earlier that

$$||a^{(n)}|| \to 1 \tag{5.66}$$

in probability as $n \to \infty$. Hence

$$\sqrt{n}\lambda^{(n)'}(\hat{\eta}^{(n)} - \eta^{(n)} - c^{(n)}) \to 0 \tag{5.67}$$

in probability as $n \to \infty$. But

$$\sqrt{n}\lambda^{(n)'}c^{(n)}$$

is asymptotically normal with mean zero and variance λ^2 by Lemma 5.5. This proves the result stated in the theorem.

As an application of the previous theorem, we get the following result.

Corollary 5.7: Let $h(.)$ be a function such that $||h|| < \infty$ in the sense of the inner product defined by (5.25). Suppose that the conditions stated in Theorem 5.6 hold. Then

$$\sqrt{n}\langle h, \hat{\theta}^{(n)} - \theta^{(n)}\rangle \tag{5.68}$$

is asymptotically normal with mean zero and variance $\langle h, h\rangle$.

Proof: Suppose that $h(t) = \sum_{j=1}^{\infty} h_j\psi_j(t)$. Note that

$$\hat{\theta}^{(n)} - \theta^{(n)} = \sum_{j=1}^{d_n}(\hat{\eta}_j^{(n)} - \eta_j)\psi_j \tag{5.69}$$

and hence

$$\langle h, \hat{\theta}^{(n)} - \theta^{(n)} \rangle = \sum_{j=1}^{d_n} h_j (\hat{\eta}_j^{(n)} - \eta_j). \tag{5.70}$$

Since

$$\sum_{j=1}^{d_n} h_j^2 \to \langle h, h \rangle = ||h||^2 \tag{5.71}$$

by Parseval's theorem, the result follows from Theorem 5.3.

Remarks:

(i) If, in addition to the conditions stated in Corollary 5.7, we have

$$\sqrt{n} \langle h, \theta^{(n)} - \theta^{(n)} \rangle \to 0 \ \text{as} \ n \to \infty, \tag{5.72}$$

then

$$\sqrt{n} \langle h, \hat{\theta}^{(n)} - \theta \rangle \tag{5.73}$$

is asymptotically normal with mean zero and variance $\langle h, h \rangle$.

(ii) It would be interesting to characterize the family of functions $\theta(t)$ and the family of processes $X(t)$ satisfying (5.4) for which

$$E \left(\int_0^T Q_H^2, \theta(s) dw_s^H \right) < \infty.$$

Note that

$$\begin{aligned} Q_{H,\theta}(t) &= \frac{d}{dw_t^H} \int_0^t k_H(t,s) \theta(s) X(s) ds \\ &= \frac{dt}{dw_t^H} \frac{d}{dt} \left[\int_0^t k_H(t,s) \theta(s) X(s) \right] ds \\ &= \frac{dt}{dw_t^H} \int_0^t \frac{dk_H(t,s)}{dt} \theta(s) X(s) ds \end{aligned} \tag{5.74}$$

from the form of the function $k_H(t,s)$. Hence

$$Q_{H,\theta}^2(t) = \left(\frac{dt}{dw_t^H} \right)^2 \int_0^t \int_0^t \frac{dk_H(t,s)}{dt} \theta(s) X(s) \frac{dk_H(t,u)}{dt} \theta(u) X(u) ds du. \tag{5.75}$$

Therefore

$$
\begin{aligned}
E\left(\int_0^T Q_{H,\theta}^2(t)dw_t^h\right) &= E\left[\int_0^T \left(\frac{dt}{dw_t^H}\right)^2 \left(\int_0^t \int_0^t \frac{dk_H(t,s)}{dt}\right.\right.\\
&\quad \left.\left.\times\ \theta(s)X(s)\frac{dk_H(t,u)}{dt}\theta(u)X(u)dsdu\right) dw_t^H\right]\\
&= \int_0^T \left(\frac{dt}{dw_t^H}\right)^2 \left(\int_0^t \int_0^t \frac{dk_H(t,s)}{dt}\right.\\
&\quad \left.\times\ \theta(s)\frac{dk_H(t,u)}{dt}\theta(u)E(X(s)X(u))dsdu\right) dw_t^H\\
&= D_{H1}\int_0^T t^{4H-2}\left(\int_0^t \int_0^t s^{\frac{1}{2}-H}(t-s)^{-\frac{1}{2}-H}u^{\frac{1}{2}-H}\right.\\
&\quad \left.\times (t-u)^{-\frac{1}{2}-H}\theta(s)\theta(u)E(X(s)X(u))dsdu\right) dw_t^H\\
&= D_{H2}\int_0^T t^{2H-1}\left(\int_0^t \int_0^t s^{\frac{1}{2}-H}(t-s)^{-\frac{1}{2}-H}u^{\frac{1}{2}-H}\right.\\
&\quad \left.\times (t-u)^{-\frac{1}{2}-H}\theta(s)\theta(u)E(X(s)X(u))dsdu\right) dt
\end{aligned}
\tag{5.76}
$$

where D_{H1} and D_{H2} are constants which can be explicitly computed. The finiteness of the term on the right of the above equation depends on the function $\theta(t)$ and the covariance function of the process $X(t)$. It is not clear whether it is possible to characterize the class of functions $\theta(t)$ for which

$$E\left(\int_0^T Q_{H,\theta}^2(t)dw_t^h\right) < \infty$$

explicitly without knowing the covariance structure of the process $X(t)$.

5.3 Nonparametric estimation of trend

Let $W^H = \{W_t^H, t \geq 0\}$ be standard fBm with Hurst parameter $H \in (\frac{1}{2}, 1)$, that is, a Gaussian process with continuous sample paths such that $W_0^H = 0, E(W_t^H) = 0$ and

$$E(W_s^H W_t^H) = \frac{1}{2}[s^{2H} + t^{2H} - |s-t|^{2H}],\ t \geq 0, s \geq 0. \tag{5.77}$$

Let us consider the SDE

$$dX_t = S(X_t)\ dt + \epsilon\ dW_t^H,\ X_0 = x_0, 0 \leq t \leq T \tag{5.78}$$

where the function $S(.)$ is unknown. The problem is to estimate the function $S(.)$ based on the observation $\{X_t, 0 \leq t \leq T\}$. Suppose $\{x_t, 0 \leq t \leq T\}$ is the solution of the differential equation

$$\frac{dx_t}{dt} = S(x_t),\ x_0, 0 \leq t \leq T. \tag{5.79}$$

We assume that the trend coefficient $S(X_t)$ satisfies the following conditions which ensure the existence and uniqueness of the solution of Equation (5.78):

(A_1): for all $N>0$, there exists $L_N>0$ such that

$$|S(X_t) - S(Y_t)| \leq L_N|X_t - Y_t|$$

whenever $|X_t| \leq N$ and $|Y_t| \leq N$ and $t \in [0, T]$.

(A_2): There exists a constant $M>0$ such that

$$|S(X_t)| \leq M(1 + |X_t|),\ t \in [0, T].$$

Existence and uniqueness of the solution of the SDE (5.78) follow as a special case of the results in Nualart and Rascanu (2002).

Lemma 5.8: Let the function $S(.)$ satisfy the conditions (A_1) and (A_2) and suppose that $L_N = L$ for all $N \geq 1$. Let X_t and x_t be the solutions of Equations (5.78) and (5.79) respectively. Then, with probability one,

$$\text{(a) } |X_t - x_t|<e^{Lt}\epsilon|W_t^H| \tag{5.80}$$

and

$$\text{(b) } \sup_{0\leq t\leq T} E(X_t - x_t)^2 \leq e^{2LT}\epsilon^2T^{2H}. \tag{5.81}$$

Proof:

(a) Let $u_t = |X_t - x_t|$. Then, by (A_1), we have

$$\begin{aligned} u_t &\leq \int_0^t |S(X_v) - S(x_v)|\,dv + \epsilon\ |W_t^H| \\ &\leq L\int_0^t u_v dv + \epsilon\ |W_t^H|. \end{aligned} \tag{5.82}$$

Applying Gronwall's lemma (cf. Lemma 1.12 in Kutoyants (1994), p. 26), it follows that

$$u_t \leq \epsilon|W_t^H|e^{Lt}. \tag{5.83}$$

(b) From (5.81), we have

$$E(X_t - x_t)^2 \le e^{2Lt}\epsilon^2 E(|W_t|^H)^2 = e^{2Lt}\epsilon^2 t^{2H}. \tag{5.84}$$

Hence

$$\sup_{0\le t\le T} E(X_t - x_t)^2 \le e^{2LT}\epsilon^2 T^{2H}. \tag{5.85}$$

Let $\Theta(L)$ denote the class of trend coefficients $S(X_t)$ satisfying conditions (A_1) and (A_2) with the same constant L. For $0<\alpha\le 1$, let $\Theta_{\alpha,k}$ denote the class of non-random functions $g(t)$ defined in the interval $[0,T]$ which are k-times differentiable with respect to t and which satisfy the condition

$$|g^{(k)}(t) - g^{(k)}(s)| \le L_\alpha |t-s|^\alpha,\ t,s\in[0,T]$$

for some constant $L_\alpha>0$. Here $g^{(k)}(t)$ denotes the kth derivative of $g(.)$ at t for $k\ge 0$. If $k=0$, we interpret $g^{(0)}$ as g. Let $\Theta_{1,k}(L)$ be the class of trend coefficients $S(X_t)$ belonging to $\Theta(L)$ such that $S(x_t)$ belongs to $\Theta_{1,k}$ when $\epsilon=0$. Let $G(u)$ be a bounded function with finite support $[A,B]$ satisfying the condition

(A_3): $G(u)=0$ for $u<A$ and $u>B$, and $\int_A^B G(u)du = 1.$

It is obvious that the following conditions are satisfied by the function $G(.)$:

(i) $\int_{-\infty}^{\infty} G^2(u)du<\infty$;

(ii) $\int_{-\infty}^{\infty} u^{2(k+1)}G^2(u)du<\infty$; and

(iii) $\int_{-\infty}^{\infty} (G(u))^{1/H}du<\infty$.

Following the procedure adapted in Kutoyants (1994), we define a kernel-type estimator of the trend $S(X_t)$ as

$$\widehat{S}_t = \frac{1}{\varphi_\epsilon}\int_0^T G\left(\frac{\tau - t}{\varphi_\epsilon}\right) dX_\tau \tag{5.86}$$

where the normalizing function $\varphi_\epsilon \to 0$ with $\epsilon^2\varphi_\epsilon^{-1}\to 0$ as $\epsilon\to 0$.

Theorem 5.9: Suppose that the trend function $S(X_t)\in\Theta(L)$ and the function $\varphi_\epsilon\to 0$ such that $\epsilon^2\varphi_\epsilon^{-1}\longrightarrow 0$ as $\epsilon\to 0$. Suppose the conditions $(A_1), (A_2)$ and (A_3) are satisfied. Then, for any $0<c\le d<T$, the estimator $\widehat{S}_t$ is uniformly consistent, that is,

$$\lim_{\epsilon\to 0}\sup_{S(X_t)\in\Theta(L)}\sup_{c\le t\le d} E_S(|\widehat{S}_t - S(x_t)|^2) = 0. \tag{5.87}$$

In addition to the conditions (A_1)–(A_3), assume that

(A_4): $\int_{-\infty}^{\infty} u^j G(u)du = 0$ for $j = 1, 2, \ldots, k$.

Theorem 5.10: Suppose that the function $S(x_t) \in \Theta_{1,k}(L)$ and $\varphi_\epsilon = \epsilon^{1/(k-H+2)}$. Then, under the conditions $(A_1), (A_2), (A_3)$ and (A_4),

$$\limsup_{\epsilon\to 0} \sup_{S(x)\in\Theta_{1,k}(L)} \sup_{c\le t\le d} E_S(|\widehat{S}_t - S(x_t)|^2)\epsilon^{-2(k+1)/(k-H+2)} < \infty. \tag{5.88}$$

Theorem 5.11: Suppose that the function $S(x_t) \in \Theta_{1,k+1}(L)$ and $\varphi_\epsilon = \epsilon^{1/(k-H+2)}$. Then, under the conditions $(A_1), (A_2), (A_3)$ and (A_4), the asymptotic distribution of

$$\epsilon^{-(k+1)/(k-H+2)}(\widehat{S}_t - S(x_t))$$

is Gaussian with the mean

$$m = \frac{S^{(k+1)}(x_t)}{(k+1)!}\int_{-\infty}^{\infty} G(u)u^{k+1}du$$

and the variance

$$\sigma^2 = H(2H-1)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(u)G(v)|u-v|^{2H-2}dudv$$

as $\epsilon \to 0$.

Suppose the trend function $S^{(0)}(X_t) \in \Theta_{1,k}(L/2)$. Fix $t_0 \in (0, T]$ and $0 \le t \le T$. Suppose the function $g_t(v, X_t) \in \Theta_{1,k+1}(L/2)$ where

$$g_{t_0}(0, x_t) = 1$$
$$g_{t_0}(v, x_t) = 0 \text{ if } |v| > \frac{1}{2}.$$

Here $x = \{x_s, 0 \le s \le t\}$ is a solution of Equation (5.79) with $S(x_t) = S^{(0)}(x_t)$ and suppose that

$$\gamma = \int_{-1/2}^{1/2} [g_{t_0}(v, x_t)]^2 dv > 0. \tag{5.89}$$

Let

$$S^{(\epsilon)}(\theta, X_t) = S^{(0)}(X_t) + \theta\epsilon^{(k+1)/(k-H+2)} g_{t_0}(\gamma(t-t_0)\epsilon^{-2(1-H)/(k-H+2)}, X_t) \tag{5.90}$$

where $\theta \in \Theta = \{\theta : |\theta - \theta_0| < \gamma\}$. Consider the family of processes

$$dX_t = S^{(\epsilon)}(\theta, X_t)dt + \epsilon\, dW_t^H,\ X_0 = x_0, 0 \le t \le T, \theta \in \Theta. \tag{5.91}$$

Then, we have

$$dX_t = (S^{(0)}(X_t) + \theta \epsilon^{(k+1)/(k-H+2)} g_{t_0}(\gamma(t-t_0)\epsilon^{-2(1-H)/(k-H+2)}, X_t))dt + \epsilon\, dW_t^H. \tag{5.92}$$

Following the notation introduced in Chapter 1, let

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

$$k_H(t,u) = c_H^{-1} u^{\frac{1}{2}-H}(t-u)^{\frac{1}{2}-H},$$

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

$$w_t^H = \lambda_H^{-1} t^{2-2H},$$

and

$$Q_{H,\theta}(t) = \frac{d}{dw_t^H}\frac{1}{\epsilon}\int_0^t k_H(t,s) S^{(\epsilon)}(\theta, X_s)ds. \tag{5.93}$$

Suppose the sample paths of the process $\{Q_{H\theta}(t), 0 \le t \le T\}$ belong a.s. to $L^2(0,T)$. Define

$$Z_t = \frac{1}{\epsilon}\int_0^t k_H(t,s)dX_s,\ t \ge 0 \tag{5.94}$$

Then the process $Z = \{Z_t, t \ge 0\}$ is an $\mathcal{F}_t$-semimartingale with the decomposition

$$Z_t = \int_0^t Q_{H,\theta}(s)dw_s^H + M_t^H \tag{5.95}$$

where $\{M_t^H, \mathcal{F}_t, t \ge 0\}$ is the fundamental martingale as described in Kleptsyna and Le Breton (2002a). Note that

$$\begin{aligned} Q_{H,\theta}(t) &= \frac{d}{dw_t^H}\frac{1}{\epsilon}\int_0^t k_H(t,s)S^{(\epsilon)}(\theta, X_s)ds \\ &= \frac{d}{dw_t^H}\frac{1}{\epsilon}\int_0^t k_H(t,s)(S^{(0)}(X_s) \\ &\quad + \theta\epsilon^{(k+1)/(k-H+2)} g_{t_0}(\gamma(s-t_0)\epsilon^{-2(1-H)/(k-H+2)}, X_s))ds \\ &= J_1(t) + \theta J_2(t)\ \ \text{(say)}. \end{aligned} \tag{5.96}$$

Then

$$dZ_t = [J_1(t) + \theta J_2(t)]\, dw_t^H + dM_t^H. \tag{5.97}$$

Let $\{P_\theta^{(\epsilon)}, \theta \in \Theta\}$ be the family of probability measures induced on the space $C[0,T]$, the space of continuous functions in the interval $[0,T]$ endowed with Borel σ-algebra generated by the supremum norm. Following Kleptsyna and Le Breton (2002a), it follows that the probabilitymeasures $P^{(\epsilon)}_{\theta_0+uA_\epsilon^{-1/2}}$ and $P^{(\epsilon)}_{\theta_0}$ are absolutely continuous with respect to each other and their Radon–Nikodym derivative is given by

$$\begin{aligned}\frac{dP^{(\epsilon)}_{\theta_0+uA_\epsilon^{-1/2}}}{dP^{(\epsilon)}_{\theta_0}}(X) &= \exp\left[uA_\epsilon^{-\frac{1}{2}}\int_0^T J_2(t)dM_t^H - \frac{1}{2}u^2A_\epsilon^{-1}\int_0^T J_2^2(t)dw_t^H\right]\\ &= \exp\left[uA_\epsilon^{-\frac{1}{2}}\epsilon^{(H-1)/(k-H+2)}\int_0^T J_2^*(t)dM_t^H\right.\\ &\quad\left. - \frac{1}{2}u^2A_\epsilon^{-1}\epsilon^{2(H-1)/(k-H+2)}\int_0^T (J_2^*(t))^2 dw_t^H\right].\end{aligned} \tag{5.98}$$

where A_ϵ is a positive non-random function and

$$J_2^*(t) = \frac{d}{dw_t^H}\int_0^t k_H(t,s)g_{t_0}(\gamma(s-t_0)\epsilon^{2(H-1)/(k-H+2)}, X_s)ds. \tag{5.99}$$

Denote

$$R_{T,\epsilon} = \int_0^T J_2^*(t)dM_t^H \tag{5.100}$$

and its quadratic variation $\langle R\rangle_{T,\epsilon}$ is

$$I_{T,\epsilon}(\theta_0) = \int_0^T (J_2^*(t))^2 dw_t^H. \tag{5.101}$$

Then the representation (5.98) can be written as

$$\begin{aligned}&\frac{dP^{(\epsilon)}_{\theta_0+uA_\epsilon^{-\frac{1}{2}}}}{dP^{(\epsilon)}_{\theta_0}}(X)\\ &= \exp\left(\frac{u}{A_\epsilon^{\frac{1}{2}}\epsilon^{(1-H)/(k-H+2)}}R_{T,\epsilon} - \frac{1}{2}\frac{u^2}{A_\epsilon\epsilon^{2(1-H)/(k-H+2)}}I_{T,\epsilon}(\theta_0)\right).\end{aligned} \tag{5.102}$$

Assume that:

(A_5)(i): $A_\epsilon\epsilon^{2(1-H)/(k-H+2)} \to \infty$ as $\epsilon \to 0$;

(A_5)(ii): $P_{\theta_0}^{(\epsilon)}\left\{\left|\frac{I_{T,\epsilon}(\theta_0)}{A_\epsilon^{1/2}\epsilon^{(1-H)/(k-H+2)}}-1\right|\geq \delta_\epsilon\right\}=O\left(\delta_\epsilon^{1/2}\right)$ where $\delta_\epsilon \longrightarrow 0$ as $\epsilon \to 0$;

(A_5)(iii): $\lim_{\epsilon\to 0}\frac{I_{T,\epsilon}(\theta_0)}{A_\epsilon^{1/2}\epsilon^{(1-H)/(k-H+2)}}=1$ a.s. (P_{θ_0}).

Let $\ell(.)$, the loss function, satisfy the following conditions:

(a) $\ell(.)$ is non negative, continuous at zero and $\ell(0)=0$;

(b) symmetric, that is, $\ell(y)=\ell(-y)$; and

(c) the set $\{y:\ell(y)\leq c\}$ is convex for all $c>0$.

We will prove the following result.

Theorem 5.12: Let $\ell(.)$ be the loss function as defined above and suppose that the trend coefficient $S(.)\in\Theta_{1,k}(L)$. Then, under the conditions $(A_1)-(A_5)$,

$$\liminf_{\epsilon\to 0}\inf_{S_{t_0}^*}\sup_{S\in\Theta_{1,k}(L)} E_s\ell(\epsilon^{-(k+1)/(k-H+2)}(S_{t_0}^*-S(x_{t_0})))>0. \tag{5.103}$$

The infimum here has been taken over all possible estimators $S_{t_0}^*$ of the function $S(x_t)$ at the point t_0.

Proof of Theorem 5.9: From (5.89) we have

$$\begin{aligned}
E_S[(\widehat{S}_t-S(x_t))^2] &= E_S\left[\frac{1}{\varphi_\epsilon}\int_0^T G\left(\frac{\tau-t}{\varphi_\epsilon}\right)(S(X_\tau)-S(x_\tau))\,d\tau\right.\\
&\quad+\frac{1}{\varphi_\epsilon}\int_0^T G\left(\frac{\tau-t}{\varphi_\epsilon}\right)S(x_\tau)d\tau-S(x_t)\\
&\quad\left.+\frac{\epsilon}{\varphi_\epsilon}\int_0^T G\left(\frac{\tau-t}{\varphi_\epsilon}\right)dW_\tau^H\right]^2\\
&\leq 3E_S\left[\frac{1}{\varphi_\epsilon}\int_0^T G\left(\frac{\tau-t}{\varphi_\epsilon}\right)(S(X_\tau)-S(x_\tau))d\tau\right]^2\\
&\quad+3E_S\left[\frac{1}{\varphi_\epsilon}\int_0^T G\left(\frac{\tau-t}{\varphi_\epsilon}\right)S(x_\tau)d\tau-S(x_t)\right]^2\\
&\quad+\frac{3\epsilon^2}{\varphi_\epsilon{}^2}E_S\left[\int_0^T G\left(\frac{\tau-t}{\varphi_\epsilon}\right)dW_\tau^H\right]^2\\
&= I_1+I_2+I_3(\text{say}).
\end{aligned}\tag{5.104}$$

Note that

$$
\begin{aligned}
I_1 &= 3E_S\left[\frac{1}{\varphi_\epsilon}\int_0^T G\left(\frac{\tau-t}{\varphi_\epsilon}\right)(S(X_\tau)-S(x_\tau))d\tau\right]^2 \\
&= 3E_S\left[\int_{-\infty}^{\infty} G(u)\left(S(X_{t+\varphi_\epsilon u})-S(x_{t+\varphi_\epsilon u})\right)du\right]^2 \\
&\le 3(B-A)\int_{-\infty}^{\infty} G^2(u)L^2E\left(X_{t+\varphi_\epsilon u}-x_{t+\varphi_\epsilon u}\right)^2du \\
&\quad \text{(by using the condition } (A_1)) \\
&\le 3(B-A)\int_{-\infty}^{\infty} G^2(u)L^2 \sup_{0\le t+\varphi_\epsilon u\le T} E\left(X_{t+\varphi_\epsilon u}-x_{t+\varphi_\epsilon u}\right)^2du \\
&\le C_1\epsilon^2 \ \text{(by using (5.85))}
\end{aligned}
\tag{5.105}
$$

for some positive constant C_1 depending on H, T, L and $B-A$. Furthermore,

$$
\begin{aligned}
I_2 &= 3E_S\left[\frac{1}{\varphi_\epsilon}\int_0^T G\left(\frac{\tau-t}{\varphi_\epsilon}\right)S(x_\tau)d\tau - S(x_t)\right]^2 \\
&\le 3L^2E_S\left[\int_{-\infty}^{\infty} G(u)\left(S(x_{t+\varphi_\epsilon u})-S(x_t)\right)du\right]^2 \\
&\le 3\left[\int_{-\infty}^{\infty} G(u)u\varphi_\epsilon du\right]^2 \\
&\le C_2\varphi_\epsilon^2\int_{-\infty}^{\infty} G^2(u)u^2du \\
&\le C_3\varphi_\epsilon^2 \ \text{(by } (A_3)\text{(ii))}
\end{aligned}
\tag{5.106}
$$

for some positive constant C_3 depending on T, L and $B-A$. Furthermore, the last term tends to zero as $\epsilon \to 0$. In addition, for $\frac{1}{2}<H<1$,

$$
\begin{aligned}
I_3 &= \frac{3\epsilon^2}{\varphi\epsilon^2}E_S\left(\int_0^T G\left(\frac{\tau-t}{\varphi_\epsilon}\right)dW_\tau^H\right)^2 \\
&\le \frac{3\epsilon^2}{\varphi_\epsilon^2}C_4(2,H)\left\{\int_0^T\left[G\left(\frac{\tau-t}{\varphi_\epsilon}\right)\right]^{1/H}d\tau\right\}^{2H} \\
&\quad \text{(cf. Memin } et\ al. \text{ (2001))} \\
&\le \frac{C_5\epsilon^2}{\varphi_\epsilon^2}\varphi_\epsilon^{2H} \ \text{((by using } (A_3)\text{(iii))} \\
&= C_6\frac{\epsilon^2}{\varphi_\epsilon}\varphi_\epsilon^{2H-1}
\end{aligned}
\tag{5.107}
$$

for some positive constant C_6 depending on H and T. Theorem 5.9 is now proved by using Equations (5.103) to (5.106).

Proof of Theorem 5.10: By Taylor's formula,

$$S(x_{t+\varphi_\epsilon u}) = S(x_t) + \sum_{j=1}^{k} S^{(j)}(x_t)\frac{(\varphi_\epsilon u)^j}{j!} + \left[S^{(k)}(x_{t+\tau(\varphi_\epsilon u)}) - S^{(k)}(x_t)\right]\frac{(\varphi_\epsilon u)^k}{k!} \tag{5.108}$$

for some $0 \leq \tau \leq 1$. Using this expansion in I_2 defined in the proof of Theorem 5.9 and using the conditions (A_4) and (A_5), we obtain that, for sufficiently small ϵ,

$$\begin{aligned}
I_2 &\leq 3E_S\left[\int_{-\infty}^{\infty} G(u)\left(S(x_{t+\varphi_\epsilon u}) - S(x_t)\right)du\right]^2 \\
&= 3E_S\left[\sum_{j=1}^{k} S^{(j)}(x_t)\left(\int_{-\infty}^{\infty} G(u)u^j du\right)\varphi_\epsilon^j (j!)^{-1}\right. \\
&\quad \left. + \left(\int_{-\infty}^{\infty} G(u)u^k(S^{(k)}(x_{t+\tau\varphi_\epsilon u}) - S^{(k)}(x_t))du\right)\varphi_\epsilon^k (k!)^{-1}\right]^2 \\
&\leq C_7 L^2 E_S\left[\int_{-\infty}^{\infty} G(u)u^{k+1}\varphi_\epsilon^{k+1}(k!)^{-1}du\right]^2 \\
&\leq C_8(B-A)(k!)^{-2}\varphi_\epsilon^{2(k+1)}\int_{-\infty}^{\infty} G^2(u)u^{2(k+1)}du \leq C_9\varphi_\epsilon^{2(k+1)}
\end{aligned} \tag{5.109}$$

for some positive constant C_9 depending on H, T, L and $B-A$. Combining the relations (5.104), (5.106) and (5.108),we get that there exists a positive constant C depending on H, T, L and $B-A$ such that

$$\sup_{c\leq t\leq d} E_S|\widehat{S}_t - S(x_t)|^2 \leq C(\epsilon^2\varphi_\epsilon^{2H-2} + \varphi_\epsilon^{2(k+1)} + \epsilon^2). \tag{5.110}$$

Choosing $\varphi_\epsilon = \epsilon^{1/(k-H+2)}$, we get

$$\limsup_{\epsilon\to 0}\ \sup_{S(x)\in\Theta_{1,k}(L)}\ \sup_{c\leq t\leq d} E_S|\widetilde{S}_t - S(x_t)|^2\epsilon^{-2(k+1)/(k-H+2)} < \infty. \tag{5.111}$$

This completes the proof of Theorem 5.10.

Remarks: Choosing $\varphi_\epsilon = \epsilon^{1/(2-H)}$ and without assuming condition (A_4), it can be shown that

$$\limsup_{\epsilon\to 0}\ \sup_{S(x)\in\Theta_{1,0}(L)}\ \sup_{c\leq t\leq d} E_S|\widehat{S}_t - S(x_t)|^2\epsilon^{-2/(2-H)} < \infty \tag{5.112}$$

which gives a slower rate of convergence than the one obtained in Theorem 5.10.

Proof of Theorem 5.11: From (5.89), we obtain that

$$\begin{aligned}
\widehat{S}_t - S(x_t) &= \left[\frac{1}{\varphi_\epsilon}\int_0^T G\left(\frac{\tau - t}{\varphi_\epsilon}\right)(S(X_\tau) - S(x_\tau))d\tau \right.\\
&\quad + \frac{1}{\varphi_\epsilon}\int_0^T G\left(\frac{\tau - t}{\varphi_\epsilon}\right)S(x_\tau)d\tau - S(x_t)\\
&\quad \left. + \frac{\epsilon}{\varphi_\epsilon}\int_0^T G\left(\frac{\tau - t}{\varphi_\epsilon}\right)dW_\tau^H\right]\\
&= \left[\int_{-\infty}^{\infty} G(u)(S(X_{t+\varphi_\epsilon} u) - S(x_{t+\varphi_\epsilon u}))du \right.\\
&\quad + \int_{-\infty}^{\infty} G(u)(S(x_{t+\varphi_\epsilon u}) - S(x_t))du\\
&\quad \left. + \frac{\epsilon}{\varphi_\epsilon}\int_0^T G\left(\frac{\tau - t}{\varphi_\epsilon}\right)dW_\tau^H\right]. \qquad (5.113)
\end{aligned}$$

Let $\varphi_\epsilon = \epsilon^{1/(k-H+2)}$ and

$$[\widehat{S}_t - S(x_t)] = R_1 + R_2 + R_3 \ \text{(say)}. \qquad (5.114)$$

By Taylor's formula

$$\begin{aligned}
S(x_{t+\varphi_\epsilon u}) &= S(x_t) + \sum_{j=1}^{k} S^{(j)}(x_t)\frac{(\varphi_\epsilon u)^j}{j!}\\
&\quad + \frac{S^{(k+1)}(x_t)}{(k+1)!}(\varphi_\epsilon u)^{k+1} + \left[S^{(k+1)}(x_{t+\tau(\varphi_\epsilon u)})\right.\\
&\quad \left. - S^{(k+1)}(x_t)\right]\frac{(\varphi_\epsilon u)^{k+1}}{(k+1)!}. \qquad (5.115)
\end{aligned}$$

In view of relation (5.113), applying condition (A_4), we get

$$\begin{aligned}
E_S(R_2 - m)^2 &= E_S\left[\int_{-\infty}^{\infty} G(u)\left(S^{(k+1)}(x_{t+\tau(\varphi_\epsilon u)}) - S^{(k+1)}(x_t)\right)\frac{(\varphi_\epsilon u)^{k+1}}{(k+1)!}du\right]^2\\
&\leq C_{10}L^2 E_S\left(\int_{\infty}^{\infty} G(u)u^{(k+2)}\frac{\varphi_\epsilon^{k+2}}{(k+1)!}du\right)^2 \quad \text{(by } (A_3))\\
&\leq C_{11}\varphi_\epsilon^{2(k+2)} \ \ \text{(by } \ (A_3)) \qquad (5.116)
\end{aligned}$$

for some positive constant C_{11} depending on T, L, H and $B-A$. Therefore

$$\epsilon^{-2(k+1)/(k-H+2)}E_S(R_2 - m)^2 \to 0 \ \text{ as } \ \epsilon \to 0. \qquad (5.117)$$

Furthermore,

$$0 \leq \epsilon^{-2(k+1)/(k-H+2)} E_S[R_1^2] = \epsilon^{-2(k+1)/(k-H+2)} O(\phi_\epsilon^{2(k+2)})$$

by arguments similar to those given for proving the inequality (5.116). Hence

$$\epsilon^{-2(k+1)/(k-H+2)} E_S[R_1^2] \to 0 \text{ as } \epsilon \to 0. \tag{5.118}$$

In addition, it follows that $E_S[R_3^2]$ is finite by (A_3)(iii) and the variance of the Gaussian random variable

$$\int_{-\infty}^{\infty} G(t) dW_t^H$$

is

$$H(2H-1) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(u)G(v)|u-v|^{2H-2} du dv$$

by Equation (1.23) in Chapter 1. Combining these observations, an application of Slutsky's lemma proves Theorem 5.11.

Proof of Theorem 5.12: From the results on the representation of locally continuous square integrable martingales (cf. Ikeda and Watanabe (1981), Chapter II, Theorem 7.2), it follows that there exists a standard Wiener process $\{W^*(t), t \geq 0\}$ adapted to the filtration $\{\mathcal{F}_t, t \geq 0\}$ such that

$$\frac{R_{t,\epsilon}}{A_\epsilon^{\frac{1}{2}} \epsilon^{(1-H)/(k-H+2)}} = W^* \left(\frac{I_{t,\epsilon}(\theta_0)}{A_\epsilon^{\frac{1}{2}} \epsilon^{(1-H)/(k-H+2)}} \right), \ t \geq 0 \quad \text{a.s.}(P_{\theta_0}). \tag{5.119}$$

Hence Equation (5.98) can be written as

$$\log \frac{dP^{(\epsilon)}_{\theta_0 + uA_\epsilon^{-\frac{1}{2}}}}{dP^{(\epsilon)}_{\theta_0}} = uW^* \left(\frac{I_{t,\epsilon}(\theta_0)}{A_\epsilon^{\frac{1}{2}} \epsilon^{(1-H)/(k-H+2)}} \right) - \frac{u^2}{2}. \tag{5.120}$$

Now

$$\begin{aligned} &\left| P\left\{ W^* \left(\frac{I_{t,\epsilon}(\theta_0)}{A_\epsilon^{\frac{1}{2}} \epsilon^{(1-H)/(k-H+2)}} \right) \leq x \right\} - \Phi(x) \right| \\ &\leq (2\delta_\epsilon)^{\frac{1}{2}} + P\left\{ \left| \frac{I_{t,\epsilon}(\theta_0)}{A^{\frac{1}{2}} \epsilon^{(1-H)/(k-H+2)}} - 1 \right| \geq \delta_\epsilon \right\} \\ &\leq (2\delta_\epsilon)^{\frac{1}{2}} + \psi_{\delta_\epsilon}(\theta_0) = O(\delta_\epsilon^{1/2}) \end{aligned} \tag{5.121}$$

by the condition (A_5) where $\delta_\epsilon \to 0$ as $\epsilon \to 0$. Therefore

$$W^* \left(\frac{I_{t,\epsilon}(\theta_0)}{A^{\frac{1}{2}} \epsilon^{(1-H)/(k-H+2)}} \right) \sim N(0,1) \tag{5.122}$$

and hence, from Equation (5.98), we get

$$\log \frac{dP^{(\epsilon)}_{\theta_0 + uA_\epsilon^{-\frac{1}{2}}}}{dP_{\theta_0}} = u\Delta_\epsilon - \frac{u^2}{2} + \psi_{\delta_\epsilon}(\theta_0) \tag{5.123}$$

where $\Delta_\epsilon \xrightarrow{\mathcal{L}} N(0,1)$ as $\epsilon \to 0$ and $\lim_{\epsilon \to 0} \psi_{\delta_\epsilon}(\theta_0) = 0$. Following Remark 2.3 in Kutoyants (1994), p. 44, we can choose functions $g_{t_0}(0, x_t)$ and $\gamma > 0$ such that

$$\liminf_{\epsilon \to 0} \sup_{|\theta - \theta_0| < \gamma} E_\theta[\ell(\theta_\epsilon - \theta)] \geq \frac{1}{2\sqrt{2\pi}} \int \ell(y) \epsilon^{-y^2/2} dy. \tag{5.124}$$

Let S_t^* be any estimator of $S^{(\epsilon)}(\theta, x_t)$. Let

$$\theta_\epsilon^* = [S_{t_0}^* - S^{(0)}(X_{t_0})][g_{t_0}(0, X_{t_0})]^{-1} \epsilon^{-(k+1)/(k-H+2)}.$$

Then

$$\begin{aligned} &\sup_{S(.) \in \Theta_{1,k}(L)} E_S[\ell(\epsilon^{-(k+1)/(k-H+2)}(S_{t_0}^* - S(x_{t_0})))] \\ &\quad \geq \sup_{|\theta - \theta_0| < \gamma} E_S[\ell(\epsilon^{-(k+1)/(k-H+2)}(S_{t_0}^* - S^{(\epsilon)}(\theta, x_{t_0})))] \\ &\quad \geq \sup_{|\theta - \theta_0| < \gamma} E_\theta[\ell(\theta_\epsilon^* - \theta + O(1))] \\ &\quad \geq \frac{1}{2\sqrt{2\pi}} \int \ell(y) \epsilon^{-y^2/2} dy \end{aligned} \tag{5.125}$$

following the arguments given by Kutoyants (1994), p. 151. Hence

$$\liminf_{\epsilon \to 0} \inf_{S_{t_0}^*} \sup_{S \in \Theta_{1,k}(L)} E_s \ell \left(\epsilon^{-(k+1)/(k-H+2)} (S_{t_0}^* - S(x_{t_0})) \right) > 0. \tag{5.126}$$

This proves Theorem 5.12.

Remarks: Since the trend coefficient $S^{(\epsilon)}(\theta, X_t) \in \Theta_{1,k}(L)$, we can apply Theorem 5.10 for the function $\ell(u) = u^2$ and obtain that the estimator $\widehat{S}_t$ has an asymptotically optimal rate of convergence when $\varphi_\epsilon = \epsilon^{1/(k-H+2)}$.

The results discussed in this section are due to Mishra and Prakasa Rao (2009a).

6

Parametric inference for some SDEs driven by processes related to fBm

6.1 Introduction

We studied parametric and nonparametric inference for some processes driven by fBm in earlier chapters. We now consider extensions of these results and other problems related to processes driven by fBm or mixed fBm. We will also discuss the problem of estimation for translation of a process driven by fBm.

6.2 Estimation of the translation of a process driven by fBm

Baran and Pap (2003) considered the problem of estimation of the mean for the translation of an Ornstein–Uhlenbeck process. We now consider similar problems for processes governed by SDEs driven by fBm. Among other things, we obtain sufficient conditions for the absolute continuity of the measures generated by a stochastic process $\{Y(t), 0 \leq t \leq T\}$ driven by fBm with Hurst index $H \in (0, 1)$ and its translation $\{\tilde{Y}(t), 0 \leq t \leq T\}$, with $\tilde{Y}(t) = Y(t) + g(t)$ and $g(t)$ non-random, and obtain the Radon–Nikodym derivative in case the measures are absolutely continuous. As a consequence we study the maximum likelihood estimation of the parameter m when the function $g(t) = m\, h(t)$ with a known function $h(.)$ satisfying $h(0) = 0$ and unknown parameter m. We consider the

Statistical Inference for Fractional Diffusion Processes B. L. S. Prakasa Rao

special case of the fractional Ornstein–Uhlenbeck-type process with Hurst index $H \in (\frac{1}{2}, 1)$ in more detail. Hu (2001) studied the prediction and translation problems for fBm using fractional calculus methods. However, our approach to the problem is via the techniques developed by Kleptsyna *et al.* (2000a). Norros *et al.* (1999) considered the case of constant drift or equivalently the case when $\tilde{Y}(t) = Y(t) + mt$ and derived the MLE of the parameter m when Y is fBm with Hurst index $H \in [\frac{1}{2}, 1)$. Results in this section are due to Prakasa Rao (2005c).

Preliminaries

Suppose a process $\{Y(t), t \geq 0\}$ satisfies the SDE

$$dY(t) = C(t)dt + B(t)dW_t^H, \;\; Y(0) = 0, t \geq 0$$

where $C = \{C(t), t \geq 0\}$ is an $(\mathcal{F}_t)$-adapted process, $B(t)$ is a known nonzero, non-random function and W^H is standard fBm with known Hurst index H. Consider now the process

$$\tilde{Y}(t) = Y(t) + g(t), \;\; t \geq 0$$

where $g(.)$ is an absolutely continuous function with $g(0) = 0$. Note that the function $g(.)$ is almost everywhere differentiable. Let $g'(t)$ denote the derivative of $g(t)$ wherever it exists and define it to be zero elsewhere. Then the process $\tilde{Y}(t)$ satisfies the integral equation

$$\tilde{Y}(t) = g(t) + \int_0^t C(s)ds + \int_0^t B(s)dW_s^H, \; t \geq 0. \tag{6.1}$$

For convenience, we write the above integral equation in the form of a SDE

$$d\tilde{Y}(t) = (C(t) + g'(t))dt + B(t)dW_t^H, \; \tilde{Y}(0) = 0, t \geq 0 \tag{6.2}$$

driven by fBm W^H. Let

$$\tilde{C}(t) = C(t) + g'(t) \tag{6.3}$$

and define

$$\tilde{Q}_H(t) = \frac{d}{dw_t^H}\int_0^t k_H(t,s)\frac{\tilde{C}(s)}{B(s)}\,ds, \quad Q_H(t) = \frac{d}{dw_t^H}\int_0^t k_H(t,s)\frac{C(s)}{B(s)}\,ds \tag{6.4}$$

for $t \geq 0$. Observe that

$$\begin{aligned}\tilde{Q}_H(t) &= Q_H(t) + \frac{d}{dw_t^H}\int_0^t k_H(t,s)\frac{g'(s)}{B(s)}\,ds \\ &= Q_H(t) + g^*_{H,B}(t) \;\; \text{(say)}.\end{aligned} \tag{6.5}$$

Let

$$Z_t = \int_0^t Q_H(s)dw_s^H + M_t^H,\ t \geq 0$$

or equivalently

$$dZ_t = Q_H(t)dw_t^H + dM_t^H,\ Z(0) = 0, t \geq 0$$

where M^H is the fundamental Gaussian martingale defined in (1.73) of Chapter 1 with $\langle M^H \rangle_t = w_t^H$.

Suppose the function $g(.)$ is such that the sample paths of the process $\tilde{Q}_H$ defined by (6.5) belong P-a.s. to $L^2([0,T], dw^H)$. Define

$$\tilde{\Lambda}_H(T) = \exp\left[-\int_0^T \tilde{Q}_H(t)\, dM_t^H - \frac{1}{2}\int_0^t \tilde{Q}_H^2(t)dw_t^H\right]. \tag{6.6}$$

Suppose that $E(\tilde{\Lambda}_H(T)) = 1$. As an application of Theorem 1.20 of Chapter 1, we get that the measure $\tilde{P}^* = \tilde{\Lambda}_H(T)P$ is a probability measure and the probability measure of the process $\tilde{Y}$ under $\tilde{P}^*$ is the same as that of the process V defined by

$$V_t = \int_0^t B(s)dW_s^H,\ 0 \leq t \leq T.$$

Hence we obtain the following result.

Theorem 6.1: Let $\tilde{P}_T^*$ and P_T^* be the probability measures generated by the processes $\tilde{Y}$ and Y respectively in the interval $[0, T]$. Then the measures are absolutely continuous with respect to each other and the Radon–Nikodym derivative of $\tilde{P}_T^*$ with respect to P_T^* is given by

$$\frac{d\tilde{P}_T^*}{dP_T^*} = \exp\left\{-\int_0^T [\tilde{Q}_H(t) - Q_H(t)]dM_t^H - \frac{1}{2}\int_0^t [\tilde{Q}_H^2(t) - Q_H^2(t)]dw_t^H\right\}. \tag{6.7}$$

Maximum likelihood estimation of translation

Let us now suppose that $g(t) = m\ h(t)$ with $h(0) = 0$ and suppose that the functions $h(.)$ and $B(.)$ are *known* and $h(.)$ is differentiable everywhere but the constant m is *unknown*. The problem is to estimate the parameter m based on the observation $\{\tilde{Y}_t, 0 \leq t \leq T\}$. Observe that

$$\begin{aligned}\tilde{Q}_H(t) &= Q_H(t) + m\left(\frac{d}{dw_t^H}\int_0^t k_H(t,s)\frac{h'(s)}{B(s)}\, ds\right)\\ &= Q_H(t) + m\ h_{H,B}^*(t)\\ &= Q_H(t) + m\ h^{**}(t)\ \text{(say)}.\end{aligned} \tag{6.8}$$

Following the notation used above, it can be seen that

$$\tilde{Z}_t = \int_0^t \tilde{Q}_H(s)dw_s^H + M_t^H, \ t \geq 0. \tag{6.9}$$

Observe that

$$d\tilde{Z}_t = dZ_t + m\ h^{**}(t)dw_t^H, \ t \geq 0.$$

Applying Theorem 6.1, we get

$$\begin{aligned}
\frac{d\tilde{P}_T^*}{dP_T^*} &= \exp\left\{-m\int_0^T h^{**}(t)dM_t^H - \frac{1}{2}\int_0^T [2m\tilde{Q}_H(t)h^{**}(t) - m^2(h^{**}(t))^2]dw_t^H\right\} \\
&= \exp\left\{-m\int_0^T h^{**}(t)(d\tilde{Z}_t - \tilde{Q}_H(t)dw_t^H)\right. \\
&\qquad \left. -\frac{1}{2}\int_0^T [2m\tilde{Q}_H(t)h^{**}(t) - m^2(h^{**}(t))^2]dw_t^H\right\} \\
&= \exp\left\{-m\int_0^T h^{**}(t)d\tilde{Z}_t + \frac{1}{2}m^2\int_0^T (h^{**}(t))^2 dw_t^H\right\}.
\end{aligned} \tag{6.10}$$

Suppose that

$$0 < \int_0^T (h^{**}(t))^2 dw_t^H < \infty$$

for all $T \geq 0$, and

$$0 < \int_0^\infty (h^{**}(t))^2 dw_t^H = \infty. \tag{6.11}$$

Then we obtain that the MLE of m based on the process $\{\tilde{Y}(t), 0 \leq t \leq T\}$ is given by

$$\hat{m}_T = \frac{\int_0^T h^{**}(t)d\tilde{Z}_t}{\int_0^T (h^{**}(t))^2 dw_t^H}. \tag{6.12}$$

Remarks: Observe that the estimator $\hat{m}_T$ does not directly depend on the process $\{C(t)\}$ but through observation of the process $\{\tilde{Z}_t, 0 \leq t \leq T\}$.

Suppose m_0 is the true value of m. Then it follows that

$$\hat{m}_T - m_0 = \frac{\int_0^T h^{**}(t)dZ_t}{\int_0^T (h^{**}(t))^2 dw_t^H} \tag{6.13}$$

$$= \frac{\int_0^T h^{**}(t)dM_t^H}{\int_0^T (h^{**}(t))^2 dw_t^H} + \frac{\int_0^T h^{**}(t)Q_H(t)dw_t^H}{\int_0^T (h^{**}(t))^2 dw_t^H}. \tag{6.14}$$

Since the process M^H is a martingale with quadratic variation w^H, it follows that

$$\frac{\int_0^T h^{**}(t) dM_t^H}{\int_0^T (h^{**}(t))^2 dw_t^H} \to 0$$

a.s. as $T \to \infty$ by the strong law of large numbers for local martingales given in Liptser (1980) (cf. Prakasa Rao (1999b)) under condition (6.11). Hence we have the following result.

Theorem 6.2: Suppose that m_0 is the true value of the parameter m. Further suppose that the following conditions hold:

$$\frac{\int_0^T h^{**}(t) Q_H(t) dw_t^H}{\int_0^T (h^{**}(t))^2 dw_t^H} \to 0$$

as $T \to \infty$ and

$$\int_0^\infty (h^{**}(t))^2 dw_t^H = \infty.$$

Then the estimator $\hat{m}_T$ of m defined by Equation (6.12) is strongly consistent.

Suppose that $h(t) \equiv t$ in the above discussion, which reduces to the constant drift case. Then

$$h^{**}(t) = \frac{d}{dw_t^H} \int_0^t k_H(t,s) \frac{1}{B(s)} ds \tag{6.15}$$

and the corresponding estimator $\hat{m}_T$ for the parameter m can be computed using Equation (6.12) once the function $B(.)$ is known.

Fractional Ornstein–Uhlenbeck-type process

As a special case of the results obtained here, we now consider the problem of estimation of translation for the fractional Ornstein–Uhlenbeck-type process. Suppose the process $\{Y_t, t \geq 0\}$ satisfies the stochastic integral equation

$$Y_t = \theta \int_0^t Y_s ds + \sigma W_t^H, \; t \geq 0 \tag{6.16}$$

or equivalently the SDE

$$dY_t = \theta Y_t dt + \sigma dW_t^H, \; Y_0 = 0, t \geq 0 \tag{6.17}$$

with known Hurst index $H \in [\frac{1}{2}, 1)$. Further suppose that we observe the process $\{\tilde{Y}_t, 0 \leq t \leq T\}$ where $\tilde{Y}_t = Y_t + m\, h(t)$ with $h(0) = 0$. In addition, assume that the function $h(.)$ is *known* and everywhere differentiable and satisfies the condition

$$0 < \int_0^T (h^{**}(t))^2 dw_t^H < \infty,$$

but the parameter m is *unknown*. The problem is to estimate the parameter m based on observation of the process $\{\tilde{Y}_t, 0 \leq t \leq T\}$. Following the results given above, we obtain that

$$\hat{m}_T = \frac{\int_0^T h^{**}(t) d\tilde{Z}_t}{\int_0^T (h^{**}(t))^2 dw_t^H} \tag{6.18}$$

where

$$\begin{aligned} d\tilde{Z}_t &= \tilde{Q}_H(t) dw_t^H + dM_t^H \\ &= (Q_H(t) + mh^{**}(t)) dw_t^H + dM_t^H \\ &= dZ_t + mh^{**}(t) dw_t^H, \end{aligned} \tag{6.19}$$

$$\tilde{Q}_H(t) = \frac{d}{dw_t^H} \int_0^t k_H(t,s) \frac{\theta Y_s + mh'(s)}{\sigma} ds,$$

$$Q_H(t) = \frac{d}{dw_t^H} \int_0^t k_H(t,s) \frac{\theta Y_s}{\sigma} ds \tag{6.20}$$

and

$$h^{**}(t) = \frac{d}{dw_t^H} \int_0^t k_H(t,s) \frac{h'(s)}{\sigma} ds. \tag{6.21}$$

Suppose m_0 is the true value of m. Then it follows that

$$\hat{m}_T - m_0 = \frac{\int_0^T h^{**}(t) dZ_t}{\int_0^T (h^{**}(t))^2 dw_t^H}. \tag{6.22}$$

It is easy to see that the solution of the SDE (6.17) is given by

$$Y_t = \sigma \int_0^t e^{\theta(t-u)} dW_u^H,\ t \geq 0. \tag{6.23}$$

Hence the process $\{Y_t, t \geq 0\}$ with $Y_0 = 0$ is a zero-mean Gaussian process with the covariance function given by

$$\begin{aligned} Cov(Y_t, Y_s) &= \sigma^2 E\left\{\int_0^t e^{\theta(t-u)} dW_u^H \int_0^s e^{\theta(s-v)} dW_v^H\right\} \\ &= \sigma^2 H(2H-1) e^{\theta(t+s)} \int_0^t \int_0^s e^{-\theta u} e^{-\theta v} |u-v|^{2H-2} dv du. \end{aligned} \tag{6.24}$$

This follows from results in Chapter 1 (cf. Pipiras and Taqqu (2000)). From the representation (6.23), we obtain that $\{Z_t, t \geq 0\}$ is a zero-mean Gaussian process. Hence it follows, from the representation given by (6.22), that $\hat{m}_T - m_0$ has a Gaussian distribution with mean zero and variance given by

$$Var(\hat{m}_T) = \frac{E(\int_0^T h^{**}(t)dZ_t)^2}{(\int_0^T (h^{**}(t))^2 dw_t^H)^2} \tag{6.25}$$

and we have the following theorem.

Theorem 6.3: Suppose that the stochastic process $\tilde{Y}_t = Y_t + m\, h(t), t \geq 0$, where Y_t is a fractional Ornstein–Uhlenbeck-type process defined by (6.17). Further suppose that the process $\tilde{Y}_t$ is observed over the interval $[0, T]$ and that the function $h(t)$ is known. Let $\hat{m}_T$ be the MLE of the parameter m based on the observed process and let m_0 be the true value of the parameter m. Then $\hat{m}_T - m_0$ has a Gaussian distribution with mean zero and variance given by (6.25).

It is obvious that the estimator $\hat{m}^T$ is a consistent estimator for m_0 if $Var(\hat{m}^T) \to 0$ as $T \to \infty$. Observe that

$$\int_0^T h^{**}(t)dZ_t = \int_0^T h^{**}(t)Q_H(t)dw_t^H + \int_0^T h^{**}(t)dM_t^H \tag{6.26}$$

and hence

$$\begin{aligned}
E\left(\int_0^T h^{**}(t)dZ_t\right)^2 &= Var\left(\int_0^T h^{**}(t)dZ_t\right) \\
&= Var\left(\int_0^T h^{**}(t)Q_H(t)dw_t^H\right) + Var\left(\int_0^T h^{**}(t)dM_t^H\right) \\
&\quad + 2\, Cov\left(\int_0^T h^{**}(t)Q_H(t)dw_t^H, \int_0^T h^{**}(t)dM_t^H\right) \\
&= Var\left(\int_0^T h^{**}(t)Q_H(t)dw_t^H\right) \\
&\quad + 2\, Cov\left(\int_0^T h^{**}(t)Q_H(t)dw_t^H, \int_0^T h^{**}(t)dM_t^H\right) \\
&\quad + \int_0^T (h^{**}(t))^2 dw_t^H.
\end{aligned} \tag{6.27}$$

Therefore

$$Var(\hat{m}_T) = \frac{Var(\int_0^T h^{**}(t)Q_H(t)dw_t^H) + 2\,Cov(\int_0^T h^{**}(t)Q_H(t)dw_t^H, \int_0^T h^{**}(t)dM_t^H)}{(\int_0^T (h^{**}(t))^2 dw_t^H)^2} + \frac{1}{\int_0^T (h^{**}(t))^2 dw_t^H}. \tag{6.28}$$

Suppose that $h(t) \equiv t$ in the above discussion. Then the problem reduces to the constant drift case. Then

$$h^{**}(t) = \frac{1}{\sigma}\frac{d}{dw_t^H}\int_0^t k_H(t,s)ds = \frac{1}{\sigma} \tag{6.29}$$

and the corresponding estimator $\hat{m}_T$ for the parameter m can be computed using Equation (6.18) once the constant σ is known. In fact

$$\hat{m}_T = \sigma\frac{\tilde{Z}_T}{w_T^H}, \tag{6.30}$$

$$\hat{m}_T - m_0 = \sigma\frac{Z_T}{w_T^H}. \tag{6.31}$$

Furthermore,

$$\begin{aligned}
Var(\hat{m}_T) &= \sigma^2\frac{Var(\int_0^T Q_H(t)dw_t^H) + 2\,Cov(\int_0^T Q_H(t)dw_t^H, M_T^H)}{(w_T^H)^2} + \sigma^2\frac{1}{w_T^H} \\
&\le \sigma^2\frac{E(\int_0^T Q_H(t)dw_t^H)^2 + 2\,\{[Var(\int_0^T Q_H(t)dw_t^H)][Var(M_T^H)]\}^{1/2}}{(w_T^H)^2} + \sigma^2\frac{1}{w_T^H} \\
&\le \sigma^2\frac{E(\int_0^T Q_H^2(t)dw_t^H)w_T^H + 2\,\{[E(\int_0^T Q_H(t)dw_t^H)^2][Var(M_T^H)]\}^{1/2}}{(w_T^H)^2} + \sigma^2\frac{1}{w_T^H} \\
&\le \sigma^2\frac{E(\int_0^T Q_H^2(t)dw_t^H)w_T^H + 2\,\{[E(\int_0^T Q_H^2(t)dw_t^H)w_T^H][Var(M_T^H)]\}^{1/2}}{(w_T^H)^2} + \sigma^2\frac{1}{w_T^H} \\
&\le \sigma^2 D_H\frac{E(\int_0^T Q_H^2(t)dw_t^H) + \{E(\int_0^T Q_H^2(t)dw_t^H)\}^{1/2}}{w_T^H} + \sigma^2\frac{1}{w_T^H}
\end{aligned} \tag{6.32}$$

for some constant $D_H > 0$ from the representation (1.73) in Chapter 1 and an application of the Cauchy–Schwartz inequality and Fubini's theorem. We have also used the observation that

$$Var(M_T^H) = C_H T^{2-2H} = O(w_T^H)$$

for some constant $C_H > 0$. This can be checked from the results in Section 2 or from Theorem 3.1 in Norros *et al.* (1999). If a bound on the term

$$E\left(\int_0^T Q_H^2(t)dw_t^H\right)$$

can be obtained as a function of T, then it is possible to obtain an upper bound on the variance term given above. It is possible to get an explicit expression for

$$\Psi_T(\theta; a) = E\left[\exp\left(-a\int_0^T Q_H^2(t)dw_t^H\right)\right],\ a > 0$$

as given in Proposition 3.2 of Kleptsyna and Le Breton (2002a) and hence

$$E\left(\int_0^T Q_H^2(t)dw_t^H\right) = -\lim_{a\to 0^+}\Psi_T'(\theta : a).$$

It is known from the arguments given in Kleptsyna and Le Breton (2002a) that

$$\int_0^T Q_H^2(t)dw_t^H \to \infty \text{ a.s. as } T \to \infty.$$

However, explicit computation of the expectation defined above seems to be difficult. If

$$E\left(\int_0^T Q_H^2(t)dw_t^H\right) = o(w_T^H)$$

as $T \to \infty$, then we obtain that

$$Var(\hat{m}_T) \to 0 \text{ as } T \to \infty$$

and hence $\hat{m}_T \xrightarrow{p} m_0$ as $T \to \infty$ since $E(\hat{m}_T) = m_0$ for all T. Hence $\hat{m}_T$ is a consistent estimator of m_0 under the above conditions.

An alternate way of viewing Equation (6.31) is by writing it in the form

$$\frac{w_T^H(\hat{m}_T - m_0)}{\sigma} = Z_T = \frac{1}{\sigma}\int_0^T k_H(T,s)dY_s \tag{6.33}$$

or equivalently

$$w_T^H(\hat{m}_T - m_0) = \int_0^T k_H(T,s)dY_s \tag{6.34}$$

which in turn shows that the distribution of the estimator $\hat{m}_T$ is normal with mean m_0 and variance

$$(w_T^H)^{-2}E\left[\int_0^T k_H(T,s)dY_s\right]^2.$$

Remarks: If the parameter $\theta = 0$, then the process $\{Y_t, t \geq 0\}$ reduces to fBm and

$$Var(\hat{m}_T) = \sigma^2 \frac{1}{w_T^H} = \sigma^2 \lambda_H T^{2H-2} \tag{6.35}$$

from (1.72) of Chapter 1 and the definition of the process $\{Q_H(t), t \geq 0\}$. Since the Hurst index $H \in [\frac{1}{2}, 1)$, it follows that

$$Var(\hat{m}_T) \to 0 \ \text{ as } \ T \to \infty. \tag{6.36}$$

Combining this observation with the fact that $E(\hat{m}_T) = m_0$, it follows that $\hat{m}_T \xrightarrow{p} m_0$ as $T \to \infty$. In other words, the estimator $\hat{m}_T$ is a consistent estimator for m_0. A stronger result also follows from the fact that, in case $\theta = 0$,

$$\hat{m}_T - m_0 = \sigma \frac{M_T^H}{w_T^H} \tag{6.37}$$

and the last term tends to zero a.s. as $T \to \infty$ by the strong law of large numbers for martingales (cf. Prakasa Rao (1999b), p. 61) since the quadratic variation of the martingale M^H is w^H and $w_T^H \to \infty$ as $T \to \infty$. The strong consistency of the estimator $\hat{m}_T$, for the case of fBm, was proved earlier in Norros *et al.* (1999).

Statistical inference from sampled data

We have assumed that the stochastic processes under consideration can be observed continuously over a specified time period and that statistical inference was based on either one or many realizations of the process over that time period. In practice, it is virtually impossible to observe a process continuously over any given time period, for example, due to limitations on the precision of the measuring instruments or due to the unavailability of observations at every time point. Hence statistical inference based on sampled data is of great importance (cf. Prakasa Rao (1988)). For a discussion on parametric inference for diffusion-type processes from sampled data, see Prakasa Rao (1999a), Chapter 3. Except for some minor discussions in the case of inference from discrete data on fractional Ornstein–Uhlenbeck-type processes, there is no significant work in this area for fractional diffusion processes. We now discuss a very special result due to Bertin *et al.* (2007) on maximum likelihood estimation based on sampled data.

Suppose $\{W_t, t \geq 0\}$ is a standard Wiener process and let

$$Y_t = \theta t + W_t, \ \ 0 \leq t \leq T.$$

with $Y_{t_0} = Y_0 = 0$. Suppose the process Y is observed at times $t_i, 0 \leq i \leq N-1$, with $\Delta t = t_{j+1} - t_j, 0 \leq j \leq N-1$. Let $Y_{t_i} = Y_i$ for convenience in notation. The problem is to estimate the parameter θ based on the observations $\{Y_i, 0 \leq i \leq N\}$

and study its properties. Observe that

$$Y_{t_{j+1}} - Y_{t_j} = \theta \Delta t + W_{t_{j+1}} - W_{t_j}.$$

The conditional probability density of Y_{j+1} given $Y_1, \ldots, Y_j$ is the same as the conditional probability density of Y_{j+1} given Y_j by the Markov property of the process $\{Y_t, t \geq 0\}$. It is given by

$$f_{Y_{j+1}|Y_j}(y_{j+1}|y_j) = (2\pi \Delta t)^{-1/2} \exp\left[-\frac{1}{2}\frac{(y_{j+1} - y_j - \theta \Delta t)^2}{\Delta t}\right]$$

and the likelihood function corresponding to the observations $y_1, \ldots, y_n$ is

$$\begin{aligned} L(\theta; y_1, \ldots, y_N) &= f_{Y_1}(y_1) \prod f_{Y_{j+1}|Y_1,\ldots,Y_j}(y_{j+1}|y_1, \ldots, y_j) \\ &= f_{Y_1}(y_1) \prod f_{Y_{j+1}|Y_j}(y_{j+1}|y_j) \\ &= (2\pi \Delta t)^{-N/2} \exp\left[-\frac{1}{2}\sum_{j=0}^{N-1} \frac{(y_{j+1} - y_j - \theta \Delta t)^2}{\Delta t}\right]. \quad (6.38) \end{aligned}$$

It is easy to check that the MLE of θ is given by

$$\hat{\theta}_N = \frac{1}{N \Delta t} \sum_{j=0}^{N-1} (Y_{j+1} - Y_j).$$

Hence

$$\hat{\theta}_N - \theta = \frac{1}{N \Delta t} \sum_{j=0}^{N-1} (W_{(j+1)/N} - W_{j/N}).$$

Therefore

$$E|\hat{\theta}_N - \theta|^2 = \frac{1}{N \Delta t}$$

and the last term converges to zero only if

(C) $N \Delta t \to \infty$ as $N \to \infty$.

Hence the estimator $\hat{\theta}_N$ is L^2-consistent if the condition (C) is satisfied. Note that the condition stated in (C) does not hold if $t_j = j/N, 0 \leq j \leq N$. We need to observe the process Y at times such that the time interval in between the times of observations is of the order $\Delta t = N^{-\alpha}$ with $0 < \alpha < 1$ so that the condition (C) holds. Equivalently, if we observe the process at N^α time points with $\alpha > 1$, that is, $T > N^{\alpha - 1}$ with $\Delta t = 1/N$, then the condition (C) holds and the estimator $\hat{\theta}_N$ is L^2-consistent for the parameter θ.

Suppose W^H is fBm with $\frac{1}{2} < H < 1$. Consider the process

$$Y_t = \theta t + W_t^H,\ 0 \leq t \leq T$$

with $Y_{t_0} = Y_0 = 0$. Suppose the process Y is observed at times $t_i = i/N, 0 \leq i \leq N^\alpha - 1$, with $\alpha > 1$. Let $Y_{t_i} = Y_i$ for convenience in notation. The problem is to estimate the parameter θ based on the observations $\{Y_i, 0 \leq i \leq N^\alpha - 1\}$ and study its properties. Observe that

$$Y_{t_{j+1}} - Y_{t_j} = \theta \Delta t + W_{t_{j+1}}^H - W_{t_j}^H.$$

Since the process W^H neither is Markov nor has independent increments, it is not possible to write down the likelihood function in the form of (6.38) as it was done in the case of the Wiener process. Suppose we interpret the function as partial likelihood and write down the partial likelihood function in the form

$$PL(\theta; y_1, \ldots, y_N) = f_{Y_1}(y_1) \prod_{j=0}^{N^\alpha - 1} f_{Y_{j+1}|Y_j}(y_{j+1}|y_j) =$$

$$(2\pi(\Delta t)^{2H})^{-N^\alpha/2} \exp\left[-\frac{1}{2}\sum_{j=0}^{N^\alpha-1} \frac{(y_{j+1} - y_j - \theta \Delta t)^2}{(\Delta t)^{2H}}\right]. \quad (6.39)$$

Maximizing the partial likelihood function, we obtain the pseudo-likelihood estimator for θ as

$$\hat{\theta}_N = \frac{1}{N^\alpha \Delta t} \sum_{j=0}^{N^\alpha - 1} (Y_{j+1} - Y_j).$$

Therefore

$$\hat{\theta}_N - \theta = \frac{N}{N^\alpha} \sum_{j=0}^{N^\alpha - 1} (W_{(j+1)/N}^H - W_{j/N}^H). \quad (6.40)$$

Hence

$$E|\hat{\theta}_N - \theta|^2$$

$$= N^{2-2\alpha} \sum_{i=j=0}^{N^\alpha - 1} E[(W_{(i+1)/N}^H - W_{i/N}^H)(W_{(j+1)/N}^H - W_{j/N}^H)]$$

$$= N^{2-2H-2\alpha} + 2N^{2-2\alpha}N^{-2H} \sum_{i<j} [|i-j+1|^{2H} + |i-j-1|^{2H} - 2|i-j|^{2H}]$$

$$= N^{2-2H-2\alpha} + 2N^{2-2\alpha}N^{-2H}\sum_{k=1}^{N^\alpha}(N^\alpha - k)[|k+1|^{2H} + |k-1|^{2H} - 2k^{2H}]$$

$$= N^{2-2H-2\alpha} + 2N^{2-2\alpha}N^{-2H}O(N^{2H\alpha}). \tag{6.41}$$

From the bound derived above on the mean square error of the estimator $\hat{\theta}_N$, it follows that the estimator $\hat{\theta}_N$ is L^2-consistent for the parameter θ as $N \to \infty$ since $\alpha > 1$.

Let us again consider the model

$$Y_t = \theta t + \sigma W_t^H,\ t \geq 0$$

where $H > \frac{1}{2}$ is known and θ and σ are unknown parameters to be estimated from observations of the process $\{Y_t, t \geq 0\}$ at discrete time instants $t_k = kh, k = 1, 2, \ldots, N$, for some fixed $h > 0$. Then the observation vector is $\mathbf{Y} = (Y_{t_1}, Y_{t_2}, \ldots, Y_{t_N})'$. Hu *et al.* (2009) obtained the MLEs of μ and σ^2 and studied their asymptotic properties. For convenience, let $\mathbf{t} = (h, 2h, \ldots, Nh)'$ and $\mathbf{W}_t^H = (W_h^H, W_{2h}^H, \ldots, W_{Nh}^H)'$. Then the probability density function of the random vector $\mathbf{Y}$ is

$$f(\mathbf{y}) = (2\pi\sigma^2)^{-N/2}|\Gamma_H|^{-1/2}\exp\left[-\frac{1}{2\sigma^2}(\mathbf{y} - \theta\mathbf{t})'\Gamma_H^{-1}(\mathbf{y} - \theta\mathbf{t})\right]$$

where

$$\Gamma_H = ((cov(W_{ih}^H, W_{jh}^H)))_{i,j=1,2,\ldots,N}.$$

Note that

$$cov(W_{ih}^H, W_{jh}^H)) = \frac{1}{2}h^{2H}(i^{2H} + j^{2H} - |i-j|^{2H}),\ i, j = 1, 2, \ldots, N.$$

It is easy to check that the MLEs of θ and σ^2 are given by

$$\hat{\theta}_N = \frac{\mathbf{t}'\Gamma_H^{-1}\mathbf{Y}}{\mathbf{t}'\Gamma_H^{-1}\mathbf{t}} \tag{6.42}$$

and

$$\hat{\sigma}_N^2 = \frac{1}{N}\frac{(\mathbf{Y}'\Gamma_H^{-1}\mathbf{Y})(\mathbf{t}'\Gamma_H^{-1}\mathbf{t}) - (\mathbf{t}'\Gamma_H^{-1}\mathbf{Y})^2}{\mathbf{t}'\Gamma_H^{-1}\mathbf{t}}. \tag{6.43}$$

The following result is due to Hu *et al.* (2009).

Theorem 6.4: The estimators $\hat{\theta}_N$ and $\hat{\sigma}_N^2$ are strongly consistent as $N \to \infty$. Furthermore,

$$\sqrt{\mathbf{t}'\Gamma_H^{-1}\mathbf{t}}\ (\hat{\theta}_N - \theta) \xrightarrow{\mathcal{L}} N(0, \sigma^2) \tag{6.44}$$

and

$$\sqrt{\frac{N}{2}}(\hat{\sigma}_N^2 - \sigma^2) \xrightarrow{\mathcal{L}} N(0, \sigma^4) \tag{6.45}$$

as $N \to \infty$.

For a proof of the asymptotic normality of the estimator $\hat{\sigma}_N^2$, Hu *et al.* (2009) use the central limit theorem for stochastic integrals proved in Nualart and Ortiz (2008) using Malliavin calculus.

6.3 Parametric inference for SDEs with delay governed by fBm

Gushchin and Kuchler (1999) investigated asymptotic inference for linear SDEs with a time delay of the type

$$dX(t) = (aX(t) + bX(t-1))dt + dW_t,\ t \geq 0$$

driven by standard Brownian motion $\{W_t, t \geq 0\}$ with the initial condition $X(t) = X_0(t), -1 \leq t \leq 0$, where $X_0(t)$ is a continuous process independent of W. They investigated the asymptotic properties of the MLE of the parameter $\theta = (a, b)$. They showed that the asymptotic behavior of the MLE depends on the ranges of the values of a and b.

We now consider the linear SDE

$$dX(t) = (aX(t) + bX(t-1))dt + dW_t^H, t \geq 0$$

with a time delay driven by fBm $\{W_t^H, t \geq 0\}$. We investigate the asymptotic properties of the MLE of the parameter $\theta = (a, b)$.

Maximum likelihood estimation

Let us consider the SDE

$$dX(t) = (aX(t) + bX(t-1))dt + dW_t^H,\ t \geq 0 \tag{6.46}$$

where $\theta = (a, b) \in \Theta \subset R^2$ and $W = \{W_t^H, t \geq 0\}$ is fBm with a *known* Hurst parameter H with the initial condition $X(t) = X_0(t), t \in [-1, 0]$, where $X_0(.)$ is a continuous Gaussian stochastic process independent of W^H. In other words, $X = \{X_t, t \geq 0\}$ is a stochastic process satisfying the stochastic integral equation

$$\begin{aligned} X(t) &= X(0) + \int_0^t [aX(s) + bX(s-1)]ds + W_t^H,\ t \geq 0, \\ X(t) &= X_0(t), -1 \leq t \leq 0. \end{aligned} \tag{6.47}$$

Let

$$C(\theta, t) = aX(t) + \ bX(t-1),\ t \geq 0 \tag{6.48}$$

and assume that the sample paths of the process $\{C(\theta, t)\}, t \geq 0$, are smooth enough so that the process

$$Q_{H,\theta}(t) = \frac{d}{dw_t^H} \int_0^t k_H(t, s) C(\theta, s) ds,\ t \geq 0 \tag{6.49}$$

is well defined where w_t^H and $k_H(t, s)$ are as defined in (1.72) and (1.70) of Chapter 1 respectively. Suppose the sample paths of the process $\{Q_{H,\theta}, 0 \leq t \leq T\}$ belong a.s. to $L^2([0, T], dw_t^H)$. Define

$$Z_t = \int_0^t k_H(t, s) dX_s,\ t \geq 0. \tag{6.50}$$

Then the process $Z = \{Z_t, t \geq 0\}$ is an $(\mathcal{F}_t)$-semimartingale with the decomposition

$$Z_t = \int_0^t Q_{H,\theta}(s) dw_s^H + M_t^H \tag{6.51}$$

where M^H is the fundamental martingale defined by (1.73) of Chapter 1 and the process X admits the representation

$$\begin{aligned} X(t) &= X(0) + \int_0^t K_H(t, s) dZ_s,\ t \geq 0 \\ X(t) &= X_0(t), -1 \leq t \leq 0 \end{aligned} \tag{6.52}$$

where the function K_H is as defined by (1.75) of Chapter 1 with $f \equiv 1$. Let P_T^θ be the measure induced by the process $\{X_t, -1 \leq t \leq T\}$ on $C[-1, T]$ when θ is the true parameter conditional on $X(t) = X_0(t), -1 \leq t \leq 0$. Following Theorem 1.20 in Chapter 1, we get that the Radon–Nikodym derivative of P_T^θ with respect to $P_T^{(0,0)}$ is given by

$$\frac{dP_T^\theta}{dP_T^{(0,0)}} = \exp\left[\int_0^T Q_{H,\theta}(s) dZ_s - \frac{1}{2} \int_0^T Q_{H,\theta}^2(s) dw_s^H\right]. \tag{6.53}$$

We now consider the problem of estimation of the parameter $\theta = (a, b)$ based on observation of the process $X = \{X_t, 0 \leq t \leq T\}$ conditional on $X(t) = X_0(t), -1 \leq t \leq 0$, and study its asymptotic properties as $T \to \infty$.

Let $L_T(\theta)$ denote the Radon–Nikodym derivative $dP_T^\theta / dP_T^{(0,0)}$. The MLE is defined by the relation

$$L_T(\hat{\theta}_T) = \sup_{\theta \in \Theta} L_T(\theta). \tag{6.54}$$

We assume that there exists a measurable MLE. Sufficient conditions can be given for the existence of such an estimator (cf. Lemma 3.1.2 in Prakasa Rao (1987)). Note that

$$\begin{aligned} Q_{H,\theta}(t) &= \frac{d}{dw_t^H}\int_0^t k_H(t,s)C(\theta,s)ds \\ &= \frac{d}{dw_t^H}\int_0^t k_H(t,s)aX(s)ds + \frac{d}{dw_t^H}\int_0^t k_H(t,s)bX(s-1)ds \\ &= aJ_1(t) + bJ_2(t) \ \ \text{(say).} \end{aligned} \tag{6.55}$$

Then

$$\log L_T(\theta) = \int_0^T (aJ_1(t)+bJ_2(t))dZ_t - \frac{1}{2}\int_0^T (aJ_1(t)+bJ_2(t))^2 dw_t^H \tag{6.56}$$

and the likelihood equations are given by

$$\int_0^T J_1(t)dZ_t = a\int_0^T J_1^2(t)dw_t^H + b\int_0^T J_1(t)J_2(t)dw_t^H \tag{6.57}$$

and

$$\int_0^T J_2(t)dZ_t = b\int_0^T J_2^2(t)dw_t^H + a\int_0^T J_1(t)J_2(t)dw_t^H. \tag{6.58}$$

Solving the above equations, we obtain that the MLE $\hat{\theta}_T$ of $\theta = (a,b)'$ is given by

$$\hat{\theta}_T = (I_T^0)^{-1}V_T^0 \tag{6.59}$$

where

$$V_T^0 = \left(\int_0^T J_1(t)dZ_t, \int_0^T J_2(t)dZ_t\right) \tag{6.60}$$

and

$$I_T^0 = ((I_{ij})) \tag{6.61}$$

is the observed Fisher information matrix with

$$I_{ii} = \int_0^T J_i^2(t)dw_t^H,\ i = 1,2 \tag{6.62}$$

and

$$I_{12} = I_{21} = \int_0^T J_1(t)J_2(t)dw_t^H. \tag{6.63}$$

We can write the log-likelihood function in the form

$$\log \frac{dP_T^{\theta}}{dP_T^{(0,0)}} = \theta' V_T^0 - \frac{1}{2}\theta' I_T^0 \theta,\ \theta \in R^2. \tag{6.64}$$

Let $\theta_0 = (a, b)' \in R^2$ be arbitrary but fixed. Let $\theta = \theta_0 + \phi_T \gamma$ where $\gamma = (\alpha, \beta)' \in R^2$ and $\phi_T = \phi_T(\theta_0)$ is a normalizing matrix with $||\phi_T|| \to 0$ as $T \to \infty$. It is easy to see that

$$\log \frac{dP_T^{\theta}}{dP_T^{\theta_0}} = \gamma' V_T - \frac{1}{2}\gamma' I_T \gamma \tag{6.65}$$

where

$$V_T' = \left(\int_0^T J_1(t) dM_t^H, \int_0^T J_2(t) dM_t^H \right) \phi_T \tag{6.66}$$

and

$$I_T = \phi_T' I_T^0 \phi_T. \tag{6.67}$$

For linear SDEs with a time delay driven by the standard Wiener process, Gushchin and Kuchler (1999) discussed different conditions under which the family of measures $\{P_T^{\theta}\}$ is locally asymptotically normal (LAN) or locally asymptotically mixed normal (LAMN) or in general locally asymptotically quadratic (LAQ). For a discussion of these concepts, see Prakasa Rao (1999b), Chapter 6.

In view of the representation (6.65) for the log-likelihood ratio process, the family of measures $\{P_T^{\theta}\}$ is LAQ at θ_0 if we can choose the normalizing matrix $\phi_T(\theta_0)$ in such a way that (i) the vectors V_T and I_T are bounded in probability as $T \to \infty$; (ii) if (V_{T_n}, I_{T_n}) converges in distribution to a limit (V_∞, I_∞) for a subsequence $T_n \to \infty$, then

$$E\left(\exp\left(\gamma' V_\infty - \frac{1}{2} I_\infty \gamma\right)\right) = 1$$

for every $\gamma \in R^2$; and (iii) if I_{T_n} converges in distribution to a limit I_∞ for a subsequence $T_n \to \infty$, then I_∞ is a.s. positive definite. The family of measures is LAMN at θ_0 if (V_T, I_T) converges in distribution to $(I_\infty^{1/2} Z, I_\infty)$ as $T \to \infty$ where the matrix I_∞ is a.s. positive definite and Z is a standard Gaussian vector independent of I_∞. If, in addition, I_∞ is non random, then the family of measures is LAN at θ_0. For the case $b = 0$, the process $X(t)$ reduces to the fractional Ornstein–Uhlenbeck-type process. Strong consistency of the MLE was proved for such a process in Chapter 3 (cf. Kleptsyna and Le Breton (2002a)). Properties such as the strong consistency and the existence of the limiting distribution of the MLE for this process, as well as for more general processes governed by linear SDEs driven by fBm, were studied in Chapter 2 and in Chapter 3 (cf. Prakasa Rao (2003, 2005a)). The results discussed in this section are from Prakasa Rao (2008b).

Suppose we are able to obtain a normalizing matrix ϕ_T such that $||\phi_T|| \to 0$ as $T \to \infty$ and

$$(V_T, I_T) \xrightarrow{\mathcal{L}} (V_\infty, I_\infty)$$

as $T \to \infty$. Then we have

$$\phi_T^{-1}(\hat{\theta}_T - \theta_0) \xrightarrow{\mathcal{L}} I_\infty^{-1} V_\infty$$

which shows the asymptotic behavior of the MLE $\hat{\theta}_T$ as $T \to \infty$. If the family of measures $\{P_T^\theta\}$ is LAMN, then the local asymptotic minimax bound holds for any arbitrary estimator $\tilde{\theta}_T$ of θ and is given by

$$\lim_{r\to\infty} \liminf_{T\to\infty} \sup_{||\phi_T^{-1}(\theta-\theta_0)||<r} E_\theta[\ell(\phi_T^{-1}(\tilde{\theta}_T - \theta))] \geq E[\ell(I_\infty^{-1} V_\infty)]$$

$$= E[\ell(I_\infty^{-1/2}\mathbf{Z})] \quad (6.68)$$

where $\mathbf{Z}$ is a bivariate vector with independent components with standard Gaussian distributions and $\ell : R^2 \to [0, \infty)$ is a bowl-shaped loss function. The MLE is asymptotically efficient in the sense that the Hajek–Le Cam lower bound obtained above is achieved by the MLE $\hat{\theta}_T$. These results are consequences of the LAMN property for the family of measures $\{P_T^\theta\}$. We will discuss sufficient conditions for LAMN later in this section.

A representation for the solution of (6.46)

Let us again consider the SDE

$$dX(t) = (aX(t) + bX(t-1))dt + dW_t^H, \ t \geq 0 \quad (6.69)$$

where $\theta = (a, b) \in \Theta \subset R^2$ and $W = \{W_t^H, t \geq 0\}$ is fBm with Hurst parameter H with the initial condition $X(t) = X_0(t), t \in [-1, 0]$, where $X_0(.)$ is a continuous Gaussian stochastic process independent of W^H. Observe that the process $\{W_t^H, t \geq 0\}$ is a process with stationary increments. Applying the results in Mohammed and Scheutzow (1990), we find that there exists a unique solution $X = \{X(t), t \geq -1\}$ of Equation (6.69) and it can be represented in the form

$$X(t) = x_0(t)X_0(0) + b\int_{-1}^{0} x_0(t-s-1)X_0(s)ds + \int_0^t x_0(t-s)dW_s^H, \ t \geq 0. \quad (6.70)$$

This process has continuous sample paths for $t \geq 0$ a.s. and, conditionally on X_0, the process X is a Gaussian process. Furthermore, the function $x_0(.)$, defined for

$t \geq -1$, is the *fundamental solution* of the differential equation

$$\frac{dx(t)}{dt} = ax(t) + bx(t-1),\ t > 0 \tag{6.71}$$

subject to the conditions $x(0) = 1, x(t) = 0, t \in [-1, 0)$.

Consider the characteristic equation

$$\lambda - a - be^{-\lambda} = 0 \tag{6.72}$$

of the above differential equation. A complex number λ is a solution of (6.72) if and only if the function $e^{\lambda t}$ is a solution of the differential equation

$$\frac{dx(t)}{dt} = ax(t) + bx(t-1),\ t \geq 0.$$

Let Λ be the set of solutions of Equation (6.72). Define

$$v_0 = \max\{Re\lambda|\lambda \in \Lambda\}$$

and

$$v_1 = \max\{Re\lambda|\lambda \in \Lambda, Re\lambda < v_0\}.$$

A complete discussion of the existence and representation of the fundamental solution $x_0(t)$ of (6.71) is given in Lemma 1.1 and the following discussion in Gushchin and Kuchler (1999). For the class of linear SDEs driven by the standard Wiener process, Gushchin and Kuchler (1999) have proved that the corresponding family of measures $\{P_T^\theta\}$ form (i) a LAN family if $v_0 < 0$, (ii) a LAQ family if $v_0 = 0$, and (iii) a LAMN family if $v_0 > 0$, $v_0 \in \Lambda$ and $v_1 < 0$ or $v_1 > 0$ and $v_1 \in \Lambda$ for suitable a and b.

Local asymptotic mixed normality

Observe that the processes

$$R_i(T) = \int_0^T J_i(t)dM_t^H,\ i = 1, 2 \tag{6.73}$$

are zero-mean local martingales with the quadratic covariation processes

$$\langle R_m, R_n\rangle_T = \int_0^T J_m(t)J_n(t)dw_t^H,\ 1 \leq m, n \leq 2. \tag{6.74}$$

Let

$$R'_T = (R_1(T), R_2(T)). \tag{6.75}$$

Let $\{\langle R, R\rangle_t, t \geq 0\}$ be the matrix of covariate processes defined above. Suppose that there exist positive functions $Q_i(t) \to 0$ as $t \to \infty$ possibly random such that

$$Q_t \langle R, R\rangle_t Q_t \xrightarrow{p} \eta^2$$

as $t \to \infty$ where η^2 is a random positive-definite symmetric matrix with probability one and Q_t is a diagonal matrix of order 2 with diagonal elements $Q_1(t)$ and $Q_2(t)$. Applying a modified version of the multidimensional version of the central limit theorem for continuous local martingales (cf. Theorem 1.49 and Remark 1.47 in Prakasa Rao (1999b)) to take into account possibly different norming factors for different components $R_i(t), i = 1, 2$ (following the techniques in Theorem A.1 of Sorensen (1991)), we obtain that

$$(Q_T R_T, Q_T \langle R, R\rangle_T Q_T) \xrightarrow{\mathcal{L}} (\mathbf{Z}, \eta^{\mathbf{2}})$$

as $T \to \infty$, where $\mathbf{Z}'$ is a bivariate random vector with the characteristic function

$$\phi(u_1, u_2) = E\left(\exp\left(-\frac{1}{2}\mathbf{u}'\eta^2\mathbf{u}\right)\right),$$

where $\mathbf{u} = (u_1, u_2)'$. From the representation (6.53) and the above observations, we obtain the following result leading to sufficient conditions for the LAMN property of the family of measures $\{P_T^\theta\}$ (cf. Prakasa Rao (1999b), p. 271). No further proof or additional arguments are needed to prove the LAMN property of the family of measures $\{P_T^\theta\}$ from the results in Prakasa Rao (1999b, p. 271).

Theorem 6.5: Suppose the parameters a and b are such that there exists a matrix norming function $Q_t \to 0$ as $t \to \infty$ as described above such that

$$Q_t \langle R, R\rangle_t Q_t \xrightarrow{p} \eta^2$$

as $t \to \infty$, where η^2 is a random positive-definite symmetric matrix with probability one. Then the family of measures $\{P_T^\theta\}$ forms a LAMN family.

Remarks: If the matrix η^2 is non-random, then the family of measures $\{P_T^\theta\}$ forms a LAN family. We conjecture that, in general, the family is (i) LAN if $v_0 < 0$ with the norming diagonal matrix with diagonal elements $(T^{-1/2}, T^{-1/2})$; (ii) LAMN if $v_0 > 0$ and $v_1 < 0$ with the norming diagonal matrix with diagonal elements $(e^{-v_0 T}, T^{-1/2})$; and (iii) LAMN if $v_0 > 0, v_1 > 0$ and $v_1 \in \Lambda$ with the norming diagonal matrix with diagonal elements $(e^{-v_0 T}, e^{-v_1 T})$. This conjecture is supported by the results obtained by Gushchin and Kuchler (1999) for linear SDEs with a time delay driven by the Wiener process and by the results in Kleptsyna and Le Breton (2002a) for the fractional Ornstein–Uhlenbeck-type process (the case $b = 0$) which implies that $v_0 = a$. In order to check this conjecture, one method is to obtain the moment generating function of the matrix

R_T explicitly using the methods developed in Kleptsyna and Le Breton (2002a) and then study the asymptotic behavior of the matrix $\langle R, R \rangle_T$ under different conditions on the parameters a and b. It is clear that the asymptotic behavior of the log-likelihood ratio depends on the norming function Q_T which in turn depends on the asymptotic behavior of the random variables

$$\int_0^T J_m(t) J_n(t) dw_t^H, \ 1 \leq m, n \leq 2.$$

6.4 Parametric estimation for linear system of SDEs driven by fBms with different Hurst indices

Geometric Brownian motion has been widely used for modeling the fluctuation of share prices in the stock market and geometric fBm, that is, a process governed by a SDE of the type

$$dX(t) = \theta X(t)dt + \sigma_1 X(t)dW^h(t), \ X(0) = x_0 \in R, \ 0 \leq t \leq T, \tag{6.76}$$

has also been studied for modeling fluctuations of share prices in the stock market in the presence of long-range dependence. In the present scenario where the fluctuations of share prices in one country are influenced by the same in another country or within the same country from different regions, it is reasonable to model the share prices by a system of SDEs driven by noise components coming from different environments which could be dependent or independent.

We now discuss estimation of the trend for a linear system of SDEs and specialize the results to a linear system of geometric fBm later in this section. The results obtained in this section are from Prakasa Rao (2008a).

General case

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space and $\{\mathcal{F}_t, t \geq 0\}$ be a filtration defined on the same. Consider the linear stochastic differential system

$$\begin{aligned} dX(t) &= \theta\, a_1(t, X(t))dt + b_1(t, X(t))dW^H(t), \ X(0) = x_0 \in R, \ 0 \leq t \leq T_1 \\ dY(t) &= \theta\, a_2(t, Y(t))dt + b_2(t, Y(t))dW^h(t), \ Y(0) = y_0 \in R, \ 0 \leq t \leq T_2 \end{aligned} \tag{6.77}$$

where $\theta \in \Theta \subset R - \{0\}$. The functions $b_1(.,.)$ and $b_2(.,.)$ are assumed to be known and nonvanishing. We further assume that the functions $a_1(.,.)$ and $a_2(.,.)$ are also known and that the fBms $\{W^h(t), 0 \leq t \leq T\}$ and $\{W^H(t), 0 \leq t \leq T\}$ with known Hurst indices $h \in [\frac{1}{2}, 1)$ and $H \in [\frac{1}{2}, 1)$ respectively are independent and adapted to the filtration $\{\mathcal{F}_t, t \geq 0\}$. This can be achieved if necessary by choosing $\mathcal{F}_t$ to be the σ-algebra generated by the family $\{W^h(s), 0 \leq s \leq t; W^H(s), 0 \leq s \leq t\}$. We use path wise construction of the stochastic integral with respect to the fBm discussed in Zahle (1998). Suppose the system defined by (6.77) has a unique pathwise solution $\{X(s), 0 \leq s \leq T_1; Y(s), 0 \leq s \leq T_2\}$. Sufficient conditions for the existence and uniqueness of the solution are given in Nualart and Rascanu (2002) (cf. Mishura and Rudomino-Dusyatska (2004)). In addition to these conditions, We assume that $a_1(t, X(t))/b_1(t, X(t))$ is Lebesgue integrable on $[0, T_1]$ and $a_2(t, Y(t))/b_2(t, Y(t))$ is Lebesgue integrable on $[0, T_2]$. Let P_t^X be the measure generated by the process X on $[0, t]$ and P_t^Y be the measure generated by the process Y on $[0, t]$. We will now calculate the Radon–Nikodym derivative for probability measures Q on $(\Omega, \mathcal{F})$ such that P_t^X is equivalent to $Q_t^X, 0 \leq t \leq T_1$, and the process X has zero drift under the measure Q. Let

$$\phi_t \equiv \psi(t, X(t)) = \frac{a_1(t, X(t))}{b_1(t, X(t))}.$$

Define, for $0 < s < t \leq T_1$,

$$k_H(t,s) = k_H^{-1} s^{(1/2)-H}(t-s)^{(1/2)-H}, \quad 0 \leq s \leq t$$
$$= 0 \quad \text{otherwise}$$

where

$$k_H = 2H\Gamma\left(\frac{3}{2} - H\right)\Gamma\left(H + \frac{1}{2}\right). \tag{6.78}$$

Suppose $\psi(t, x) \in C^1[0, T_1] \cap C^2(R)$. Then, by Lemma 1 in Mishura and Rudomino-Dusyatska (2004), there exists another $\mathcal{F}_t$-predictable process $\{\delta_s, 0 \leq s \leq T_1\}$ such that

$$\int_0^t \delta_s ds < \infty \text{ a.s. } [P],\ 0 \leq t \leq T_1 \tag{6.79}$$

and

$$\int_0^t k_H(t,s)\phi_s ds = \int_0^t \delta_s ds,\ 0 \leq t \leq T_1. \tag{6.80}$$

Norros *et al.* (1999) proved that

$$\int_0^t k_H(t,s) dW_s^H = \int_0^t s^{\frac{1}{2}-H} d\tilde{W}_s,\ 0 \leq t \leq T_1 \tag{6.81}$$

where the stochastic integral on the left exists as a path wise integral with respect to fBm W^H,

$$\tilde{W}_t = \int_0^t s^{H-\frac{1}{2}}\, d\tilde{M}_s,\ 0 \leq t \leq T_1$$

and

$$\tilde{M}_t = \int_0^t k_H(t,s)dW_s^H,\ 0 \le t \le T_1.$$

Furthermore, the process $\{\tilde{W}_s, 0 \le t \le T_1\}$ is a standard Wiener process. Suppose that

$$E\left(\int_0^{T_1} s^{2H-1}\delta_s^2 ds\right) < \infty. \tag{6.82}$$

Define

$$\tilde{L}_t = \int_0^t s^{H-\frac{1}{2}}\delta_s d\tilde{W}_s,\ 0 \le t \le T_1.$$

Under the conditions stated above, the process $\{\tilde{L}_t, \mathcal{F}_t, 0 \le t \le T_1\}$ is a square integrable martingale. Suppose the martingale $\{\tilde{L}_t, \mathcal{F}_t, 0 \le t \le T_1\}$ satisfies the condition

$$E\left[\exp\left(\tilde{L}_t - \frac{1}{2}\langle\tilde{L}\rangle_t\right)\right] = 1,\ 0 \le t \le T_1.$$

Then it is known that the process

$$B_t^H = W_t^H - \int_0^t \phi_s ds,\ 0 \le t \le T_1$$

is fBm with respect to the probability measure Q_H defined on $(\Omega, \mathcal{F})$ by

$$\frac{dQ_H}{dP}\bigg|_{\mathcal{F}_t} = \exp\left(\tilde{L}_t - \frac{1}{2}\langle\tilde{L}\rangle_t\right),\ 0 \le t \le T_1.$$

Note that

$$\frac{dQ_H}{dP}\bigg|_{\mathcal{F}_t} = \exp\left(\int_0^t s^{H-\frac{1}{2}}\delta_s d\tilde{W}_s - \frac{1}{2}\int_0^t s^{2H-1}\delta_s^2 ds\right),\ 0 \le t \le T_1.$$

In analogy with the above discussion, we construct another probability measure Q_h defined on $(\Omega, \mathcal{F})$ such that

$$\frac{dQ_h}{dP}\bigg|_{\mathcal{F}_t} = \exp\left(\int_0^t s^{h-\frac{1}{2}}\psi_s d\hat{W}_s - \frac{1}{2}\int_0^t s^{2h-1}\psi_s^2\, ds\right),\ 0 \le t \le T_2.$$

Here $\hat{W}$ is the Wiener process corresponding to fBm W^h, and $\{\psi_t, 0 \le t \le T_2\}$ and $\{\eta_t, 0 \le t \le T_2\}$ are processes such that

$$\int_0^t k_h(t,s)\eta_s ds = \int_0^t \psi_s ds,\ 0 \le t \le T_2. \tag{6.83}$$

Observe that

$$B_t^h = W_t^h - \int_0^t \eta_s ds,\ 0 \le t \le T_2$$

is fBm with respect to the probability measure Q_h defined by

$$\frac{dQ_h}{dP}\bigg|_{\mathcal{F}_t} = \exp\left(\int_0^t s^{h-\frac{1}{2}}\psi_s d\hat{W}_s - \frac{1}{2}\int_0^t s^{2h-1}\psi_s^2 ds\right),\ 0 \le t \le T_2.$$

With respect to the probability measures Q_H and Q_h, the trend term should be zero for the first equation in the system over the interval $[0, T_1]$, and it should be zero for the second equation in the system over the interval $[0, T_2]$. Hence

$$b_1(t, X(t))\phi_t = -\theta\ a_1(t, X(t))$$

and

$$b_2(t, Y(t))\eta_t = -\theta\ a_2(t, Y(t)).$$

Observe that

$$\begin{aligned}\frac{dQ}{dP} &= \frac{dQ_h}{dP}\bigg|_{\mathcal{F}_{T_2}}\frac{dQ_H}{dP}\bigg|_{\mathcal{F}_{T_1}}\\ &= \exp\left(\int_0^{T_2} s^{h-\frac{1}{2}}J_s' d\hat{W}_s - \frac{1}{2}\int_0^{T_2} s^{2h-1}J_s'^2 ds\right)\\ &\quad\times\exp\left(\int_0^{T_1} s^{H-\frac{1}{2}}I_s' d\tilde{W}_s - \frac{1}{2}\int_0^{T_1} s^{2H-1}I_s'^2 ds\right)\\ &= \exp\left[\int_0^{T_2} s^{h-\frac{1}{2}}J_s' d\hat{W}_s + \int_0^{T_1} s^{H-\frac{1}{2}}I_s' d\tilde{W}_s\right.\\ &\quad\left.-\frac{1}{2}\left(\int_0^{T_2} s^{2h-1}J_s'^2 ds + \int_0^{T_1} s^{2H-1}I_s'^2 ds\right)\right].\end{aligned}$$

Note that

$$\begin{aligned}J_t' &= \left(\int_0^t k_h(t,s)\eta_s ds\right)_t'\\ &= \left[\int_0^t k_h(t,s)\left(\frac{-\theta\ a_2(s,Y(s))}{b_2(s,Y(s))}\right)ds\right]_t'\\ &= -\theta\left[\int_0^t k_h(t,s)\left(\frac{a_2(s,Y(s))}{b_2(s,Y(s))}\right)ds\right]_t'\\ &= -\theta\Delta_t\ \ \text{(say)}.\end{aligned}$$

Similarly

$$I'_t = -\theta \left[\int_0^t k_H(t,s) \left(\frac{a_1(s, X(s))}{b_1(s, X(s))} \right) ds \right]'_t$$
$$= -\theta \Gamma_t \quad \text{(say)}.$$

From the above relations, we get

$$\log \frac{dQ}{dP} = \int_0^{T_2} s^{h-\frac{1}{2}} J'_s d\hat{W}_s + \int_0^{T_1} s^{H-\frac{1}{2}} I'_s \, d\tilde{W}_s$$
$$-\frac{1}{2} \left(\int_0^{T_1} s^{2h-1} J'^2_s ds + \int_0^{T_2} s^{2H-1} I'^2_s ds \right)$$
$$= \int_0^{T_2} s^{h-\frac{1}{2}} (-\theta \Delta_s) d\hat{W}_s + \int_0^{T_1} s^{H-\frac{1}{2}} (-\theta \Gamma_s) d\tilde{W}_s$$
$$-\frac{1}{2} \left[\int_0^{T_2} s^{2h-1} (-\theta \Delta_s)^2 \; ds + \int_0^{T_1} s^{2H-1} (-\theta \Gamma_s)^2 \; ds \right].$$

In order to estimate the parameter θ based on observations of the process $\{X(s), 0 \le s \le T_1\}$ and of the process $\{Y(s), 0 \le s \le T_2\}$, we maximize the function dQ/dP or equivalently $\log dQ/dP$. Differentiating the function $\log dQ/dP$ with respect to θ and equating the derivative to zero, we obtain the likelihood equation

$$\theta \left(\int_0^{T_1} s^{2H-1} \Gamma_s^2 ds + \int_0^{T_2} s^{2h-1} \Delta_s^2 ds \right)$$
$$= -\left(\int_0^{T_1} s^{H-\frac{1}{2}} \Gamma_s d\tilde{W}_s + \int_0^{T_2} s^{h-\frac{1}{2}} \Delta_s d\hat{W}_s \right).$$

The solution $\hat{\theta}_{T_1,T_2}$ of this equation is given by

$$\hat{\theta}_{T_1,T_2} = -\frac{\int_0^{T_1} s^{H-\frac{1}{2}} \Gamma_s d\tilde{W}_s + \int_0^{T_2} s^{h-\frac{1}{2}} \Delta_s d\hat{W}_s}{\int_0^{T_1} s^{2H-1} \Gamma_s^2 ds + \int_0^{T_2} s^{2h-1} \Delta_s^2 ds} \tag{6.84}$$

which is the MLE in this general case. It can be checked that (see Eq. (14) in Mishura and Rudomino-Dusyatska (2004))

$$d\tilde{W}_s = dW_s^{(1)} - \theta s^{H-\frac{1}{2}} \Gamma(s) ds$$

and

$$d\hat{W}_s = dW_s^{(2)} - \theta s^{h-\frac{1}{2}} \Delta(s) ds$$

where $W^{(1)}$ and $W^{(2)}$ are independent Wiener processes under the measure Q. Using these relations, it follows that

$$\hat{\theta}_{T_1,T_2} - \theta = -\frac{\int_0^{T_1} s^{H-\frac{1}{2}}\Gamma_s dW_s^{(1)} + \int_0^{T_2} s^{h-\frac{1}{2}}\Delta_s dW_s^{(2)}}{\int_0^{T_1} s^{2H-1}\Gamma_s^2 ds + \int_0^{T_2} s^{2h-1}\Delta_s^2 ds}. \tag{6.85}$$

Let

$$J(T_1, T_2) \equiv \int_0^{T_1} s^{2H-1}\Gamma_s^2 ds + \int_0^{T_2} s^{2h-1}\Delta_s^2 ds.$$

Then we get

$$J(T_1, T_2)(\hat{\theta}_{T_1,T_2} - \theta) = -\left[\int_0^{T_1} s^{H-\frac{1}{2}}\Gamma_s dW_s^{(1)} + \int_0^{T_2} s^{h-\frac{1}{2}}\Delta_s dW_s^{(2)}\right].$$

Let

$$M^{(1)}(t) = \int_0^t s^{H-\frac{1}{2}}\Gamma_s dW_s^{(1)},\ 0 \le t \le T_1$$

$$M^{(2)}(t) = \int_0^t s^{h-\frac{1}{2}}\Delta_s dW_s^{(2)},\ 0 \le t \le T_2$$

and let $\mathbf{M}(T_1, T_2)$ be a diagonal matrix with diagonal elements $(M^{(1)}(T_1), M^{(2)}(T_2))$. Note that $M^{(i)}, i = 1, 2$, are independent continuous local martingales with quadratic variations

$$\langle M^{(1)}, M^{(1)}\rangle_t = \int_0^t s^{2H-1}\Gamma_s^2 ds,\ 0 \le t \le T_1$$

and

$$\langle M^{(2)}, M^{(2)}\rangle_t = \int_0^t s^{2h-1}\Delta_s^2 ds,\ 0 \le t \le T_2.$$

Let $\langle \mathbf{M}\rangle_{t,u}$ be a diagonal matrix with diagonal elements $(\langle M^{(1)}, M^{(1)}\rangle_t, \langle M^{(2)}, M^{(2)}\rangle_u)$. Suppose there exists a vector-valued function (k_{1,t_1}, k_{2,t_2}) such that $k_{i,t_i} > 0, i = 1, 2$, increasing to infinity as $t_i \to \infty$. Let $K_{t,u}$ be a diagonal matrix with diagonal elements $(k_{1,t}, k_{2,u})$. Suppose that

$$K_{T_1,T_2}^{-1}\langle \mathbf{M}\rangle_{(T_1,T_2)} K_{T_1,T_2}^{-1} \xrightarrow{p} \eta^2 \text{ as } T_1 \text{ and } T_2 \to \infty \tag{6.86}$$

where η^2 is a random positive diagonal matrix. Following the results in Theorem 1.50 in Prakasa Rao (1999b) and Theorem A.1 in Sorensen (1991), it follows that

$$K_{T_1,T_2}^{-1}\mathbf{M}(T_1, T_2) \xrightarrow{\mathcal{L}} \mathbf{Z}\eta \text{ as } T_1 \text{ and } T_2 \to \infty$$

where $\mathbf{Z}$ is a diagonal matrix with diagonal elements as independent standard Gaussian random variables and the random matrices $\mathbf{Z}$ and η are independent. As a consequence of this result, we can give a set of sufficient conditions for asymptotic normality of the estimator $\hat{\theta}_{T_1,T_2}$ as $T_1 \to \infty$ and $T_2 \to \infty$. We will now discuss a special case.

Special case

The asymptotic properties of the estimator $\hat{\theta}_{T_1,T_2}$ depend on the processes Δ_s and Γ_s which in turn depend on the functions $a_1(.,.), b_1(.,.), a_2(.,.), b_2(.,.)$, the process $\{X(s), 0 \leq s \leq T_1\}$ and the process $\{Y(s), 0 \leq s \leq T_2\}$. However, if these functions and the process $X(.)$ and $Y(.)$ are such that $\Delta_s = \alpha s^{1-2h}$ and $\Gamma_s = \beta s^{1-2H}$ for some constants α and β, then it follows that

$$\begin{aligned}
\hat{\theta}_{T_1,T_2} - \theta &= -\frac{\int_0^{T_1} s^{H-\frac{1}{2}}\beta s^{1-2H} dW_s^{(1)} + \int_0^{T_2} s^{h-\frac{1}{2}}\alpha s^{1-2h} dW_s^{(2)}}{\int_0^{T_1} s^{2H-1}\beta^2 s^{2-4H} ds + \int_0^{T_2} s^{2h-1}\alpha^2 s^{2-4h} ds} \\
&= -\frac{\beta\int_0^{T_1} s^{\frac{1}{2}-H} dW_s^{(1)} + \alpha\int_0^{T_2} s^{\frac{1}{2}-h} dW_s^{(2)}}{\beta^2\int_0^{T_1} s^{1-2H} ds + \alpha^2\int_0^{T_2} s^{1-2h} ds} \\
&= -\frac{\beta\int_0^{T_1} s^{\frac{1}{2}-H} dW_s^{(1)} + \alpha\int_0^{T_2} s^{\frac{1}{2}-h} dW_s^{(2)}}{\beta^2 T_1^{2-2H}(2-2H)^{-1} + \alpha^2 T_2^{2-2h}(2-2h)^{-1}}.
\end{aligned} \tag{6.87}$$

Since the processes $W^{(1)}$ and $W^{(2)}$ are independent Wiener processes, it easy to see that the estimator $\hat{\theta}_{T_1,T_2}$ has a Gaussian distribution with mean θ and variance

$$[\beta^2 T_1^{2-2H}(2-2H)^{-1} + \alpha^2 T_2^{2-2h}(2-2h)^{-1}]^{-1}.$$

It is clear that the processes $\Delta_s = \alpha s^{1-2h}$ and $\Gamma_s = \beta s^{1-2H}$ hold for some constants α and β if $a_1(.,.) = \beta\, b_1(.,.)$ and $a_2(.,.) = \alpha\, b_2(.,.)$ hold.

Geometric fBm

We now specialize the results discussed earlier to a linear system generated by geometric fBms.

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space. Consider the linear system of SDEs

$$\begin{aligned}
dX(t) &= \theta X(t)dt + \sigma_1 X(t) dW^H(t),\ X(0) = x_0 \in R,\ 0 \leq t \leq T_1, \\
dY(t) &= \theta Y(t)dt + \sigma_2 Y(t) dW^h(t),\ Y(0) = y_0 \in R,\ 0 \leq t \leq T_2,
\end{aligned} \tag{6.88}$$

defined on $(\Omega, \mathcal{F}, P)$ where $\{\theta, \sigma_1, \sigma_2\} \subset R - \{0\}$ and the fBms $\{W^h(t), 0 \leq t \leq T_1\}$ and $\{W^H(t), 0 \leq t \leq T_2\}$ with known Hurst indices $h \in [\frac{1}{2}, 1)$ and $H \in [\frac{1}{2}, 1)$ respectively are independent. We further assume that the parameters σ_1 and σ_2 are known positive constants.

Following the notation introduced earlier in this section, let Q be the product measure of the probability measures Q^h induced on $C[0, T_1]$ and Q^H induced on $C[0, T_2]$. With respect to this probability measure Q, the processes $\{\phi_s\}$ and $\{\eta_s\}$ chosen above should be such that the trend term should be zero in order that the process $\{X(s), 0 \leq s \leq T_1\}$ is a solution of the first SDE in the system given by (6.88) in the interval $[0, T_1]$ and the trend term should be zero in order that the process $\{(Y(s), 0 \leq s \leq T_2\}$ is a solution of the second SDE in the system given by (6.88) in the interval $[0, T_2]$. Hence

$$\sigma_1 \int_0^t \phi_s ds = -\theta t,\ 0 \leq t \leq T_1$$

and

$$\sigma_2 \int_0^t \eta_s ds = -\theta t,\ 0 \leq t \leq T_2.$$

Observe that

$$\begin{aligned}
\frac{dQ}{dP} &= \frac{dQ_H}{dP}\bigg|_{\mathcal{F}_{T_1}} \frac{dQ_h}{dP}\bigg|_{\mathcal{F}_{T_2}} \\
&= \exp\left(\int_0^{T_1} s^{H-\frac{1}{2}} \delta_s d\tilde{W}_s - \frac{1}{2}\int_0^{T_1} s^{2H-1}\delta_s^2 ds\right) \\
&\quad \times \exp\left(\int_0^{T_2} s^{h-\frac{1}{2}} \psi_s d\hat{W}_s - \frac{1}{2}\int_0^{T_2} s^{2h-1}\psi_s^2 ds\right) \\
&= \exp\left[\int_0^{T_2} s^{h-\frac{1}{2}} \psi_s d\hat{W}_s + \int_0^{T_1} s^{H-\frac{1}{2}} \delta_s d\tilde{W}_s \right. \\
&\quad \left. -\frac{1}{2}\left(\int_0^{T_2} s^{2h-1}\psi_s^2 ds + \int_0^{T_1} s^{2H-1}\delta_s^2 ds\right)\right]. \qquad (6.89)
\end{aligned}$$

Note that

$$\delta_t = \left[\int_0^t k_H(t,s)\phi_s ds\right]_t' = -\frac{\theta}{\sigma_1}\left[\int_0^t k_H(t,s) ds\right]_t' \qquad (6.90)$$

and

$$\psi_t = \left[\int_0^t k_h(t,s)\eta_s ds\right]_t' = -\frac{\theta}{\sigma_2}\left[\int_0^t k_h(t,s) ds\right]_t' \qquad (6.91)$$

where g_t' denotes the derivative of g with respect to t evaluated at t. It is easy to see, from the computations given in Norros *et al.* (1999), that

$$\int_0^t k_H(t,s) ds = D_H^2 t^{2-2H}$$

where

$$D_H = \frac{C_H}{2H(2-2H)^{1/2}}$$

and

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

Hence

$$\left[\int_0^t k_H(t,s)ds\right]_t' = D_H^2(2-2H)t^{1-2H}.$$

Let

$$\gamma_H(t) = D_H^2(2-2H)t^{1-2H}.$$

It can be checked that

$$\gamma_H(t) = \frac{\Gamma(\frac{3}{2}-H)}{2H\Gamma(H+\frac{1}{2})\Gamma(2-2H)}t^{1-2H} = J_H t^{1-2H} \text{ (say)}.$$

From the above relations, we get

$$\begin{aligned}\log\frac{dQ}{dP} = &-\frac{\theta}{\sigma_1}\int_0^{T_1} s^{H-\frac{1}{2}}\gamma_H(s)d\tilde{W}_s - \frac{\theta}{\sigma_2}\int_0^{T_2} s^{h-\frac{1}{2}}\gamma_h(s)d\hat{W}_s \\ &-\frac{\theta^2}{2\sigma_1^2}\int_0^{T_1} s^{2H-1}\gamma_H^2(s)ds - \frac{\theta^2}{2\sigma_2^2}\int_0^{T_2} s^{2h-1}\gamma_h^2(s)ds. \quad (6.92)\end{aligned}$$

Estimation

Note that the Radon–Nikodym derivative dQ/dP obtained above is the Radon–Nikodym derivative of the product measure of the probability measure generated by the process $\{X(s), 0 \le s \le T_1\}$ on the space $C[0, T_1]$ and the probability measure generated by the independent process $\{Y(s), 0 \le s \le T_2\}$. In order to estimate the parameter θ based on observation of the process $\{X(s), 0 \le s \le T_1\}$ and of the process $\{Y(s), 0 \le s \le T_2\}$, we maximize the likelihood function dQ/dP or equivalently $\log dQ/dP$. Differentiating the function $\log dQ/dP$ with respect to θ and equating the derivative to zero, we obtain the likelihood equation

$$\theta\left[\frac{J_H^2}{\sigma_1^2}\frac{T_1^{2-2H}}{2-2H} + \frac{J_h^2}{\sigma_2^2}\frac{T_2^{2-2h}}{2-2h}\right] = -\left[\frac{J_H}{\sigma_1}\int_0^{T_1} s^{\frac{1}{2}-H}d\tilde{W}_s + \frac{J_h}{\sigma_2}\int_0^{T_2} s^{\frac{1}{2}-h}d\hat{W}_s\right]$$

which leads to the estimator, namely, the MLE, given by

$$\hat{\theta}_{T_1,T_2} = -\frac{[(J_H/\sigma_1)\int_0^{T_1} s^{\frac{1}{2}-H} d\tilde{W}_s + (J_h/\sigma_2)\int_0^{T_2} s^{\frac{1}{2}-h} d\hat{W}_s]}{\left[\frac{J_H^2}{\sigma_1^2}\frac{T_1^{2-2H}}{2-2H} + \frac{J_h^2}{\sigma_2^2}\frac{T_2^{2-2h}}{2-2h}\right]}.$$

It can be checked that (see Eq. (14) in Mishura and Rudomino-Dusyatska (2004))

$$d\tilde{W}_s = dW_s^{(1)} - \frac{\theta}{\sigma_1} s^{H-\frac{1}{2}} \gamma_H(s) ds$$

and

$$d\hat{W}_s = dW_s^{(2)} - \frac{\theta}{\sigma_2} s^{h-\frac{1}{2}} \gamma_h(s) ds$$

where $W^{(1)}$ and $W^{(2)}$ are independent Wiener processes under the measure Q. Using these relations, it follows that

$$\hat{\theta}_{T_1,T_2} - \theta = -\frac{[(J_H/\sigma_1)\int_0^{T_1} s^{\frac{1}{2}-H} dW_s^{(1)} + (J_h/\sigma_2)\int_0^{T_2} s^{\frac{1}{2}-h} dW_s^{(2)}]}{\left[\frac{J_H^2}{\sigma_1^2}\frac{T_1^{2-2H}}{2-2H} + \frac{J_h^2}{\sigma_2^2}\frac{T_2^{2-2h}}{2-2h}\right]}.$$

In particular, it follows that the estimator $\hat{\theta}_{T_1,T_2} - \theta$ has a Gaussian distribution with mean zero and variance

$$\left[\frac{J_H^2}{\sigma_1^2}\frac{T_1^{2-2H}}{2-2H} + \frac{J_h^2}{\sigma_2^2}\frac{T_2^{2-2h}}{2-2h}\right]^{-1}.$$

Suppose $h \geq H$. Further suppose that we observe the process X governed by the first equation in the system up to time $T_1 = T$ and observe the process Y governed by the second equation in the system up to time $T_2 = T^{(1-H)/(1-h)}$. Then the variance of the MLE is given by

$$\left[\frac{J_H^2}{\sigma_1^2}\frac{T_1^{2-2H}}{2-2H} + \frac{J_h^2}{\sigma_2^2}\frac{T_2^{2-2h}}{2-2h}\right]^{-1}$$

which is of the order $O(T^{2H-2})$. A better estimator with smaller variance can be obtained by suitably choosing $T_1 = T$ and $T_2 = cT^{(1-H)/(1-h)}$ where c is defined by the relation

$$\frac{J_H^2}{\sigma_1^2}\frac{T^{2-2H}}{2-2H} = c^{2-2h}\frac{J_h^2}{\sigma_2^2}\frac{T_2^{2-2h}}{2-2h}.$$

Remarks: The methods of this section can be extended to study the problem of estimation of the parameter θ for more general linear systems of the type

$$dX_i(t) = [\theta a_i(t, X_i(t)) + c_i(t, X_i(t))]dt + b_i(t, X_i(t))dW^{H_i}(t),\ 0 \le t \le T_i, 1 \le i \le n.$$

6.5 Parametric estimation for SDEs driven by mixed fBm

We have discussed estimation problems for SDEs of the form

$$dX(t) = \theta X(t)dt + \sigma_1 X(t)dW^h(t),\ X(0) = x_0 \in R,\ 0 \le t \le T \tag{6.93}$$

for modeling fluctuations of share prices in the stock market in the presence of long-range dependence. It is reasonable to model the share prices by SDEs driven by two or more noise components coming from different environments which could be dependent or independent. Recently there has been interest in studying the problem of estimation of a parameter θ for a mixed Brownian and fractional Brownian model of the type

$$dX_t = \theta X(t)dt + \sigma_1 X(t)dW(t) + \sigma_2 X(t)dW^H(t),\ X(0) = x_0 \in R,\ 0 \le t \le T \tag{6.94}$$

where $\{\theta, \sigma_1, \sigma_2\} \subset R-\{0\}$, W is the standard Wiener process and W^H is fBm with Hurst index H (cf. Rudomino-Dusyatska (2003)). We studied properties of the MLE of the parameter θ for a linear system of SDEs of the type

$$\begin{aligned} dX(t) &= \theta a_1(t, X(t))dt + b_1(t, X(t))dW^H(t),\ X(0) = x_0,\ 0 \le t \le T_1 \\ dY(t) &= \theta a_2(t, X(t))dt + b_2(t, X(t))dW^h(t),\ Y(0) = y_0,\ 0 \le t \le T_2 \end{aligned} \tag{6.95}$$

driven by two fBms possibly with different Hurst indices, in the previous section. We now consider a SDE of the type

$$\begin{aligned} dX(t) &= \theta a(t, X(t))dt + b_1(t, X(t))dW^h(t) + b_2(t, X(t))dW^H(t), \\ X(0) &= x_0 \in R, \end{aligned} \tag{6.96}$$

where $0 \le t \le T$, driven by a mixture of fBms with possibly different Hurst indices and study the properties of a pseudo-likelihood estimator of the trend parameter θ based on the observation $\{X(s), 0 \le s \le T\}$. The results discussed in this section are due to Prakasa Rao (2009).

Preliminaries

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space and $\{\mathcal{F}_t, t \geq 0\}$ be a filtration defined on the same. Consider the SDE

$$dX(t) = \theta a(t, X(t))dt + b_1(t, X(t))dW^h(t) + b_2(t, X(t))dW^H(t) \qquad (6.97)$$

where $X(0) = x_0 \in R$, $0 \leq t \leq T$ and $\theta \in \Theta \subset R - \{0\}$. The functions $b_1(., .)$ and $b_2(., .)$ are assumed to be known and non-vanishing. We further assume that the function $a(., .)$ is also known and that the fBms $\{W^h(t), 0 \leq t \leq T\}$ and $\{W^H(t), 0 \leq t \leq T\}$ with known Hurst indices $h \in [\frac{1}{2}, 1)$ and $H \in [\frac{1}{2}, 1)$ respectively are independent and adapted to the filtration $\{\mathcal{F}_t, t \geq 0\}$. This can be achieved if necessary by choosing $\mathcal{F}_t$ to be the σ-algebra generated by the family $\{W^h(s), 0 \leq s \leq t; W^H(s), 0 \leq s \leq t\}$.

Define

$$\begin{aligned} k_H(t, s) &= k_H^{-1} s^{(1/2)-H}(t-s)^{(1/2)-H}, \quad 0 \leq s \leq t \\ &= 0 \text{ otherwise} \end{aligned}$$

where

$$k_H = 2H\Gamma\left(\frac{3}{2} - H\right)\Gamma\left(H + \frac{1}{2}\right). \qquad (6.98)$$

Suppose there exists a $\mathcal{F}_t$-predictable process $\{\phi_s, s \geq 0\}$ such that

$$\int_0^t k_H(t, s)|\phi_s|ds < \infty \text{ a.s. } [P], \quad t \geq 0 \qquad (6.99)$$

and another $\mathcal{F}_t$-predictable process $\{\delta_s, s \geq 0\}$ such that

$$\int_0^t k_H(t, s)\phi_s ds = \int_0^t \delta_s ds < \infty \text{ a.s. } [P], \quad t \geq 0. \qquad (6.100)$$

Let

$$I_t \equiv \int_0^t k_H(t, s)\phi_s ds = \int_0^t \delta_s ds, \; t \geq 0. \qquad (6.101)$$

Note that

$$I_t' = \delta_t \qquad (6.102)$$

where g_t' denotes the derivative of g with respect to t evaluated at t.

Norros *et al.* (1999) proved that

$$\int_0^t k_H(t, s)dW_s^H = \int_0^t s^{\frac{1}{2}-H} d\hat{W}_s \qquad (6.103)$$

where the stochastic integral on the left exists as a path wise integral with respect to fBm W^H,

$$\hat{W}_t = \int_0^t s^{H-\frac{1}{2}} d\hat{M}_s$$

and

$$\hat{M}_s = \int_0^t k_H(t, s) dW_s^H.$$

Furthermore, the process $\{\hat{W}_s, s \geq 0\}$ is a standard Wiener process. Suppose that

$$E\left(\int_0^t s^{2H-1}\delta_s^2 ds\right) < \infty. \tag{6.104}$$

Define

$$\hat{L}_t = \int_0^t s^{H-\frac{1}{2}}\delta_s d\hat{W}_s,\ t \geq 0.$$

Under the conditions stated above, the process $\{\hat{L}_t, \mathcal{F}_t, t \geq 0\}$ is a square integrable martingale. Suppose the martingale $\{\hat{L}_t, \mathcal{F}_t, t \geq 0\}$ satisfies the condition

$$E\left[\exp\left(\hat{L}_t - \frac{1}{2} < \hat{L} >_t\right)\right] = 1.$$

Then it is known that the process

$$B_t^H = W_t^H - \int_0^t \phi_s ds,\ t \geq 0$$

is fBm with respect to the probability measure Q_H defined by

$$\left.\frac{dQ_H}{dP}\right|_{\mathcal{F}_t} = \exp\left(\hat{L}_t - \frac{1}{2}\langle\hat{L}\rangle_t\right).$$

Note that

$$\left.\frac{dQ_H}{dP}\right|_{\mathcal{F}_t} = \exp\left(\int_0^t s^{H-\frac{1}{2}} I_s' d\hat{W}_s - \frac{1}{2}\int_0^t s^{2H-1} I_s'^2 ds\right).$$

In analogy with the above discussion, we construct another probability measure Q_h such that

$$\left.\frac{dQ_h}{dP}\right|_{\mathcal{F}_t} = \exp\left(\int_0^t s^{h-\frac{1}{2}} J_s' d\tilde{W}_s - \frac{1}{2}\int_0^t s^{2h-1} J_s'^2 ds\right).$$

Here $\tilde{W}$ is the Wiener process corresponding to fBm W^h, and $\{\psi_t, t \geq 0\}$ and $\{\eta_t, t \geq 0\}$ are $\mathcal{F}_t$-predictable processes such that

$$J_t \equiv \int_0^t k_h(t,s)\eta_s ds = \int_0^t \psi_s ds,\ t \geq 0. \tag{6.105}$$

Observe that

$$B_t^h = W_t^h - \int_0^t \eta_s ds,\ t \geq 0$$

is fBm with respect to the probability measure Q_h defined by

$$\frac{dQ_h}{dP}\bigg|_{\mathcal{F}_t} = \exp\left(\int_0^t s^{h-\frac{1}{2}} J_s' d\tilde{W}_s - \frac{1}{2}\int_0^t s^{2h-1} J_s'^2 ds\right).$$

Since the processes W^h and W^H are independent, we define Q to be the product measure of the probability measures Q^h and Q^H. With respect to this probability measure Q, the trend term should be $\theta a(t, X(t))$. Hence

$$b_1(t, X(t))\phi_t + b_2(t, X(t))\eta_t = -\theta a(t, X(t)).$$

Therefore the processes $\{\eta_t, t \geq 0\}$ and $\{\phi_t, t \geq 0\}$ are connected by the relation

$$\eta_t = \frac{-b_1(t, X(t))\phi_t - \theta a(t, X(t))}{b_2(t, X(t))}.$$

Observe that

$$\begin{aligned}
\frac{dQ}{dP}\bigg|_{\mathcal{F}_t} &= \frac{dQ_h}{dP}\bigg|_{\mathcal{F}_t}\frac{dQ_H}{dP}\bigg|_{\mathcal{F}_t} \\
&= \exp\left(\int_0^t s^{h-\frac{1}{2}} J_s' d\tilde{W}_s - \frac{1}{2}\int_0^t s^{2h-1} J_s'^2 ds\right) \\
&\quad \times \exp\left(\int_0^t s^{H-\frac{1}{2}} I_s' d\hat{W}_s - \frac{1}{2}\int_0^t s^{2H-1} I_s'^2 ds\right) \\
&= \exp\left[\int_0^t s^{h-\frac{1}{2}} J_s' d\tilde{W}_s + \int_0^t s^{H-\frac{1}{2}} I_s' d\hat{W}_s \right. \\
&\quad \left. -\frac{1}{2}\int_0^t [s^{2h-1} J_s'^2 + s^{2H-1} I_s'^2] ds\right].
\end{aligned}$$

Note that

$$\begin{aligned}
J'_t &= \left[\int_0^t k_h(t,s)\eta_s ds\right]'_t \\
&= \left[\int_0^t k_h(t,s)\left(\frac{-\theta a(s,X(s)) - b_1(s,X(s))\phi_s}{b_2(s,X(s))}\right) ds\right]'_t \\
&= -\theta\left[\int_0^t k_h(t,s)\left(\frac{a(s,X(s))}{b_2(s,X(s))}\right) ds\right]'_t - \left[\int_0^t k_h(t,s)\left(\frac{b_1(s,X(s))\phi_s}{b_2(s,X(s))}\right) ds\right]'_t \\
&= -\theta\Delta_t - R_t \text{ (say).}
\end{aligned}$$

From the above relations, we get that

$$\begin{aligned}
\log\frac{dQ}{dP}\bigg|_{\mathcal{F}_t} &= \int_0^t s^{h-\frac{1}{2}} J'_s d\tilde{W}_s + \int_0^t s^{H-\frac{1}{2}} I'_s d\hat{W}_s \\
&\quad -\frac{1}{2}\int_0^t [s^{2h-1}J'^2_s + s^{2H-1}I'^2_s]ds \\
&= \int_0^t s^{h-\frac{1}{2}}(-\theta\Delta_s - R_s)d\tilde{W}_s + \int_0^t s^{H-\frac{1}{2}} I'_s d\hat{W}_s \\
&\quad -\frac{1}{2}\int_0^t [s^{2h-1}(-\theta\Delta_s - R_s)^2 + s^{2H-1}I'^2_s]ds.
\end{aligned}$$

Pseudo-likelihood estimation

In order to estimate the parameter θ based on observation of the process $\{X(s), 0 \le s \le T\}$, we maximize the pseudo-likelihood function $dQ/dP|_{\mathcal{F}_T}$ or equivalently $\log dQ/dP|_{\mathcal{F}_T}$. Note that the function $dQ/dP|_{\mathcal{F}_T}$ is not the usual likelihood function based on observation of the process $\{X(s), 0 \le s \le T\}$, but the product of the likelihood functions as if the process $\{X(s), 0 \le s \le T\}$ was driven by independent fBms W^H and W^h with a common trend and possibly different diffusion coefficients separately. Differentiating the pseudo-likelihood function $\log dQ/dP|_{\mathcal{F}_T}$ with respect to θ and equating the derivative to zero, we obtain the pseudo-likelihood equation

$$\theta\int_0^T s^{2h-1}\Delta_s^2 ds + \int_0^T s^{h-\frac{1}{2}}\Delta_s d\tilde{W}_s + \int_0^T s^{2h-1}\Delta_s R_s ds = 0$$

after some simplification.

The solution $\tilde{\theta}_T$ of the pseudo-likelihood equation is given by

$$\begin{aligned}
\tilde{\theta}_T\int_0^T s^{2h-1}\Delta_s^2 ds &= -\left(\int_0^T s^{h-\frac{1}{2}}\Delta_s d\tilde{W}_s + \int_0^T s^{2h-1}\Delta_s R_s ds\right) \\
&= -\left(\int_0^T s^{h-\frac{1}{2}}\Delta_s d\tilde{W}_s + \int_0^T s^{2h-1}\Delta_s(-\theta\Delta_s - \psi_s)ds\right).
\end{aligned}$$

Hence

$$(\tilde{\theta}_T - \theta) \int_0^T s^{2h-1} \Delta_s^2 ds = -\int_0^T s^{h-\frac{1}{2}} \Delta_s d\tilde{W}_s + \int_0^T s^{2h-1} \Delta_s \psi_s ds. \quad (6.106)$$

Special case

The asymptotic properties of the pseudo-likelihood estimator $\tilde{\theta}_T$ depend on the processes Δ_s and ψ_s which in turn depend on the functions $a(.,.), b_1(.,.)$ and $b_2(.,.)$ and the process $X(.), 0 \leq s \leq T$. However, if these functions and the process $X(.)$ are such that $\Delta_s = \gamma s^{1-2h}$ for some constant γ, then the above equation reduces to

$$(\tilde{\theta}_T - \theta) \int_0^T s^{2h-1} \gamma^2 s^{2-4h} ds = -\int_0^T s^{h-\frac{1}{2}} \gamma s^{1-2h} d\tilde{W}_s + \int_0^T s^{2h-1} \gamma s^{1-2h} \psi_s ds \quad (6.107)$$

which implies that

$$\begin{aligned}
\gamma(\tilde{\theta}_T - \theta) \int_0^T s^{1-2h} ds &= -\int_0^T s^{\frac{1}{2}-h} d\tilde{W}_s + \int_0^T \psi_s ds \\
&= -\int_0^T k_h(T,s) dW_s^h + \int_0^T k_h(T,s) \eta_s ds \\
&= -\int_0^T k_h(T,s)(dW_s^h - \eta_s ds) \\
&= -\int_0^T k_h(T,s) dB_s^h.
\end{aligned}$$

Since the random variable

$$\int_0^T k_h(T,s) dB_s^h$$

has the same distribution as the random variable

$$\int_0^t s^{\frac{1}{2}-h} dZ_s$$

where Z_s is a standard Wiener process, it follows that the random variable

$$\gamma \frac{T^{2-2h}}{2-2h}(\tilde{\theta}_T - \theta)$$

has a normal distribution with mean zero and variance $T^{2-2h}/(2-2h)$. This in turn proves that the pseudo-likelihood estimator $\tilde{\theta}_T$ is unbiased, consistent and

$$Var(\tilde{\theta}_T) = \gamma^{-2} \left(\frac{2-2h}{T^{2-2h}} \right).$$

Note that the process

$$G_T = \int_0^T s^{\frac{1}{2}-h} dZ_s,\ T \geq 0$$

is a square integrable martingale with $\langle G \rangle_T = T^{2-2h}/(2-2h) \to \infty$ as $T \to \infty$. Hence

$$\frac{G_T}{\langle G \rangle_T} \to 0$$

a.s. as $T \to \infty$ (Liptser (1980), Prakasa Rao (1999b), Remark 1.46). Hence

$$\tilde{\theta}_T \to \theta$$

a.s. as $T \to \infty$. Therefore the pseudo-likelihood estimator $\tilde{\theta}_T$ is strongly consistent for estimating the parameter θ.

Remarks: It is clear that the process $\Delta_s = \gamma s^{1-2h}$ for some constant γ if the functions $a(.,)$ and $b_1(.,.)$ satisfy the condition $a(.,.) = \gamma\, b_1(.,.)$. A similar estimator $\hat{\theta}_T$ can be obtained if the functions $a(.,.)$ and $b_2(.,)$ are proportional to each other.

General case

In general, it is not possible to find the limiting distribution of the pseudo-likelihood estimator $\tilde{\theta}_T$ unless the functions $a(.,.), b_1(.,.)$ and $b_2(.,.)$ are specified.

Note that

$$\begin{aligned}
\tilde{\theta}_T - \theta &= \frac{-\int_0^T s^{h-\frac{1}{2}} \Delta_s d\tilde{W}_s + \int_0^T s^{2h-1} \Delta_s \psi_s ds}{\int_0^T s^{2h-1} \Delta_s^2 ds} \\
&= \frac{-\int_0^T s^{h-\frac{1}{2}} \Delta_s d\tilde{W}_s}{\int_0^T s^{2h-1} \Delta_s^2 ds} + \frac{\int_0^T s^{2h-1} \Delta_s \psi_s ds}{\int_0^T s^{2h-1} \Delta_s^2 ds} \\
&= -(\alpha_T)(\zeta_T)^{-1} + \beta_T (\zeta_T)^{-1} \ \text{(say)}.
\end{aligned}$$

Since the process $\tilde{W}_s$ is the standard Wiener process, the quadratic variation of the stochastic integral

$$\alpha_T = \int_0^T s^{h-\frac{1}{2}} \Delta_s d\tilde{W}_s$$

is

$$\zeta_T = \int_0^T s^{2h-1} \Delta_s^2 ds.$$

Strong consistency

Theorem 6.6: Suppose that $Z_T \to \infty$ a.s. as $T \to \infty$. Further suppose that

$$\frac{\beta_T}{\zeta_T} \to 0$$

a.s. as $T \to \infty$. Then

$$\tilde{\theta}_T \to \theta$$

a.s. as $T \to \infty$.

Proof: This result follows from the strong law of large numbers for martingales (cf. Liptser (1980), Prakasa Rao (1999b)).

Limiting distribution

Theorem 6.7: Suppose that $\beta_T = o_p(\sqrt{\zeta_T})$. Then

$$\zeta_T^{1/2}(\tilde{\theta}_T - \theta) \xrightarrow{\mathcal{L}} N(0, 1)$$

as $T \to \infty$.

Proof: This result follows from the central limit theorem for local martingales (cf. Prakasa Rao (1999b)).

Remarks:

(i) The conditions in Theorems 6.6 and 6.7 are difficult to verify in general. However, in the case of geometric mixed fBm

$$\begin{aligned} dX(t) &= \theta X(t)dt + \sigma_1 X(t)dW^H(t) \\ &\quad + \sigma_2 X(t)dW^h(t),\ X(0) = x_0,\ 0 \le t \le T \end{aligned} \tag{6.108}$$

these properties can be derived. Applying the Cauchy–Schwartz inequality, we can see that

$$\zeta_T^{1/2}(\tilde{\theta}_T - \theta) = -\alpha_t \zeta_T^{-1/2} + O\left(\sqrt{\int_0^T s^{2h-1}\psi^2(s)ds}\right).$$

(ii) The methods discussed above can be extended to study estimation of the parameter θ for SDEs of the type

$$\begin{aligned} dX(t) &= [\theta\ a(t, X(t)) + c(t, X(t))]dt \\ &\quad + b_1(t, X(t))dW^h(t) + b_2(t, X(t))dW^H(t),\ t \ge 0. \end{aligned} \tag{6.109}$$

(iii) It is not clear under what conditions the probability measure generated by the process $\{X(s), 0 \leq s \leq T\}$ satisfying the SDE (6.94) or (6.96) is absolutely continuous with respect to the probability measure generated by fBm or Brownian motion and, if so, how to obtain a Girsanov-type theorem for the corresponding Radon–Nikodym derivatives. If this problem is solved, then we can discuss the problem of genuine maximum likelihood estimation for the parameter θ instead of the method of pseudo-likelihood estimation presented above.

6.6 Alternate approach for estimation in models driven by fBm

Our approach for estimation in models driven by fBm has been via the fundamental martingale associated with fBm. In some recent papers, Berzin and Leon (2006, 2007) proposed a new method for the simultaneous estimation of parameters σ and H in some models driven by fBm using regression methods. We will briefly discuss their results.

Let $\{W_H(t), t \geq 0\}$ be fBm with Hurst index $H > \frac{1}{2}$. Modifying slightly the notation used earlier, we suppose that the process W_H is a centered Gaussian process with covariance function

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

with $\upsilon_{2H}^2 = [\Gamma(2H+1)\ \sin(\pi H)]^{-1}$.

The stochastic integral with respect to fBm can be defined path wise as the limit of Riemann sums following the work in Lin (1995), Lyons (1994) and Zahle (1998). This allows us to study SDEs of the type

$$dX(t) = \mu(X(t))dt + \sigma(X(t))dW_H(t),\ X(0) = c,\ t \geq 0. \tag{6.110}$$

Examples of such models include

$$dX(t) = \mu dt + \sigma dW_H(t),\ t \geq 0, \tag{6.111}$$

$$dX(t) = \mu\ X(t)dt + \sigma dW_H(t),\ t \geq 0, \tag{6.112}$$

$$dX(t) = \mu\ X(t)dt + \sigma\ X(t)dW_H(t),\ t \geq 0, \tag{6.113}$$

and

$$dX(t) = \mu dt + \sigma\ X(t)dW_H(t),\ t \geq 0 \tag{6.114}$$

with $X(0) = c$. The solutions of the SDEs given above are:

$$X(t) = \sigma\ W_H(t) + \mu\ t + c\ \ \text{(cf. Lin (1995))}; \tag{6.115}$$

$$X(t) = \sigma\ W_H(t) + \exp(\mu t)\left[\sigma\mu\ \left(\int_0^t W_H(s)\exp(-\mu s)ds\right) + c\right]; \tag{6.116}$$

$$X(t) = c\ \exp[\sigma\ W_H(t) + \mu t] \text{ (cf. Klingenhofer and Zahle (1999));} \quad (6.117)$$

and

$$X(t) = \exp(\sigma\ W_H(t)) \left[c + \mu \int_0^t \exp(-\sigma\ W_H(s))ds \right] \quad (6.118)$$

respectively. For any process $Y(.)$ and $\epsilon > 0$, let

$$Y_\epsilon^{(2)}(t) = Y(t + 2\epsilon) - 2\ Y(t + \epsilon) + Y(t).$$

Observe that the process $Y_\epsilon^{(2)}(.)$ denotes the second-order difference of the process $Y(.)$ with difference ϵ. For a probability density function ϕ in $C^2(R)$ with compact support contained in $(-1, 1]$, and for each $t \geq 0$ and $\epsilon > 0$, define

$$W_H^\epsilon(t) = \frac{1}{\epsilon} \int_{-\infty}^{\infty} \phi\left(\frac{t-x}{\epsilon}\right) W_H(x)dx,$$

$$R_\epsilon(t) = \frac{\epsilon^{2-H} W_H^{\epsilon\,(2)}(t)}{\sigma_{2H}}$$

with

$$\sigma_{2H}^2 = Var[\epsilon^{(2-H)} W_H^{\epsilon\,(2)}(t)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} |x|^{3-2H} |\hat{\phi}(-x)|^2 dx.$$

Here $\hat{\phi}(.)$ is the Fourier transform of the function $\phi(.)$. Berzin and Leon (2007) proved that, for $0 < H < 1$ and any integer $k \geq 1$,

$$\int_0^1 [R_\epsilon(u)]^k du \overset{\text{a.s.}}{\to} E(N^k) \text{ as } \epsilon \to 0 \quad (6.119)$$

where N is a standard Gaussian random variable.

Suppose we observe, instead of the process $X(t)$, a smoothed process $X_\epsilon(t) = \phi_\epsilon * X(t)$ obtained by convolution where $\phi_\epsilon(u) = \frac{1}{\epsilon}\phi(u/\epsilon)$. We extend the process $\{X(.), t \geq 0\}$ to the interval $(-\infty, 0)$ by defining $X(t) = c$, for $t < 0$. For the model specified by (6.114), it is assumed that the parameter μ and the constant c have the same sign where μ could also be equal to zero.

Let

$$Z_\epsilon(t) = \frac{\epsilon^{2-H} X_\epsilon^{(2)}(t)}{\sigma_{2H}} \text{ for the first two models}$$

$$= \frac{\epsilon^{2-H} X_\epsilon^{(2)}(t)}{\sigma_{2H} X_\epsilon(t)} \text{ for the last two models.} \quad (6.120)$$

Let

$$M_k(\epsilon) = \int_0^1 |X_\epsilon^{(2)}(u)|^k du \qquad \text{for the first two models}$$
$$= \int_0^1 |X_\epsilon^{(2)}(u)/X_\epsilon(u)|^k du \quad \text{for the last two models} \tag{6.121}$$

Berzin and Leon (2006) proved that

$$\frac{\epsilon^{k(2-H)} M_k(\epsilon)}{\sigma_{2H}^k \sigma^k \nu_k} \overset{\text{a.s.}}{\to} 1 \text{ as } \epsilon \to 0 \tag{6.122}$$

for every integer $k \geq 1$. Here ν_k denotes the kth absolute moment of the standard Gaussian distribution. Hence

$$\log(M_k(\epsilon)) = k(H-2)\log(\epsilon) + \log(\sigma_{2H}^k \sigma^k \nu_k) + o(1) \text{ a.s.} \tag{6.123}$$

as $\epsilon \to 0$. Let $h_i > 0, i = 1, \ldots, r$. Let $Y_i = \log(M_k(h_i)), x_i = \log(h_i)$ and $a_k = k(H-2), b_k = \log(\sigma_{2H}^k \sigma^k \nu_k)$. Then Equation (6.123) can be written in the form of a regression model

$$Y_i = a_k x_i + b_k + \eta_i,\ i = 1, \ldots, r \tag{6.124}$$

where $\eta_i, i = 1, \ldots, r$, are the errors. Applying the method of least squares, the estimators $\hat{H}_k$ of H and $\hat{B}_k$ of b_k are given by

$$k(\hat{H}_k - 2) = \sum_{i=1}^r z_i \log(M_k(h_i)) \tag{6.125}$$

and

$$\hat{B}_k = \frac{1}{r}\sum_{i=1}^r \log(M_k(h_i)) - k(\hat{H}_k - 2)\frac{1}{r}\sum_{i=1}^r \log(h_i) \tag{6.126}$$

where

$$z_i = \frac{d_i}{\sum_{i=1}^r d_i^2},\ d_i = \log(c_i) - \frac{1}{r}\sum_{i=1}^r \log(c_i). \tag{6.127}$$

Note that

$$\sum_{i=1}^r z_i = 0 \text{ and } \sum_{i=1}^r z_i d_i = 1. \tag{6.128}$$

Let us choose

$$\sigma_{2H}^{\hat{k}} = \sigma_{2\hat{H}_k}^k \tag{6.129}$$

as an estimator of σ_{2H}^{k} and

$$\sigma^{\hat{k}} = \frac{\exp(\hat{B}_k)}{\sigma_{2H}^{\hat{k}} \nu_k} \tag{6.130}$$

as an estimator of σ^k. Let $\hat{\sigma}_k$ be the estimator of σ defined by

$$\hat{\sigma}_k = (\sigma^{\hat{k}})^{1/k}. \tag{6.131}$$

Berzin and Leon (2006) proved that the estimator $\hat{H}_k$ is a strongly consistent estimator for H and

$$\frac{1}{\sqrt{\epsilon}}(\hat{H}_k - H)$$

is asymptotically Gaussian with mean zero and suitable variance as $\epsilon \to 0$. Furthermore, the estimator $\hat{\sigma}_k$ is a weakly consistent estimator for σ and

$$\frac{1}{\sqrt{\epsilon}\log(\epsilon)}(\hat{\sigma}_k - \sigma)$$

is asymptotically Gaussian with mean zero and suitable variance. It was shown in Berzin and Leon (2006) that the estimators with minimum asymptotic variance are obtained for $k = 2$.

If H is known, then Berzin and Leon (2006) prove that the estimator

$$\tilde{\sigma}_k = \frac{[\int_0^1 |Z_\epsilon(u)|^k du]^{1/k}}{\nu_k^{1/k}} \tag{6.132}$$

is a strongly consistent estimator for σ and $\epsilon^{-1/2}(\tilde{\sigma}_k - \sigma)$ is asymptotically Gaussian with mean zero and suitable variance. It can be shown that the asymptotic variance is minimum when $k = 2$.

6.7 Maximum likelihood estimation under misspecified model

We will now study the asymptotic behavior of the MLE of a parameter in the drift term of a fractional diffusion process under conditions in which the true drift function does not coincide with that specified by the parametric model. The need for such an analysis arises from the desirability of using estimators that are robust under departures from the underlying model.

Let us consider the problem of estimating the drift function of a fractional diffusion process satisfying the SDE given by

$$dX_t = b(X_t)dt + dW_T^H,\ X_0 = x,\ 0 \le t \le T \tag{6.133}$$

where W^H is standard fBm with Hurst index $H \in (\frac{1}{2}, 1)$ and the process X is observed over the interval $[0, T]$. Suppose that the true model (6.133) is misspecified and a parametric model given by

$$dX_t = f(\theta, X_t)dt + dW_t^H, \; X_0 = x, \; 0 \leq t \leq T \tag{6.134}$$

is used for estimating the trend parameter $\theta \in \Theta$. We now study the asymptotic behavior of the MLE of θ under departure of the true drift function $b(x)$ from the function $f(x, \theta)$ specified by the parametric model. Define the process

$$Q_{H,\theta}(t) = \frac{d}{dw_t^H} \int_0^t k_H(t, s) f(\theta, X_s) ds, \; 0 \leq t \leq T \tag{6.135}$$

where w_t^H and $k_H(t, s)$ are as defined in (1.72) and (1.70) in Chapter 1 respectively. Suppose the sample paths of the process $\{Q_{H,\theta}(t), 0 \leq t \leq T\}$ belong a.s. to $L^2([0, T], dw_t^H)$. Define

$$Z_t = \int_0^t k_H(t, s) dX_s, \; 0 \leq t \leq T. \tag{6.136}$$

Under the parametric model, the process $Z = \{Z_t, 0 \leq t \leq T\}$ is an $(\mathcal{F}_t)$-semimartingale with the decomposition

$$Z_t = \int_0^t Q_{H,\theta}(s) dw_s^H + M_t^H \tag{6.137}$$

where M^H is the fundamental martingale defined by (1.73) in Chapter 1 and the process X admits the representation

$$X_t = X_0 + \int_0^t K_H(t, s) dZ_s \tag{6.138}$$

where the function K_H is as defined by (1.75) in Chapter 1 with $f \equiv 1$. Let P_θ^T be the measure induced by the fractional diffusion process $\{X_t, 0 \leq t \leq T\}$ specified by the model defined by (6.134). Let P_0^T be the probability measure induced by fBm $\{W^H(t), 0 \leq t \leq T\}$. Following Theorem 1.20 of Chapter 1, we get that the Radon–Nikodym derivative of P_θ^T with respect to P_0^T is given by

$$\frac{dP_\theta^T}{dP_0^T} = \exp\left[\int_0^T Q_{H,\theta}(s) dZ_s - \frac{1}{2} \int_0^T Q_{H,\theta}^2(s) dw_s^H\right]. \tag{6.139}$$

In analogy with (6.135), we define the processes

$$Q_H(t) = \frac{d}{dw_t^H} \int_0^t k_H(t, s) b(X_s) ds, \; 0 \leq t \leq T. \tag{6.140}$$

Under the true model, the process Z defined by (6.136) can be represented in the form

$$Z_t = \int_0^t Q_H(s)dw_s^H + M_t^H,\ 0 \le t \le T \tag{6.141}$$

or equivalently

$$dZ_t = Q_H(t)dw_t^H + dM_t^H,\ 0 \le t \le T \tag{6.142}$$

From (6.139) and (6.142), we obtain the log-likelihood function and it is given by

$$\begin{aligned}\ell_T(\theta) &= \log \frac{dP_\theta^T}{dP_0^T} \\ &= \int_0^T Q_{H,\theta}(s)dZ_s - \frac{1}{2}\int_0^T Q_{H,\theta}^2(s)dw_s^H \\ &= \int_0^T Q_{H,\theta}(s)(Q_H(s)dw_s^H + dM_s^H) - \frac{1}{2}\int_0^T Q_{H,\theta}^2(s)dw_s^H \\ &= -\frac{1}{2}\int_0^T (Q_{H,\theta}(s) - Q_H(s))^2 dw_s^H + \frac{1}{2}\int_0^T Q_H^2(s)dw_s^H \\ &\quad + \int_0^T Q_{H,\theta}(s)dM_s^H. \end{aligned}\tag{6.143}$$

Let

$$I_{H,\theta}(T) = \int_0^T (Q_{H,\theta}(s) - Q_H(s))^2 dw_s^H \tag{6.144}$$

and

$$\eta_H(T) = \int_0^T Q_H^2(s)dw_s^H. \tag{6.145}$$

Then

$$\ell_T(\theta) = -\frac{1}{2}I_{H,\theta}(T) + \frac{1}{2}\eta_H(T) + \int_0^T Q_{H,\theta}(s)dM_s^H. \tag{6.146}$$

Suppose there exists a non random function $\gamma_H(T) \to \infty$ as $T \to \infty$ such that

(C1) $[\gamma_H(T)]^{-1} I_{H,\theta}(T) \overset{\text{a.s.}}{\to} I_{H,\theta}$ uniformly in $\theta \in \Theta$ as $T \to \infty$

(C2) $[\gamma_H(T)]^{-1} \eta_H(T) \overset{\text{a.s.}}{\to} \eta_H \ge 0$ as $T \to \infty$

(C3) $[\gamma_H(T)]^{-1} \sup_{\theta\in\Theta} |\int_0^T Q_{H,\theta}(s)dM_s^H| \overset{\text{a.s.}}{\to} 0$ in as $T \to \infty$

under the true model.

Let

$$\theta^* = \text{argmin}_{\theta \in \Theta} I_{H,\theta}. \tag{6.147}$$

Suppose it is unique. Let $\hat{\theta}_T$ be the MLE of θ under the parametric model. Note that the asymptotic behavior of the MLE or the maximizer of $\ell_T(\theta)$ is equivalent to the asymptotic behavior of the minimizer of $I_{H,\theta}(T)$ as $T \to \infty$ in view of the condition (C3). This implies the following result.

Theorem 6.8: Under conditions (C1)–(C3),

$$\hat{\theta}_T \overset{\text{a.s.}}{\to} \theta^* \text{ as } T \to \infty.$$

Suppose the log-likelihood function $\ell_T(\theta)$ is twice differentiable with respect to $\theta \in \Theta$. Then $\ell_T'(\hat{\theta}_T) = 0$. Expanding $\ell_T'(\theta)$ in a neighborhood of θ^*, we get that

$$\ell_T'(\theta^*) = \ell_T'(\hat{\theta}_T) + (\theta^* - \hat{\theta}_T)\ell_T''(\bar{\theta}_T) \tag{6.148}$$

where $|\bar{\theta}_T - \theta^*| \leq |\hat{\theta}_T - \theta^*|$ for T large. Since $\ell_T'(\hat{\theta}_T) = 0$, it follows that

$$\begin{aligned}\ell_T'(\theta^*) &= (\theta^* - \hat{\theta}_T)\ell_T''(\bar{\theta}_T)\\ &= (\theta^* - \hat{\theta}_T)[\ell_T''(\theta^*) + (\ell_T''(\bar{\theta}_T) - \ell_T''(\theta^*))].\end{aligned} \tag{6.149}$$

Theorem 6.9: In addition to conditions (C1)–(C3), suppose that the following conditions hold under the true model:

(C4) the log-likelihood function $\ell_T(\theta)$ is twice differentiable with respect to $\theta \in \Theta$ and the function $\ell_T''(\theta)$ is uniformly continuous in $\theta \in \Theta$;

(C5) $[\gamma_H(T)]^{-1/2}\ell_T'(\theta) \overset{\mathcal{L}}{\to} N(0, \sigma^2(\theta))$ as $T \to \infty$, for every $\theta \in \Theta$; and

(C6) $[\gamma_H(T)]^{-1}\ell_T''(\theta) \overset{p}{\to} C(\theta)$ as $T \to \infty$ for every $\theta \in \Theta$, where $C(\theta)$ is non-random and non-vanishing. Then

$$[\gamma_H(T)]^{1/2}(\hat{\theta}_T - \theta^*) \overset{\mathcal{L}}{\to} N\left(0, \frac{\sigma^2(\theta^*)}{C^2(\theta^*)}\right) \text{ as } T \to \infty$$

under the true model.

This theorem can be proved by standard arguments using the relation

$$[\gamma_H(T)]^{1/2}(\hat{\theta}_T - \theta^*) = \frac{[\gamma_H(T)]^{-1/2}\ell_T'(\theta^*)}{[\gamma_H(T)]^{-1}[\ell_T''(\theta^*) + o_p(1)]} \tag{6.150}$$

as $T \to \infty$. This follows by the condition (C4). Applying conditions (C5) and (C6), we get

$$[\gamma_H(T)]^{1/2}(\hat{\theta}_T - \theta^*) \xrightarrow{\mathcal{L}} N\left(0, \frac{\sigma^2(\theta^*)}{C^2(\theta^*)}\right) \text{ as } T \to \infty$$

under the true model.

7

Parametric estimation for processes driven by fractional Brownian sheet

7.1 Introduction

We have studied parametric and nonparametric inference for stochastic processes driven by fBm or mixed fBm. We will now study similar problems for random fields driven by a fractional Brownian sheet. Some applications to modeling are mentioned in the next section.

7.2 Parametric estimation for linear SDEs driven by a fractional Brownian sheet

We now study the asymptotic behavior of the MLE and the Bayes estimator for the parameter θ for random fields governed by SDEs of the type

$$X_{t,s} = \theta \int_0^t \int_0^s b(X_{v,u}) du dv + W_{t,s}^{\alpha,\beta}, \ 0 \leq t, \ s \leq T \tag{7.1}$$

where $W^{\alpha,\beta}$ is a fractional Brownian sheet (fBs) with Hurst parameters $\alpha, \beta \in (\frac{1}{2}, 1)$. The Bernstein–von Mises-type theorem and its application for Bayes estimation of parameters for diffusion fields were discussed in Prakasa Rao (1984). Sottinen and Tudor (2008) have recently investigated the problem of maximum likelihood estimation for linear SDEs driven by a fractional

Statistical Inference for Fractional Diffusion Processes B. L. S. Prakasa Rao

Brownian sheet. Models of the form (7.1) have been used in the context of identification or detection in signal estimation problems. Here the coefficient θ is unobservable and it has to be estimated from observation of the process X. Two-dimensional models where the noise has fractional behavior are used in satellite imaging (cf. Pesquet-Popescu and Levy Vehel (2002)), in radar image data classification (cf. Jung Lae and Chyi-Chying (1997)), in the classification and segmentation of hydrological basins (cf. Maitre and Pinciroli (1999)) and in medical applications such as the detection of osteoporosis from X-ray images (cf. Leger (2001), Jeanne *et al.* (2001)).

Two-parameter martingales

We now give a brief introduction to two-parameter martingale theory which is needed for results discussed later in this chapter.

Let R_+^2 be the positive orthant of R^2 with the partial ordering $\mathbf{z}_1 < \mathbf{z}_2$ with $\mathbf{z}_1 = (t_1, s_1)$ and $\mathbf{z}_2 = (t_2, s_2), t_i \geq 0, s_i \geq 0, i = 1, 2,$ if $t_1 \leq t_2$ and $s_1 \leq s_2$. If $t_1 < t_2$ and $s_1 < s_2$, we denote the ordering by $\mathbf{z}_1 \ll \mathbf{z}_2$. If $\mathbf{z}_1 \ll \mathbf{z}_2$, then $(\mathbf{z}_1, \mathbf{z}_2]$ will be the rectangle $(t_1, t_2] \times (s_1, s_2]$ and if $f_{\mathbf{z}}$ is a function defined on R_+^2, then $f((\mathbf{z}_1, \mathbf{z}_2])$ will denote $f_{t_2,s_2} - f_{t_2,s_1} - f_{t_1,s_2} + f_{t_1,s_1}$. We will denote the rectangle $\{\zeta : 0 < \zeta < \mathbf{z}\}$ by $R_{\mathbf{z}}$.

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\{\mathcal{F}_{\mathbf{z}}, \mathbf{z} \in R_+^2\}$ be a family of sub-σ-algebras of $\mathcal{F}$ satisfying the following conditions:

(F1) $\mathcal{F}_{\mathbf{z}_1} \subset \mathcal{F}_{\mathbf{z}_2}$ if $\mathbf{z}_1 < \mathbf{z}_2$;

(F2) $\mathcal{F}_0$ contains all null sets of $\mathcal{F}$;

(F3) for each $\mathbf{z}$, $\mathcal{F}_{\mathbf{z}} = \cap_{\mathbf{z} \ll \zeta} \mathcal{F}_\zeta$;

(F4) for each $\mathbf{z} = (t, s)$, $\mathcal{F}_{\mathbf{z}}^1$ and $\mathcal{F}_{\mathbf{z}}^2$ are conditionally independent. Here $\mathcal{F}_{\mathbf{z}}^1 = \mathcal{F}_{t,\infty}$ and $\mathcal{F}_{\mathbf{z}}^2 = \mathcal{F}_{\infty,s}$ with

$$\mathcal{F}_{t,\infty} = \vee_{s \geq 0} \mathcal{F}_{t,s} = \sigma(\cup_{s \geq 0} \mathcal{F}_{t,s})$$

and

$$\mathcal{F}_{\infty,s} = \vee_{t \geq 0} \mathcal{F}_{t,s} = \sigma(\cup_{t \geq 0} \mathcal{F}_{t,s}).$$

Note that the condition (F4) is equivalent to the condition:

$(F4')$ for all bounded random variables X and for all $\mathbf{z} \in R_+^2$,

$$E(X|\mathcal{F}_{\mathbf{z}}) = E[E(X|\mathcal{F}_{\mathbf{z}}^1)|\mathcal{F}_{\mathbf{z}}^2] = E[E(X|\mathcal{F}_{\mathbf{z}}^2)|\mathcal{F}_{\mathbf{z}}^1] \text{a.s.}$$

Definition: Let $\{\mathcal{F}_{\mathbf{z}}, \mathbf{z} \in \mathbf{R}_+^2\}$ be a filtration satisfying conditions (F1)–(F4).

(i) A process $X = \{X_{\mathbf{z}}, \mathbf{z} \in R_+^2\}$ is called a *two-parameter martingale* with respect to $\{\mathcal{F}_{\mathbf{z}}, \mathbf{z} \in R_+^2\}$ if (a) for each $\mathbf{z} \in R_+^2$, $X_{\mathbf{z}}$ is adapted to $\mathcal{F}_{\mathbf{z}}$ and integrable, and (b) for each $\mathbf{z} < \mathbf{z}'$, $E(X_{\mathbf{z}'}|\mathcal{F}_{\mathbf{z}}) = X_{\mathbf{z}}$ a.s.

(ii) The process $X = \{X_{\mathbf{z}}, \mathbf{z} \in \mathbf{R}^{\mathbf{2}}_{+}\}$ is called a *weak martingale* with respect to $\{\mathcal{F}_{\mathbf{z}}, \mathbf{z} \in \mathbf{R}^{\mathbf{2}}_{+}\}$ if (a) for each $\mathbf{z} \in \mathbf{R}^{\mathbf{2}}_{+}, \mathbf{X_z}$ is adapted to $\mathcal{F}_{\mathbf{z}}$ and integrable, and (b) for each $\mathbf{z} < \mathbf{z}', \mathbf{E}(\mathbf{X}((\mathbf{z}, \mathbf{z}'])|\mathcal{F}_{\mathbf{z}}) = \mathbf{0}$ a.s.

(iii) The process $X = \{X_{\mathbf{z}}, \mathbf{z} \in \mathbf{R}^{\mathbf{2}}_{+}\}$ is called an *i-martingale*, $i = 1, 2$, with respect to $\{\mathcal{F}_{\mathbf{z}}, \mathbf{z} \in \mathbf{R}^{\mathbf{2}}_{+}\}$ if (a) for each $\mathbf{z} \in \mathbf{R}^{\mathbf{2}}_{+}, \mathbf{X_z}$ is adapted to $\mathcal{F}^{i}_{\mathbf{z}}$ and integrable, and (b) for each $\mathbf{z} \ll \mathbf{z}', \mathbf{E}(\mathbf{X}((\mathbf{z}, \mathbf{z}'])|\mathcal{F}^{\mathbf{i}}_{\mathbf{z}}) = \mathbf{0}$ a.s.

(iv) The process $X = \{X_{\mathbf{z}}, \mathbf{z} \in \mathbf{R}^{\mathbf{2}}_{+}\}$ is called a *strong martingale*, if, in addition, X vanishes on the axes, that is, $X_{0,s} = X_{t,0} = 0$ for all $0 \le t, s, < \infty$, and $E(X((\mathbf{z}, \mathbf{z}'])|\mathcal{F}^{\mathbf{1}}_{\mathbf{z}} \vee \mathcal{F}^{\mathbf{2}}_{\mathbf{z}}) = \mathbf{0}$ a.s. for all $\mathbf{z} \ll \mathbf{z}'$.

It can be checked that any two-parameter martingale is both a 1-and 2-martingale. Conversely, if X is both a 1- and 2-martingale, then X is a two-parameter martingale provided $\{X_{t,0}, \mathcal{F}^{1}_{t,0}, t \in R_{+}\}$ and $\{X_{0,s}, \mathcal{F}^{1}_{0,s}, s \in R_{+}\}$ are both martingales. Furthermore, any two-parameter martingale is a weak martingale and any strong martingale is a two-parameter martingale.

A process $\{X_{\mathbf{z}}\}$ is said to be *right-continuous* if $\lim_{\mathbf{z}' \to \mathbf{z}: \mathbf{z} \ll \mathbf{z}'} = X_{\mathbf{z}}$ a.s.

Given a filtration $\{\mathcal{F}_{\mathbf{z}}, \mathbf{z} \in R^{2}_{+}\}$ satisfying properties (F1)–(F4), a process $X = \{X_{\mathbf{z}}, \mathbf{z} \in R^{2}_{+}\}$ is called an increasing process if (i) X is right-continuous and adapted to $\{\mathcal{F}_{\mathbf{z}}, \mathbf{z} \in R^{2}_{+}\}$, (ii)$X_{\mathbf{z}} = 0$ a.s. on the axes and (iii) $X(A) \ge 0$ for every rectangle $A \subset R^{2}_{+}$.

Let $\mathbf{T} = (T_1, T_2) \in R^{2}_{+}$. Let $M^{2}_{S}(\mathbf{T})$ be the class of strong martingales on $R_{\mathbf{T}}$ such that $E|M_{\mathbf{z}}|^2 < \infty$ for all $\mathbf{z} \in R_{\mathbf{T}}$. Let $M \in M^{2}_{S}(\mathbf{T})$. It is known that there exists a unique $\mathcal{F}^{1}_{\mathbf{z}}$-predictable increasing process $[M]^{(1)}$ and a unique $\mathcal{F}^{2}_{\mathbf{z}}$-predictable increasing process $[M]^{(2)}$ such that $M^{2}_{\mathbf{z}} - [M]^{(i)}$ is an i-martingale for $i = 1, 2$ respectively. Following Cairoli and Walsh (1975), for a strong martingale M, one can choose either the process $[M]^{(1)}$ or the process $[M]^{(2)}$ as the increasing process $A_{\mathbf{z}}$ such that the process $\{M^{2}_{\mathbf{z}} - A_{\mathbf{z}}, \mathbf{z} \in \mathbf{T}\}$ is a two-parameter martingale. In general, it is not necessary that $[M]^{(1)} = [M]^{(2)}$ a.s. However, if $\{\mathcal{F}_{\mathbf{z}}, \mathbf{z} \in R^{2}_{+}\}$ is generated by the Brownian sheet defined below, then $[M]^{(1)} = [M]^{(2)}$ a.s. and this process is called the quadratic variation of the two-parameter strong martingale M and is denoted by $\langle M \rangle$.

Let ϕ be a bi-measurable function of $(\omega, \mathbf{z})$ such that $\phi_{\mathbf{z}}$ is $\mathcal{F}_{\mathbf{z}}$-measurable and

$$\int_{\mathbf{T}} E[\phi^{2}_{\mathbf{z}}] d\mathbf{z} < \infty.$$

Then it is possible to define the integral $\int_{\mathbf{T}} \phi dM$.

The Brownian sheet is a Gaussian process $\{W_{t,s}, 0 \le t, s \le T\}$ starting from zero with mean zero and covariance function

$$E(W_{t,s} W_{v,u}) = \min(t, v)\ \min(s, u),\ 0 \le t, s, v, u \le T.$$

An alternative way of describing the process is as follows.

The Wiener sheet (Brownian sheet) or the Wiener random field $\mathbf{W} = \{\mathbf{W}_{\mathbf{Z}}, \mathcal{F}_{\mathbf{Z}}, \mathbf{Z} \in \mathbf{R}^{\mathbf{2}}_{+}\}$ is a two-parameter Wiener process with continuous sample paths with the property that, if $\mathbf{Z} = (t, s)$, then for any fixed t, the process $\{W_{t,s}, 0 \leq s \leq \infty\}$ is a standard Wiener process and for any fixed s, the process $\{W_{t,s}, 0 \leq t \leq \infty\}$ is a standard Wiener process.

For the case of the Wiener sheet described above, stochastic integrals of the type

$$\int_{\mathbf{T}} \phi_\zeta dW_\zeta$$

are studied in Wong and Zakai (1976) and Cairoli and Walsh (1975). Suppose that $\phi = \{\phi_{\mathbf{Z}}, \mathbf{Z} \in R^2_+\}$ is an $\mathcal{F}_{\mathbf{Z}}$-adapted measurable random field such that

$$E\left(\int_{R_{\mathbf{Z}}} \phi^2_\zeta d\zeta\right) < \infty.$$

Following Cairoli and Walsh (1975), the integral

$$\mathrm{I}_{\mathbf{Z}} = \int_{R_{\mathbf{Z}}} \phi_\zeta \; dW_\zeta$$

can be defined. The following result is due to Etemadi and Kallianpur (1977).

Theorem 7.1: Let $\{X_{\mathbf{z}}, \mathcal{F}_{\mathbf{z}}, \mathbf{z} \in \mathbf{T}\}$ be a square integrable strong martingale such that the filtration $\{\mathcal{F}_{\mathbf{z}}, \mathbf{z} \in \mathbf{T}\}$ is generated by a Brownian sheet. Then

$$\exp\left(X_{\mathbf{z}} - \frac{1}{2}\langle X\rangle_{\mathbf{z}}\right)$$

is a martingale if and only if

$$\mathrm{E}\left[\exp\left(X_{\mathbf{z}} - \frac{1}{2}\langle X\rangle_{\mathbf{z}}\right)\right] = 1.$$

Our discussion here is based on Amirdjanova and Linn (2007) and Prakasa Rao (1984).

Fractional Brownian sheet

Thefractional Brownian sheet with Hurst index $(\alpha, \beta) \in (0, 1) \times (0, 1)$ is a Gaussian process $W^{\alpha,\beta} \equiv \{W^{\alpha,\beta}_{t,s}, 0 \leq t, s \leq T\}$ starting from zero with mean zero and covariance function

$$E(W^{\alpha,\beta}_{t,s} W^{\alpha,\beta}_{v,u}) = \frac{1}{2}(t^{2\alpha} + v^{2\alpha} - |t-v|^{2\alpha})\frac{1}{2}(s^{2\beta} + u^{2\beta} - |s-u|^{2\beta}), 0 \leq t, s, v, u \leq T.$$

As in the one-dimensional case, we can associate a two-parameter strong martingale with the fractional Brownian sheet following Sottinen and Tudor (2008). We explain this in some detail to introduce the notation and for completeness. Following the notation introduced in Chapter 1, let

$$M_{t,s}^{\alpha,\beta} = \int_0^t \int_0^s k_\alpha(t,v) k_\beta(s,u) dW_{u,v}^{\alpha,\beta}, \quad 0 \le t,s,v,u \le T \tag{7.2}$$

where $k_\alpha(t,v)$ is as defined by (1.70) in Chapter 1. It can be shown that the process $M^{\alpha,\beta} \equiv \{M_{t,s}^{\alpha,\beta}, 0 \le t,s,v,u \le T\}$ is a Gaussian strong martingale with the quadratic variation equal to $w_t^\alpha w_s^\beta$ where w_t^α is as given by (1.72) in Chapter 1. The integral (7.2) can be defined as a Wiener integral with respect to the fractional Brownian sheet $W^{\alpha,\beta}$. The filtration generated by the martingale $M^{\alpha,\beta}$ coincides with the filtration generated by the fractional Brownian sheet $W^{\alpha,\beta}$.

Linear SDE driven by fractional Brownian sheet

Let us consider the SDE

$$dX_\zeta = [\eta_\zeta(X) + \theta\psi_\zeta(X)]d\zeta + \sigma(\zeta)\, dW_\zeta^{\alpha,\beta}, \quad \zeta \in R_+^2 \tag{7.3}$$

where $\theta \in \Theta \subset R$, $W = \{W_\zeta^{\alpha,\beta}, \zeta \in R_+^2\}$ is a fractional Brownian sheet with parameter (α, β) and $\sigma(\zeta)$ is a nonrandom positive function on $[0, \infty)$. In other words, $X = \{X_t, t \ge 0\}$ is a stochastic process satisfying the stochastic integral equation

$$X(\mathbf{Z}) = \int_{R_Z} [\eta_\zeta(X) + \theta\psi_\zeta(X)]d\zeta + \int_{R_Z} \sigma(\zeta) dW_\zeta^{\alpha,\beta}, \;\; \zeta \in R_+^2. \tag{7.4}$$

Let

$$C(\theta,\zeta) = [\eta_\zeta(X) + \theta\psi_\zeta(X)], \;\; \zeta \in R_+^2 \tag{7.5}$$

and suppose that the random field is such that

$$R_{t,s} = \frac{d}{dw_t^\alpha} \frac{d}{dw_s^\beta} \int_0^t \int_0^s k_\alpha(t,v) k_\beta(s,u) [C(\theta,(u,v)) \backslash \sigma(u,v))] du dv \tag{7.6}$$

is well defined. For $0 < \alpha, \beta < 1$, and if $C(\theta, \zeta)$ is Lipschitz, then the sample paths of the random field R over the region $[0,T] \times [0,T]$ belong to $L^2([0,T] \times [0,T], w^\alpha \times w^\beta)$. Following the arguments in Sottinen and Tudor (2008), define

$$\begin{aligned} Z_{t,s} &= \int_0^t \int_0^s k_\alpha(t,v) k_\beta(s,u) [\sigma(v,u)]^{-1} dX_{v,u} \\ &= \int_0^t \int_0^s k_\alpha(t,v) k_\beta(s,u) [\sigma(v,u)]^{-1} C(\theta,(u,v)) du dv + M_{t,s}^{\alpha,\beta}. \end{aligned} \tag{7.7}$$

Suppose $\frac{1}{2} < \alpha, \beta < 1$ and $\sigma(t, s) \equiv 1$. Then

$$X_{t,s} = \int_0^t \int_0^s K_\alpha(t, v) K_\beta(s, u) dZ_{u,v} \tag{7.8}$$

where

$$K_\alpha(t, v) = \alpha(2\alpha - 1) \int_0^t r^{2\alpha-1}(r - v)^{\alpha-\frac{3}{2}} dr.$$

The above representation is given in Sottinen and Tudor (2008) following construction in Kleptsyna *et al.* (2000a) and Tudor and Tudor (2002). It is easy to see that

$$Z_{t,s} = \int_0^t \int_0^s R_{u,v} dw_u^\alpha dw_v^\beta + M_{t,s}^{\alpha,\beta}. \tag{7.9}$$

It can be shown that the filtrations generated by the random fields $\{Z_{t,s}, 0 \leq t, s \leq T\}$ and $\{X_{t,s}, 0 \leq t, s \leq T\}$ are the same (see Theorem 1.19 in Chapter 1 in the one-dimensional case of fBm). Hence no information is lost by using the data $\{Z_{t,s}, 0 \leq t, s \leq T\}$ instead of $\{X_{t,s}, 0 \leq t, s \leq T\}$ for inferential purposes. We use this observation for developing the methods of estimation for the parameter θ.

Let P_θ^T be the probability measure induced by the random field $\{X_{t,s}, 0 \leq t \leq T, 0 \leq s \leq T\}$. Following Theorem 4.7 in Sottinen and Tudor (2008), the Radon–Nikodym derivative of the measure P_θ^T with respect to P_0^T is given by

$$\frac{dP_\theta^T}{dP_0^T} = \exp\left(\int_0^T \int_0^T R_{t,s} dZ_{t,s} - \frac{1}{2}\int_0^T \int_0^T R_{t,s}^2 dw_t^\alpha dw_s^\beta\right). \tag{7.10}$$

Maximum likelihood estimation

We now consider the problem of estimation of the parameter θ based on observation of the process $X = \{X_{t,s}, 0 \leq t, s \leq T\}$ and study its asymptotic properties as $T \to \infty$.

Strong consistency

Let $L_T(\theta)$ denote the Radon–Nikodym derivative dP_θ^T/dP_0^T. The MLE is defined by the relation

$$L_T(\hat{\theta}_T) = \sup_{\theta \in \Theta} L_T(\theta). \tag{7.11}$$

We assume that there exists a measurable MLE. Sufficient conditions can be given for the existence of such an estimator (cf. Prakasa Rao (1987), Lemma 3.1.2). Note that

$$
\begin{aligned}
R_{t,s} &= \frac{d}{dw_t^\alpha}\frac{d}{dw_s^\beta}\int_0^t\int_0^s k_\alpha(t,v)k_\beta(s,u)[C(\theta,(u,v))/\sigma(u,v)]dudv \\
&= \frac{d}{dw_t^\alpha}\frac{d}{dw_s^\beta}\int_0^t\int_0^s k_\alpha(t,v)k_\beta(s,u)[\eta(u,v)/\sigma(u,v)]dudv \\
&\quad +\theta\frac{d}{dw_t^\alpha}\frac{d}{dw_s^\beta}\int_0^t\int_0^s k_\alpha(t,v)k_\beta(s,u)[\psi(u,v)/\sigma(u,v)]dudv \\
&= J_1(t,s)+\theta J_2(t,s)(\text{say}).
\end{aligned}
\tag{7.12}
$$

Then

$$
\begin{aligned}
\log L_T(\theta) &= \int_0^T\int_0^T (J_1(t,s)+\theta J_2(t,s))dZ_{t,s} \\
&\quad -\frac{1}{2}\int_0^T\int_0^T (J_1(t,s)+\theta J_2(t,s))^2 dw_t^\alpha dw_s^\beta
\end{aligned}
\tag{7.13}
$$

and the likelihood equation is given by

$$
\int_0^T\int_0^T J_2(t,s)dZ_{t,s} - \int_0^T\int_0^T (J_1(t,s)+\theta J_2(t,s))J_2(t,s)\,dw_t^\alpha dw_s^\beta = 0.
\tag{7.14}
$$

Hence the MLE $\hat{\theta}_T$ of θ is given by

$$
\hat{\theta}_T = \frac{\int_0^T\int_0^T J_2(t,s)dZ_{t,s} - \int_0^T\int_0^T J_1(t,s)J_2(t,s)dw_t^\alpha dw_s^\beta}{\int_0^T\int_0^T J_2^2(t,s)dw_t^\alpha dw_s^\beta}.
\tag{7.15}
$$

Let θ_0 be the true parameter. Using the fact that

$$
dZ_{t,s} = (J_1(t,s)+\theta_0 J_2(t,s))dw_t^\alpha dw_s^\beta + dM_{t,s}^{\alpha,\beta},
\tag{7.16}
$$

it can be shown that

$$
\begin{aligned}
\frac{dP_\theta^T}{dP_{\theta_0}^T} &= \exp\Bigg[(\theta-\theta_0)\int_0^T\int_0^T J_2(t,s)dM_{t,s}^{\alpha,\beta} - \frac{1}{2}(\theta-\theta_0)^2 \\
&\qquad \int_0^T\int_0^T J_2^2(t,s)dw_t^\alpha dw_s^\beta\Bigg].
\end{aligned}
\tag{7.17}
$$

Following this representation of the Radon–Nikodym derivative, we obtain that

$$
\hat{\theta}_T - \theta_0 = \frac{\int_0^T\int_0^T J_2(t,s)dM_{t,s}^{\alpha,\beta}}{\int_0^T\int_0^T J_2^2(t,s)dw_t^\alpha dw_s^\beta}.
\tag{7.18}
$$

We now discuss the problem of estimation of the parameter θ on the basis of the observation of the random field X or equivalently the random field Z over the region $[0, T] \times [0, T]$.

(A1) Let $\{M_{t,s}^{\alpha,\beta}\}$ be a two-parameter strong martingale. Assume that a random field $\{G(t, s), 0 \leq t, s \leq \infty\}$ is such that the stochastic integral

$$\int_0^T \int_0^T G(t, s) dM_{t,s}^{\alpha,\beta}$$

exits in the Cairoli–Walsh sense and

$$\int_0^T \int_0^T G^2(t, s) dw_t^\alpha dw_s^\beta \to \infty \quad \text{a.s. as } T \to \infty.$$

Further suppose that

$$\frac{\int_0^T \int_0^T G(t, s) dM_{t,s}^{\alpha,\beta}}{\int_0^T \int_0^T G^2(t, s) dw_t^\alpha dw_s^\beta} \to 0 \text{ a.s. as } T \to \infty. \tag{7.19}$$

Theorem 7.2: Suppose the condition (A1) holds. Then the MLE $\hat{\theta}_T$ is strongly consistent, that is,

$$\hat{\theta}_T \to \theta_0 \quad \text{a.s.} \quad [P_{\theta_0}] \text{ as } T \to \infty \tag{7.20}$$

provided

$$\int_0^T \int_0^T J_2^2(t, s) dw_t^\alpha dw_s^\beta \to \infty \quad \text{a.s.} \quad [P_{\theta_0}] \text{ as } T \to \infty. \tag{7.21}$$

Limiting distribution

We now study the limiting distribution of the MLE $\hat{\theta}_T$ as $T \to \infty$.

(A2) Let $\{M_{t,s}^{\alpha,\beta}\}$ be a two-parameter strong martingale. Assume that the random field $\{G(t, s), 0 \leq t, s \leq \infty\}$ is such that the stochastic integral

$$\nu_T \equiv \int_0^T \int_0^T G(t, s) dM_{t,s}^{\alpha,\beta}$$

exists in the Cairoli–Walsh sense and let

$$R_T \equiv \int_0^T \int_0^T G^2(t, s) dw_t^\alpha dw_s^\beta.$$

Further assume that the process $\{R_t, t \geq 0\}$ is such that there exists a norming function $\{I_t, t \geq 0\}$ such that

$$I_T^2 R_T = I_T^2 \int_0^T \int_0^T G^2(t,s) dw_t^\alpha dw_s^\beta \to \eta^2 \text{in probability as } T \to \infty \tag{7.22}$$

where $I_T \to 0$ as $T \to \infty$ and η is a random variable such that $P(\eta > 0) = 1$. Furthermore,

$$(I_T \nu_T, I_T^2 R_T) \to (\eta Z, \eta^2) \text{ in law as } T \to \infty \tag{7.23}$$

where the random variable Z and η are independent.

Observe that

$$I_T^{-1}(\hat{\theta}_T - \theta_0) = \frac{I_T \nu_T}{I_T^2 R_T} \tag{7.24}$$

Then we obtain the following result under the condition (A2).

Theorem 7.3: Suppose that the condition (A2) holds. Then

$$I_T^{-1}(\hat{\theta}_T - \theta_0) \to \frac{Z}{\eta} \text{ in law} \quad \text{as} \quad t \to \infty \tag{7.25}$$

where the random variable Z has the standard normal distribution and the random variables Z and η are independent.

Remarks: The problem of strong law of large numbers for normalized stochastic integrals with respect to two-parameter martingales is open. The central limit theorem for normalized stochastic integrals with respect to two-parameter martingales is also an open problem. Knopov (1982) stated a central limit theorem for stochastic integrals with respect to a Wiener sheet. The central limit theorem and strong law of large numbers for stochastic integrals with respect to one-parameter martingales are known (cf. Prakasa Rao (1999b)). Central limit theorems for stochastic integrals with respect to martingales in one dimension are proved using the fact that a one-parameter martingale can be transformed into a standard Wiener process by a random time change and a central limit theorem for stochastic integrals with respect to a Wiener process is known (cf. Prakasa Rao (1999a)). However, it is not possible to convert a two-parameter martingale into a Wiener sheet by a random time change in general. Hence it is not clear how to reduce the problem of study of the central limit theorem for two-parameter martingales to the study of the central limit theorem for stochastic integrals with respect to a Wiener sheet and apply the result stated in Knopov (1982).

Bayes estimation

Suppose that the parameter space Θ is open and Λ is a prior probability measure on the parameter space Θ. Further suppose that Λ has the density $\lambda(.)$ with respect to the Lebesgue measure and the density function is continuous and positive in an open neighborhood of θ_0, the true parameter. Let

$$A_T \equiv I_T \int_0^T \int_0^T J_2(t,s))dM_{t,s}^{\alpha,\beta} \tag{7.26}$$

and

$$B_T \equiv I_T^2 \int_0^T \int_0^T J_2^2(t,s)dw_t^\alpha dw_s^\beta. \tag{7.27}$$

We saw earlier that the MLE satisfies the relation

$$A_T = (\hat{\theta}_T - \theta_0)I_T^{-1}B_T. \tag{7.28}$$

The posterior density of θ given the observation $X^T \equiv \{X_{t,s}, 0 \le t, s \le T\}$ is given by

$$p(\theta|X^T) = \frac{(dP_\theta^T/dP_{\theta_0}^T)\lambda(\theta)}{\int_\Theta (dP_\theta^T/dP_{\theta_0}^T)\lambda(\theta)\,d\theta}. \tag{7.29}$$

Let us write $t = I_T^{-1}(\theta - \hat{\theta}_T)$ and define

$$p^*(t|X^T) = I_T p(\hat{\theta}_T + tI_T|X^T). \tag{7.30}$$

Then the function $p^*(t|X^T)$ is the posterior density of the transformed variable $t = I_T^{-1}(\theta - \hat{\theta}_T)$. Let

$$\nu_T(t) \equiv \frac{dP_{\hat{\theta}_T+tI_T}/dP_{\theta_0}}{dP_{\hat{\theta}_T}/dP_{\theta_0}} = \frac{dP_{\hat{\theta}_T+tI_T}}{dP_{\hat{\theta}_T}} \text{ a.s.} \tag{7.31}$$

and

$$C_T = \int_{-\infty}^{\infty} \nu_T(t)\lambda(\hat{\theta}_T + tI_T)dt. \tag{7.32}$$

It can be checked that

$$p^*(t|X^T) = C_T^{-1}\nu_T(t)\lambda(\hat{\theta}_T + tI_T) \tag{7.33}$$

and

$$\begin{aligned}\log \nu_T(t) &= I_T^{-1}\alpha_T[(\hat{\theta}_T + tI_T - \theta_0) - (\hat{\theta}_T - \theta_0)] \\ &\quad -\frac{1}{2}I_T^{-2}B_T[(\hat{\theta}_T + tI_T - \theta_0)^2 - (\hat{\theta}_T - \theta_0)^2] \\ &= t\alpha_T - \frac{1}{2}t^2 B_T - tB_T I_T^{-1}(\hat{\theta}_T - \theta_0) \\ &= -\frac{1}{2}B_T t^2. \end{aligned} \tag{7.34}$$

Suppose that convergence of the condition in Equation (7.22) holds a.s. under the measure P_{θ_0} and the limit is a constant $\eta^2 > 0$ with probability one. For convenience, we write $\gamma = \eta^2$. Then

$$B_T \to \gamma \quad \text{a.s.} \quad [P_{\theta_0}] \quad \text{as} \quad T \to \infty. \tag{7.35}$$

Then it is obvious that

$$\lim_{T\to\infty} \nu_T(t) = \exp\left(-\frac{1}{2}\gamma t^2\right) \quad \text{a.s.} \quad [P_{\theta_0}] \tag{7.36}$$

and, for any $0 < \varepsilon < \gamma$,

$$\log \nu_T(t) \le -\frac{1}{2}t^2(\gamma - \varepsilon) \tag{7.37}$$

for every t for T sufficiently large. Furthermore, for every $\delta > 0$, there exists $\varepsilon' > 0$ such that

$$\sup_{|t| > \delta I_T^{-1}} \nu_T(t) \le \exp\left(-\frac{1}{4}\varepsilon' I_T^{-2}\right) \tag{7.38}$$

for T sufficiently large.

Suppose that $H(t)$ is a nonnegative measurable function such that, for some $0 < \varepsilon < \gamma$,

$$\int_{-\infty}^{\infty} H(t)\exp\left[-\frac{1}{2}t^2(\gamma - \varepsilon)\right] dt < \infty. \tag{7.39}$$

Further suppose that the MLE $\hat{\theta}_T$ is strongly consistent, that is,

$$\hat{\theta}_T \to \theta_0 \quad \text{a.s.} \qquad [P_{\theta_0}] \text{ as } \quad T \to \infty \tag{7.40}$$

The following lemmas can be proved following arguments similar to those given in Chapter 2 for studying the properties of the Bayes estimators for SDEs driven by fBm. We omit the details.

Lemma 7.4: Under the conditions stated above, there exists $\delta > 0$ such that

$$\lim_{T\to\infty} \int_{|t|\le\delta I_T^{-1}} H(t) \left| \nu_T(t)\lambda(\hat{\theta}_T + tI_T) - \lambda(\theta_0)\exp\left(-\frac{1}{2}\gamma t^2\right)\right| dt = 0. \quad (7.41)$$

Suppose the following condition holds for every $\varepsilon > 0$ and $\delta > 0$:

$$\exp(-\varepsilon I_T^{-2}) \int_{|u|>\delta} H(uI_T^{-1})\lambda(\hat{\theta}_T + u)du \to 0 \quad \text{a.s.} \quad [P_{\theta_0}] \quad \text{as} \quad T\to\infty. \quad (7.42)$$

Lemma 7.5: Suppose that the conditions stated earlier hold. Then, for every $\delta > 0$,

$$\lim_{T\to\infty} \int_{|t|>\delta I_T^{-1}} H(t) \left| \nu_T(t)\lambda(\hat{\theta}_T + tI_T) - \lambda(\theta_0)\exp\left(-\frac{1}{2}\gamma t^2\right)\right| dt = 0. \quad (7.43)$$

Lemmas 7.4 and 7.5 together prove that

$$\lim_{T\to\infty} \int_{-\infty}^{\infty} H(t) \left| \nu_T(t)\lambda(\hat{\theta}_T + tI_T) - \lambda(\theta_0)\exp\left(-\frac{1}{2}\beta t^2\right)\right| dt = 0. \quad (7.44)$$

Let $H(t) \equiv 1$ in (7.44). Note that

$$C_T \equiv \int_{-\infty}^{\infty} \nu_T(t)\lambda(\hat{\theta}_T + tI_T)dt.$$

Relation (7.44) implies that

$$C_T \to \lambda(\theta_0) \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\gamma t^2\right) dt = \lambda\theta_0 \left(\frac{\gamma}{2\pi}\right)^{-1/2} \text{ a.s. } [P_{\theta_0}] \quad (7.45)$$

as $T \to \infty$. Furthermore,

$$\begin{aligned} &\int_{-\infty}^{\infty} H(t) \left| p^*(t|X^T) - \left(\frac{\gamma}{2\pi}\right)^{1/2} \exp\left(-\frac{1}{2}\gamma t^2\right)\right| dt \\ &\le C_T^{-1} \int_{-\infty}^{\infty} H(t) \left| \nu_T(t)\lambda(\hat{\theta}_T + tI_T) - \lambda(\theta_0)\exp\left(-\frac{1}{2}\gamma t^2\right)\right| dt \\ &+ \int_{-\infty}^{\infty} H(t) \left| C_T^{-1}\lambda(\theta_0) - \left(\frac{\gamma}{2\pi}\right)^{1/2}\right| \exp\left(-\frac{1}{2}\beta t^2\right) dt. \end{aligned} \quad (7.46)$$

The last two terms tend to zero a.s. $[P_{\theta_0}]$ by Equations (7.44) and (7.45). Hence we have the following theorem which is an analogue of the Bernstein–von Mises theorem for a class of processes satisfying a linear SDE driven by the standard Wiener process proved in Prakasa Rao (1981).

Theorem 7.6: Let the assumptions (7.35), (7.39), (7.40) and (7.42) hold where $\lambda(.)$ is a prior density which is continuous and positive in an open neighborhood of θ_0, the true parameter. Then

$$\lim_{T\to\infty}\int_{-\infty}^{\infty} H(t)\left|p^*(t|X^T)-\left(\frac{\gamma}{2\pi}\right)^{1/2}\exp\left(-\frac{1}{2}\gamma t^2\right)\right| dt = 0 \text{ a.s. } [P_{\theta_0}]. \quad (7.47)$$

As a consequence of the above theorem, we obtain the following result by choosing $H(t) = |t|^m$, for any integer $m \geq 0$.

Theorem 7.7: Assume that the following conditions hold:

$$\hat{\theta}_T \to \theta_0 \text{ a.s. } [P_{\theta_0}] \text{ as } T \to \infty, \quad (7.48)$$

$$B_T \to \gamma > 0 \text{ a.s. } [P_{\theta_0}] \text{ as } T \to \infty. \quad (7.49)$$

Further suppose that $\lambda(.)$ is a prior probability density on Θ which is continuous and positive in an open neighborhood of θ_0, the true parameter, and

$$\int_{-\infty}^{\infty} |\theta|^m \lambda(\theta) d\theta < \infty \quad (7.50)$$

for some integer $m \geq 0$. Then

$$\lim_{T\to\infty}\int_{-\infty}^{\infty} |t|^m \left|p^*(t|X^T)-\left(\frac{\gamma}{2\pi}\right)^{1/2}\exp\left(-\frac{1}{2}\gamma t^2\right)\right| dt = 0 \text{ a.s. } [P_{\theta_0}]. \quad (7.51)$$

In particular, choosing $m = 0$, we obtain that

$$\lim_{T\to\infty}\int_{-\infty}^{\infty} \left|p^*(t|X^T)-\left(\frac{\gamma}{2\pi}\right)^{1/2}\exp\left(-\frac{1}{2}\gamma t^2\right)\right| dt = 0 \text{ a.s. } [P_{\theta_0}] \quad (7.52)$$

whenever conditions (7.48) and (7.49) hold. This is the analogue of the Bernstein–von Mises theorem for a class of diffusion fields proved in Prakasa Rao (1984) and it shows the asymptotic convergence in L_1-mean of the posterior density to the Gaussian distribution.

As a corollary to Theorem 7.7, we also obtain that the conditional expectation, under P_{θ_0}, of $[I_T^{-1}(\hat{\theta}_T - \theta)]^m$ converges to the corresponding m-th absolute moment of the normal distribution with mean zero and variance γ^{-1}.

We define a *regular Bayes estimator* of θ, corresponding to a prior probability density $\lambda(\theta)$ and the loss function $L(\theta, \phi)$, based on the observation X^T, as an estimator which minimizes the posterior risk

$$B_T(\phi) \equiv \int_{-\infty}^{\infty} L(\theta, \phi) p(\theta|X^T) d\theta \quad (7.53)$$

over all the estimators ϕ of θ. Here $L(\theta, \phi)$ is a loss function defined on $\Theta \times \Theta$.

Suppose there exists a measurable regular Bayes estimator $\tilde{\theta}_T$ for the parameter θ (cf. Prakasa Rao (1987), Theorem 3.1.3). Suppose that the loss function $L(\theta, \phi)$ satisfies the following conditions:

$$L(\theta, \phi) = \ell(|\theta - \phi|) \geq 0 \tag{7.54}$$

and the function $\ell(t)$ is nondecreasing for $t \geq 0$. An example of such a loss function is $L(\theta, \phi) = |\theta - \phi|$. Suppose there exist nonnegative functions $R(t)$, $K(t)$ and $G(t)$ such that

(D1) $R(t)\ell(tI_T) \leq G(t)$ for all $T \geq 0$ and

(D2) $R(t)\ell(tI_T) \to K(t)$ as $T \to \infty$ uniformly on bounded intervals of t. Further suppose that the function

(D3) $\int_{-\infty}^{\infty} K(t+h)\exp(-\frac{1}{2}\beta t^2)\,dt$ has a strict minimum at $h = 0$ and

(D4) the function $G(t)$ satisfies the conditions similar to (7.39) and (7.42).

We have the following result giving the asymptotic properties of the Bayes risk of the estimator $\tilde{\theta}_T$.

Theorem 7.8: Suppose the conditions (7.48) and (7.49) in Theorem 7.7 hold and that $\lambda(.)$ is a prior probability density on Θ which is continuous and positive in an open neighborhood of θ_0, the true parameter. Further suppose that conditions (D1) to (D4) stated above hold. Then

$$I_T^{-1}(\tilde{\theta}_T - \hat{\theta}_T) \to 0 \text{ a.s. } [P_{\theta_0}] \text{ as } T \to \infty \tag{7.55}$$

and

$$\begin{aligned}\lim_{T\to\infty} R(T)B_T(\tilde{\theta}_T) &= \lim_{T\to\infty} R(T)B_T(\hat{\theta}_T)\\ &= \left(\frac{\gamma}{2\pi}\right)^{1/2}\int_{-\infty}^{\infty} K(t)\exp\left(-\frac{1}{2}\gamma t^2\right)\,dt \text{ a.s. } [P_{\theta_0}].\end{aligned} \tag{7.56}$$

We omit the proof of this theorem as it is similar to the proof of the result given in the one-dimensional case in Chapter 1 (cf. Theorem 4.1 in Borwanker *et al.* (1971)). We observed earlier that

$$I_T^{-1}(\hat{\theta}_T - \theta_0) \to \frac{Z}{\eta} \text{ in law as } T \to \infty. \tag{7.57}$$

As a consequence of Theorem 7.8, we obtain that

$$\tilde{\theta}_T \to \theta_0 \quad \text{a.s.} \quad [P_{\theta_0}] \quad \text{as} \quad T \to \infty \tag{7.58}$$

and

$$I_T^{-1}(\tilde{\theta}_T - \theta_0) \to \frac{Z}{\eta} \quad \text{in law} \quad \text{as} \quad T \to \infty. \tag{7.59}$$

In other words, the limiting distribution of the Bayes estimator is the same as that of the MLE after suitable normalization. The asymptotic Bayes risk of the estimator is given by Theorem 7.8.

8

Parametric estimation for processes driven by infinite-dimensional fBm

8.1 Introduction

We studied parametric and nonparametric inference for stochastic processes driven by fBm in the earlier chapters. We will now study similar problems for processes driven by infinite-dimensional fBm following the works of Prakasa Rao (2004e) and Cialenco *et al.* (2008).

8.2 Parametric estimation for SPDEs driven by infinite-dimensional fBm

Kallianpur and Xiong (1995) discussed the properties of solutions of stochastic partial differential equations (SPDEs) driven by infinite-dimensional fBm. They indicate that SPDEs are being used for stochastic modeling, for instance, for the study of neuronal behavior in neurophysiology and in building stochastic models of turbulence. The theory of SPDEs is investigated in Ito (1984), Rozovskii (1990) and Da Prato and Zabczyk (1992). Huebner *et al.* (1993) started the investigation of maximum likelihood estimation of parameters of two types of SPDEs and extended their results for a class of parabolic SPDEs in Huebner and Rozovskii (1995). Asymptotic properties of Bayes estimators for such problems were discussed in Prakasa Rao (2000). A short review and a comprehensive survey of these results are given in Prakasa Rao (2001, 2002). Our aim in this section is to study the problems of parameter estimation for some SPDEs driven by infinite-dimensional fBm.

SPDE with linear drift (absolutely continuous case)

Let U be a real separable Hilbert space and Q be a self-adjoint positive operator. Further suppose that the operator Q is nuclear. Then Q admits a sequence of eigenvalues $\{q_n, n \geq 1\}$ with $0 < q_n$ decreasing to zero as $n \to \infty$ and $\sum_{n=0}^{\infty} q_n < \infty$. In addition, the corresponding eigenvectors $\{e_n, n \geq 1\}$ form an orthonormal basis in U. We define *infinite-dimensional fractional Brownian motion* on U with covariance Q as

$$\mathcal{W}_Q^H(t) = \sum_{n=0}^{\infty} \sqrt{q_n} e_n W_n^H(t) \tag{8.1}$$

where $W_n^H, n \geq 1$, are real independent fBms with Hurst index H(cf. Tindel *et al.* (2003)). A formal definition is given in the next section.

Let $U = L_2[0, 1]$ and $\mathcal{W}_Q^H$ be infinite-dimensional fBm on U with Hurst index H and with the nuclear covariance operator Q.

Consider the process $u_\varepsilon(t, x), 0 \leq x \leq 1, 0 \leq t \leq T$, governed by the SPDE

$$du_\varepsilon(t, x) = (\triangle u_\varepsilon(t, x) + \theta u_\varepsilon(t, x))dt + \varepsilon d\mathcal{W}_Q^H(t, x) \tag{8.2}$$

where $\triangle = \partial^2/\partial x^2$. Suppose that $\varepsilon \to 0$ and $\theta \in \Theta \subset R$. Suppose also that the initial and boundary conditions are given by

$$u_\varepsilon(0, x) = f(x),\ f \in L_2[0, 1] \tag{8.3}$$

$$u_\varepsilon(t, 0) = u_\varepsilon(t, 1) = 0,\ 0 \leq t \leq T. \tag{8.4}$$

Let us consider a special covariance operator Q with $e_k = \sin k\pi, k \geq 1$, and $\lambda_k = (\pi k)^2, k \geq 1$. Then $\{e_k\}$ is a complete orthonormal system with eigenvalues $q_i = (1 + \lambda_i)^{-1}, i \geq 1$, for the operator Q and $Q = (I - \triangle)^{-1}$.

Tindel *et al.* (2003) have given sufficient conditions for the existence and square integrability of a solution $u_\varepsilon(t, x)$ for a SDE driven by infinite-dimensional fBm. However, their results are not directly applicable to the problem under discussion.

We assume that sufficient conditions hold so that there exists a unique square integrable solution $u_\varepsilon(t, x)$ of (8.2) under the conditions (8.3) and (8.4) and consider it as a formal sum

$$u_\varepsilon(t, x) = \sum_{i=1}^{\infty} u_{i\varepsilon}(t) e_i(x). \tag{8.5}$$

It can be checked that the Fourier coefficient $u_{i\varepsilon}(t)$ satisfies the SDE

$$du_{i\varepsilon}(t) = (\theta - \lambda_i)u_{i\varepsilon}(t)dt + \frac{\varepsilon}{\sqrt{\lambda_i + 1}}dW_i^H(t),\ 0 \le t \le T \tag{8.6}$$

with the initial condition

$$u_{i\varepsilon}(0) = v_i,\ v_i = \int_0^1 f(x)e_i(x)dx. \tag{8.7}$$

Let $P_\theta^{(\varepsilon)}$ be the probability measure generated by u_ε when θ is the true parameter. Suppose θ_0 is the true parameter. Observe that the process $\{u_{i\varepsilon}(t), 0 \le t \le T\}$ is a fractional Ornstein–Uhlenbeck-type process studied in Chapter 3 (cf. Kleptsyna and Le Breton (2002a) and Prakasa Rao (2003, 2004a,b)).

Following the notation given in Chapter 1, we define

$$M_i^H(t) = \int_0^t k_H(t,s)dW_i^H(s),\ 0 \le t \le T, \tag{8.8}$$

$$Q_{i\varepsilon}(t) = \frac{\sqrt{\lambda_i + 1}}{\varepsilon}\frac{d}{dw_t^H}\int_0^t k_H(t,s)u_{i\varepsilon}(s)ds,\ t \in [0,T], \tag{8.9}$$

$$Z_{i\varepsilon}(t) = (\theta - \lambda_i)\int_0^t Q_{i\varepsilon}(s)dw_s^H + M_i^H(t) \tag{8.10}$$

and it follows that

$$u_{i\varepsilon}(t) = \int_0^t K_H^{f_{i\varepsilon}}(t,s)dZ_{i\varepsilon}(t),\quad f_{i\varepsilon}(t) \equiv \frac{\varepsilon}{\sqrt{\lambda_i + 1}} \tag{8.11}$$

and $K_H^f(t,s)$ is as defined in (1.75) of Chapter 1. Then M_i^H is a zero-mean Gaussian martingale. Furthermore, it follows, by Theorem 1.19 in Chapter 1, that the process $\{Z_{i\varepsilon}(t)\}$ is a semimartingale and the natural filtrations $(\mathcal{Z}_{i\varepsilon_t})$ and $(\mathcal{U}_{i\varepsilon_t})$ of the processes $Z_{i\varepsilon}$ and $u_{i\varepsilon}$ respectively coincide. Let $P_{i\theta}^{T,\varepsilon}$ be the probability measure generated by the process $\{u_{i\varepsilon}(t), 0 \le t \le T\}$ when θ is the true parameter. Let θ_0 be the true parameter. It follows, by the Girsanov-type theorem discussed in Theorem 1.20 in Chapter 1, that

$$\log\frac{dP_{i\theta}^{T,\varepsilon}}{dP_{i\theta_0}^{T,\varepsilon}} = \frac{\lambda_i + 1}{\varepsilon^2}\left\{(\theta - \theta_0)\int_0^T Q_{i\varepsilon}(t)dZ_{i\varepsilon}(t) \right.$$
$$\left. -\frac{1}{2}[(\theta - \lambda_i)^2 - (\theta_0 - \lambda_i)^2]\int_0^T Q_{i\varepsilon}^2(t)dw_t^H\right\}. \tag{8.12}$$

Let $u_\varepsilon^N(t,x)$ be the projection of the solution $u_\varepsilon(t,x)$ onto the subspace spanned by the eigenvectors $\{e_i, 1 \le i \le N\}$. Then

$$u_\varepsilon^N(t,x) = \sum_{i=1}^{N} u_{i\varepsilon}(t)e_i(x). \tag{8.13}$$

From the independence of the processes $W_i^H, 1 \le i \le N$, and hence of the processes $u_{i\varepsilon}, 1 \le i \le N$, it follows that the Radon–Nikodym derivative, of the probability measure $P_\theta^{N,T,\varepsilon}$ generated by the process $u_\varepsilon^N, 0 \le t \le T$, when θ is the true parameter with respect to the probability measure $P_{\theta_0}^{N,T,\varepsilon}$ generated by the process $u_\varepsilon^n, 0 \le t \le T$, when θ_0 is the true parameter, is given by

$$\log \frac{dP_\theta^{N,T,\varepsilon}}{dP_{\theta_0}^{N,T,\varepsilon}}(u_\varepsilon^N) = \sum_{i=1}^{N} \frac{\lambda_i+1}{\varepsilon^2}\left\{(\theta-\theta_0)\int_0^T Q_{i\varepsilon}(t)dZ_{i\varepsilon}(t)\right.$$
$$\left. -\frac{1}{2}[(\theta-\lambda_i)^2-(\theta_0-\lambda_i)^2]\int_0^T Q_{i\varepsilon}^2(t)dw_t^H\right\}. \tag{8.14}$$

Furthermore, the Fisher information is given by

$$\begin{aligned} I_{N\varepsilon}(\theta) &= E_\theta\left[\frac{\partial \log(dP_\theta^{N,T\varepsilon}/dP_{\theta_0}^{N,T,\varepsilon})}{\partial\theta}\right]^2 \\ &= \sum_{i=1}^{N}\frac{\lambda_i+1}{\varepsilon^2}E_\theta\left[\int_0^T Q_{i\varepsilon}^2(t)dw_t^H\right]. \end{aligned} \tag{8.15}$$

It is easy to check that the MLE $\hat{\theta}_{N,\varepsilon}$ of the parameter θ based on the projection u_ε^N of u_ε is given by

$$\hat{\theta}_{N,\varepsilon} = \frac{\sum_{i=1}^N(\lambda_i+1)\int_0^T Q_{i\varepsilon}(t)dZ_{i\varepsilon}(t)}{\sum_{i=1}^N(\lambda_i+1)\int_0^T Q_{i\varepsilon}^2(t)dw_t^H}. \tag{8.16}$$

Suppose θ_0 is the true parameter. It is easy to see that

$$\varepsilon^{-1}(\hat{\theta}_{N,\varepsilon}-\theta_0) = \frac{\sum_{i=1}^N\sqrt{\lambda_i+1}\int_0^T Q_{i\varepsilon}(t)dM_i^H(t)}{\sum_{i=1}^N(\lambda_i+1)\int_0^T Q_{i\varepsilon}^2(t)dw_t^H}. \tag{8.17}$$

Observe that $M_i, 1 \le i \le N$, are independent zero-mean Gaussian martingales with $\langle M_i\rangle = w^H, 1 \le i \le N$.

Theorem 8.1: The MLE $\hat{\theta}_{N,\varepsilon}$ is strongly consistent, that is,

$$\hat{\theta}_{N,\varepsilon} \to \theta_0 \ \text{ a.s. } [P_{\theta_0}] \text{ as } \varepsilon \to 0 \tag{8.18}$$

provided

$$\sum_{i=1}^{N} \int_0^T (\lambda_i + 1) Q_{i\varepsilon}^2(t) dw_t^H \to \infty \text{ a.s. } [P_{\theta_0}] \text{ as } \varepsilon \to 0. \tag{8.19}$$

Proof: This theorem follows by observing that the process

$$R_\varepsilon^N \equiv \sum_{i=1}^{N} \int_0^T \varepsilon \sqrt{\lambda_i + 1} Q_{i\varepsilon}(t) dM_i^H(t),\ T \geq 0 \tag{8.20}$$

is a local martingale with the quadratic variation process

$$\langle R_\varepsilon^N \rangle_T = \sum_{i=1}^{N} \int_0^T \varepsilon^2 (\lambda_i + 1) Q_{i\varepsilon}^2(t) dw_t^H \tag{8.21}$$

and applying the strong law of large numbers (cf. Liptser (1980), Prakasa Rao (1999b), p. 61) under the condition (8.19) stated above.

Limiting distribution

We now discuss the limiting distribution of the MLE $\hat{\theta}_{N\varepsilon}$ as $\varepsilon \to 0$.

Theorem 8.2: Assume that the process $\{R_\varepsilon^N, \varepsilon \geq 0\}$ is a local continuous martingale and that there exists a norming function $I_\varepsilon^N, \varepsilon \geq 0$, such that

$$(I_\varepsilon^N)^2 \langle R_\varepsilon^N \rangle_T = (I_\varepsilon^N)^2 \sum_{i=1}^{N} \int_0^T \varepsilon^2 (\lambda_i + 1) Q_{i\varepsilon}^2(t)\ dw_t^H \to \eta^2$$
$$\text{in probability as } \varepsilon \to 0 \tag{8.22}$$

where $I_\varepsilon^N \to 0$ as $\varepsilon \to 0$ and η is a random variable such that $P(\eta > 0) = 1$. Then

$$(I_\varepsilon^N R_\varepsilon^N, (I_\varepsilon^N)^2 \langle R_\varepsilon^N \rangle_T) \to (\eta Z, \eta^2) \text{ in law as } \epsilon \to 0 \tag{8.23}$$

where the random variable Z has the standard Gaussian distribution and the random variables Z and η are independent.

Proof: This theorem follows as a consequence of the central limit theorem for local martingales (cf. Theorem 1.49 and Remark 1.47 in Prakasa Rao (1999b), p. 65).

Observe that

$$(I_\varepsilon^N)^{-1} (\hat{\theta}_{N\varepsilon} - \theta_0) = \frac{I_\varepsilon^N R_\varepsilon^N}{(I_\varepsilon^N)^2 \langle R_\varepsilon^N \rangle}. \tag{8.24}$$

Applying the above theorem, we obtain the following result.

Theorem 8.3: Suppose the conditions stated in Theorem 8.2 hold. Then

$$(I_\varepsilon^N)^{-1}(\hat{\theta}_{N\varepsilon} - \theta_0) \to \frac{Z}{\eta} \text{ in law as } \varepsilon \to 0 \tag{8.25}$$

where the random variable Z has the standard Gaussian distribution and the random variables Z and η are independent.

Remarks:

(i) If the random variable η is a constant with probability one, then the limiting distribution of the MLE is Gaussian with mean zero and variance η^{-2}. Otherwise it is a mixture of the Gaussian distributions with mean zero and variance η^{-2} with the mixing distribution as that of η.

(ii) Suppose that

$$\lim_{N\to\infty} \lim_{\varepsilon\to 0} \varepsilon^2 I_{N\varepsilon}\theta = I(\theta) \tag{8.26}$$

exists and is positive. Since the sequence of Radon–Nikodym derivatives

$$\left\{ \frac{dP_\theta^{N,T,\varepsilon}}{dP_{\theta_0}^{N,T,\varepsilon}}, n \geq 1 \right\}$$

form a nonnegative martingale with respect to the filtration generated by the sequence of random variables $\{u_\varepsilon^N, N \geq 1\}$, it converges a.s. to a random variable $\nu_{\varepsilon,\theta,\theta_0}$ as $N \to \infty$ for every $\varepsilon > 0$. It is easy to see that the limiting random variable is given by

$$\nu_{\varepsilon,\theta,\theta_0}(u_\varepsilon) = \exp\left\{ \sum_{i=1}^{\infty} \frac{\lambda_i + 1}{\varepsilon^2} \left[(\theta - \theta_0) \int_0^T Q_{i\varepsilon}(t) dZ_{i\varepsilon}(t) \right.\right.$$
$$\left.\left. -\frac{1}{2}[(\theta - \lambda_i)^2 - (\theta_0 - \lambda_i)^2] \int_0^T Q_{i\varepsilon}^2(t) dw_t^H \right] \right\}. \tag{8.27}$$

Furthermore, the sequence of random variables $u_\varepsilon^N(t)$ converges in probability to the random variable $u_\varepsilon(t)$ as $N \to \infty$ for every $\varepsilon > 0$. Hence, by Lemma 4 in Skorokhod (1965, p. 100), it follows that the measures P_θ^ε generated by the processes u_ε for different values of θ are absolutely continuous with respect to each other and the Radon–Nikodym derivative of the probability measure P_θ^ε

with respect to the probability measure $P^{\varepsilon}_{\theta_0}$ is given by

$$\frac{dP^{\varepsilon}_{\theta}}{dP^{\varepsilon}_{\theta_0}}(u_\varepsilon) = \nu_{\varepsilon,\theta,\theta_0}(u_\varepsilon)$$

$$= \exp\left\{\sum_{i=1}^{\infty}\frac{\lambda_i+1}{\varepsilon^2}\left[(\theta-\theta_0)\int_0^T Q_{i\varepsilon}(t)dZ_{i\varepsilon}(t)\right.\right.$$
$$\left.\left. -\frac{1}{2}[(\theta-\lambda_i)^2-(\theta_0-\lambda_i)^2]\int_0^T Q^2_{i\varepsilon}(t)dw^H_t\right]\right\}. \quad (8.28)$$

It can be checked that the MLE $\hat{\theta}_\varepsilon$ of θ based on u_ε satisfies the likelihood equation

$$\alpha_\varepsilon = \varepsilon^{-1}(\hat{\theta}_\varepsilon - \theta_0)\beta_\varepsilon \quad (8.29)$$

when θ_0 is the true parameter, where

$$\alpha_\varepsilon = \sum_{i=1}^{\infty}\sqrt{\lambda_i+1}\int_0^T Q_{i\varepsilon}(t)dM_i(t) \quad (8.30)$$

and

$$\beta_\varepsilon = \sum_{i=1}^{\infty}(\lambda_i+1)\int_0^T Q^2_{i\varepsilon}(t)dw^H_t. \quad (8.31)$$

One can obtain sufficient conditions for studying the asymptotic behavior of the estimator $\hat{\theta}_\varepsilon$ as in the finite projection case discussed above. We omit the details.

SPDE with linear drift (singular case)

Let $(\Omega, \mathcal{F}, P)$ be a probability space and consider the process $u_\varepsilon(t,x), 0 \le x \le 1, 0 \le t \le T$, governed by the SPDE

$$du_\varepsilon(t,x) = \theta\, \triangle u_\varepsilon(t,x)dt + \varepsilon(I-\triangle)^{-1/2}dW(t,x) \quad (8.32)$$

where $\theta > 0$ satisfies the initial and boundary conditions

$$u_\varepsilon(0,x) = f(x),\ 0 < x < 1,\ f \in L_2[0,1], \quad (8.33)$$
$$u_\varepsilon(t,0) = u_\varepsilon(t,1) = 0,\ 0 \le t \le T.$$

Here I is the identity operator, $\triangle = \partial^2/\partial x^2$ as defined above and the process $W(t,x)$ is cylindrical infinite-dimensional fBm with $H \in [\frac{1}{2}, 1)$ (cf. Tindel *et al.* (2003)). If $H = \frac{1}{2}$, then cylindrical infinite-dimensional fBm reduces to cylindrical infinite-dimensional Brownian motion and the solution for such a SPDE generates measures which are singular with respect to each other (cf. Huebner *et al.* (1993)). In analogy with this, we have designated this problem as a singular case. Following the discussion in the previous section, we assume the existence of a square integrable solution $u_\varepsilon(t,x)$ for Equation (8.32) subject to the boundary conditions (8.33). Then the Fourier coefficients $u_{i\varepsilon}(t)$ satisfy the SDEs

$$du_{i\varepsilon}(t) = -\theta\lambda_i u_{i\varepsilon}(t)dt + \frac{\varepsilon}{\sqrt{\lambda_i+1}}dW_i^H(t),\ 0 \le t \le T \tag{8.34}$$

with

$$u_{i\varepsilon}(0) = v_i,\ v_i = \int_0^1 f(x)e_i(x)dx. \tag{8.35}$$

Let $u_\varepsilon^{(N)}(t,x)$ be the projection of $u_\varepsilon(t,x)$ onto the subspace spanned by $\{e_1,\ldots,e_N\}$ in $L_2[0,1]$. In other words,

$$u_\varepsilon^{(N)}(t,x) = \sum_{i=1}^N u_{i\varepsilon}(t)e_i(x). \tag{8.36}$$

Let $P_\theta^{(\varepsilon,N)}$ be the probability measure generated by $u_\varepsilon^{(N)}$ on the subspace spanned by $\{e_1,\ldots,e_N\}$ in $L_2[0,1]$. It can be shown that the measures $\{P_\theta^{(\varepsilon,N)}, \theta\in\Theta\}$ form an equivalent family and

$$\begin{aligned}\log \frac{dP_\theta^{(\varepsilon,N)}}{dP_{\theta_0}^{(\varepsilon,N)}}(u_\varepsilon^{(N)}) &= -\frac{1}{\varepsilon^2}\sum_{i=1}^N \lambda_i(\lambda_i+1)\left[(\theta-\theta_0)\int_0^T Q_{i\varepsilon}(t)dZ_{i\varepsilon}(t)\right.\\ &\left.-\frac{1}{2}(\theta-\theta_0)^2\lambda_i\int_0^T Q_{i\varepsilon}^2(t)dw_t^H\right].\end{aligned} \tag{8.37}$$

It can be checked that the MLE $\hat\theta_{\varepsilon,N}$ of θ based on $u_\varepsilon^{(N)}$ satisfies the likelihood equation

$$\alpha_{\varepsilon,N} = -\varepsilon^{-1}(\hat\theta_{\varepsilon,N}-\theta_0)\beta_{\varepsilon,N} \tag{8.38}$$

when θ_0 is the true parameter, where

$$\alpha_{\varepsilon,N} = \sum_{i=1}^N \lambda_i\sqrt{\lambda_i+1}\int_0^T Q_{i\varepsilon}(t)dM_i(t) \tag{8.39}$$

and

$$\beta_{\varepsilon,N} = \sum_{i=1}^N(\lambda_i+1)\lambda_i^2\int_0^T Q_{i,\varepsilon}^2(t)dw_t^H. \tag{8.40}$$

Asymptotic properties of these estimators can be investigated as in the previous example. We do not go into the details as the arguments are similar. For more details, see Prakasa Rao (2004e).

Remarks: One can study the local asymptotic mixed normality (LAMN) of the family of probability measures generated by the log-likelihood ratio processes by standard arguments as in Prakasa Rao (1999b) and hence investigate the asymptotic efficiency of the MLE using Hajek–Lecam-type bounds.

8.3 Parametric estimation for stochastic parabolic equations driven by infinite-dimensional fBm

We now discuss some recent work of Cialenco *et al.* (2008) dealing with problems of estimation in models more general than those discussed in the previous section. First, we introduce some notation.

Let $\mathbf{H}$ be a separable Hilbert space with the inner product $(.,.)_0$ and with the corresponding norm $||.||_0$. Let Λ be a densely defined linear operator on $\mathbf{H}$ with the property that there exists $c > 0$ such that

$$||\Lambda u||_0 \geq c||u||_0$$

for every u in the domain of the operator Λ. The operator powers $\Lambda^\gamma, \gamma \in R$, are well defined and generate the spaces $\mathbf{H}^\gamma$ with the properties (i) for $\gamma > 0$, $\mathbf{H}^\gamma$ is the domain of Λ^γ; (ii) $\mathbf{H}^0 = \mathbf{H}$; and (iii) for $\gamma < 0$, $\mathbf{H}^\gamma$ is the completion of $\mathbf{H}$ with respect to the norm $||.||_\gamma \equiv ||\Lambda^\gamma.||_0$ (cf. Krein *et al.* (1982)). The family of spaces $\{\mathbf{H}^\gamma, \gamma \in R\}$ has the following properties:

(i) $\Lambda^\gamma(\mathbf{H}^r) = \mathbf{H}^{r-\gamma},\ \gamma, r \in R$.

(ii) For $\gamma_1 < \gamma_2$, the space $\mathbf{H}^{\gamma_2}$ is densely and continuously embedded into $\mathbf{H}^{\gamma_1}$, that is, $\mathbf{H}^{\gamma_2} \subset \mathbf{H}^{\gamma_1}$, and there exists a constant $c_{12} > 0$ such that $||u||_{\gamma_1} \leq c_{12}||u||_{\gamma_2}$.

(iii) For every $\gamma \in R$ and $m > 0$, the space $\mathbf{H}^{\gamma-m}$ is the dual of the space $\mathbf{H}^{\gamma+m}$ with respect to the inner product in $\mathbf{H}^\gamma$, with duality $\langle .,. \rangle_{\gamma,m}$ given by

$$\langle u_1, u_2 \rangle_{\gamma,m} = (\Lambda^{\gamma-m}u_1, \Lambda^{\gamma+m}u_2)_0, \quad u_1 \in \mathbf{H}^{\gamma-m}, u_2 \in \mathbf{H}^{\gamma+m}.$$

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\{W_j^H, j \geq 1\}$ be a family of independent standard fBms on this space with the *same* Hurst index H in (0, 1).

Consider the SDE

$$du(t) + (\mathcal{A}_0 + \theta\mathcal{A}_1)u(t)dt = \sum_{j\geq 1} g_j(t)dW_j^H(t),\ 0 \leq t \leq T,\ u(0) = u_0 \quad (8.41)$$

where $\mathcal{A}_0, \mathcal{A}_1$ are linear operators, $g_j, j \geq 1$, are non-random and $\theta \in \Theta \subset R$. Equation (8.41) is called *diagonalizable* if the operators $\mathcal{A}_0, \mathcal{A}_1$ have the same system of eigenfunctions $\{h_j, j \geq 1\}$ such that $\{h_j, j \geq 1\}$ is an orthonormal basis in $\mathbf{H}$ and each h_j belongs to $\cap_{\gamma \in R} \mathbf{H}^\gamma$. It is called (m, γ)-parabolic for some $m \geq 0, \gamma \in R$, if:

(i) the operator $\mathcal{A}_0 + \theta\mathcal{A}_1$ is uniformly bounded from $\mathbf{H}^{\gamma+m}$ to $\mathbf{H}^{\gamma-m}$ for every $\theta \in \Theta$, that is, there exists $C_1 > 0$ such that

$$||\mathcal{A}_0 + \theta\mathcal{A}_1||_{\gamma-m} \leq C_1||v||_{\gamma+m}, \quad \theta \in \Theta, v \in H^{\gamma+m}; \tag{8.42}$$

(ii) there exists a $\delta > 0$ and $C \in R$ such that

$$-2\langle(\mathcal{A}_0 + \theta\mathcal{A}_1)v, v\rangle_{\gamma,m} + \delta||v||^2_{\gamma+m} \leq C||v||^2_\gamma, \ v \in \mathbf{H}^{\gamma+m}, \theta \in \Theta. \tag{8.43}$$

If Equation (8.41) is (m, γ)-parabolic, then the condition (8.43) implies that

$$\langle(2\mathcal{A}_0 + 2\theta\mathcal{A}_1 + CI)v, v\rangle_{\gamma,m} \geq \delta||v||_{\gamma+m}$$

where I is the identity operator. The Cauchy–Schwartz inequality and the continuous embedding of $\mathbf{H}^{\gamma+m}$ into $\mathbf{H}^\gamma$ will imply that

$$||(2\mathcal{A}_0 + 2\theta\mathcal{A}_1 + CI)v||_\gamma \geq \delta_1||v||_\gamma$$

for some $\delta_1 > 0$ uniformly in $\theta \in \Theta$.

Let us choose $\Lambda = [2\mathcal{A}_0 + 2\theta_0\mathcal{A}_1 + CI]^{1/(2m)}$ for some fixed $\theta_0 \in \Theta$. If the operator $\mathcal{A}_0 + \theta\mathcal{A}_1$ is unbounded, we say that $\mathcal{A}_0 + \theta\mathcal{A}_1$ has order $2m$ and Λ has order 1. If Equation (8.41) is (m, γ)-parabolic and diagonalizable, we will assume that the operator Λ has the same eigenfunctions as the operators $\mathcal{A}_0$ and $\mathcal{A}_1$. This is justified by the comments made above.

Suppose Equation (8.41) is diagonalizable. Then there exists eigenvalues $\{\rho_j, j \geq 1\}$ and $\{\nu_j, j \geq 1\}$ such that

$$\mathcal{A}_0 h_j = \rho_j h_j, \ \mathcal{A}_1 h_j = \nu_j h_j.$$

Without loss of generality, we can also assume that there exists $\{\lambda_j, j \geq 1\}$ such that

$$\Lambda h_j = \lambda_j h_j.$$

Following the arguments in Cialenco *et al.* (2008), it can be shown that Equation (8.41) is (m, γ)-parabolic if and only if there exists $\delta > 0, C_1 > 0$ and $C_2 \in R$ such that, for all $j \geq 1, \theta \in \Theta$,

$$|\rho_j + \theta\nu_j| \leq C_1\lambda_j^{2m} \tag{8.44}$$

and

$$-2(\rho_j + \theta\nu_j) + \delta\lambda_j^{2m} \leq C_2. \tag{8.45}$$

As the conditions in (8.44) and (8.45) do not depend on γ, we conclude that a diagonalizable equation (8.41) is (m,γ)-parabolic for some γ if and only if it is (m,γ)-parabolic for every γ. Hereafter we will say that Equation (8.41) is m-parabolic. We will assume that it is diagonalizable and fix the basis $\{h_j, j \geq 1\}$ in $\mathbf{H}$ consisting of the eigenfunctions of $\mathcal{A}_0$, $\mathcal{A}_1$ and λ. Recall that the set of eigenfunctions is the same for all three operators. Since h_j belongs to every $\mathbf{H}^\gamma$, and since $\cap_\gamma \mathbf{H}^\gamma$ is dense in $\cup_\gamma \mathbf{H}^\gamma$, every element f of $\cup_\gamma \mathbf{H}^\gamma$ has a unique expansion $\sum_{j\geq 1} f_j h_j$ where $f_j = \langle f, h_j \rangle_{0,m}$ for suitable m.

Infinite-dimensional fBm W^H is an element of $\cup_{\gamma\in R}\mathbf{H}^\gamma$ with the expansion

$$W^H(t) = \sum_{j\geq 1} W_j^H(t) h_j. \tag{8.46}$$

The solution of the diagonalizable equation

$$du(t) + (\mathcal{A}_0 + \theta\mathcal{A}_1)u(t)dt = dW^H(t),\ 0 \leq t \leq T,\ u(0) = u_0, \tag{8.47}$$

with $u_0 \in \mathbf{H}$, is a random process with values in $\cup_\gamma \mathbf{H}^\gamma$ and has an expansion

$$u(t) = \sum_{j\geq 1} u_j(t) h_j \tag{8.48}$$

where

$$u_j(t) = (u_0, h_j)e^{-(\theta\nu_j+\rho_j)} + \int_0^t e^{-(\theta\nu_j+\rho_j)(t-s)} dW_j^H(s). \tag{8.49}$$

Let

$$\mu_j(\theta) = \theta\nu_j + \rho_j,\ j \geq 1. \tag{8.50}$$

In view of (8.45), we get that there exists a positive integer J such that

$$\mu_j(\theta) > 0 \ \text{ for } \ j \geq J \tag{8.51}$$

if Equation (8.41) is m-parabolic and diagonalizable.

Theorem 8.4: Suppose that $H \geq \frac{1}{2}$ and Equation (8.41) is m-parabolic and diagonizable. Further suppose that there exists a positive real number γ such that

$$\sum_{j\geq 1} (1 + |\mu_j(\theta)|)^{-\gamma} < \infty. \tag{8.52}$$

Then, for every $t > 0$, $W^H(t) \in L_2(\Omega, \mathbf{H}^{-m\gamma})$ and $u(t) \in L_2(\Omega, \mathbf{H}^{-m\gamma+2mH})$.

For proof, see Cialenco *et al.* (2008).

Maximum likelihood estimation

Consider the diagonalizable equation

$$du(t) + (\mathcal{A}_0 + \theta\mathcal{A}_1)u(t)dt = dW^H(t),\ u(0) = 0,\ 0 \le t \le T \tag{8.53}$$

with

$$u(t) = \sum_{j\ge 1} u_j(t)h_j \tag{8.54}$$

as given by (8.49). Suppose the process $\{u_i(t), 0 \le t \le T\}$ can be observed for $i = 1, \ldots, N$. The problem is to estimate the parameter θ using these observations.

Note that $\mu_j(\theta) = \rho_j + \nu_j\theta$, where ρ_j and ν_j are the eigenvalues of $\mathcal{A}_0$ and $\mathcal{A}_1$ respectively. Furthermore, each process u_j is a fractional Ornstein–Uhlenbeck-type process satisfying the SDE

$$du_j(t) = -\mu_j(\theta)u_j(t)dt + dW_j^H(t),\ u_j(0) = 0,\ 0 \le t \le T. \tag{8.55}$$

Since the processes $\{W_j^H, j \ge 1\}$ are independent, it follows that the process $\{u_j, 1 \le j \le N\}$ is independent. Following the notation introduced in Chapter 1, let

$$M_j^H(t) = \int_0^t k_H(t,s)dw_j^H(s),\ Q_j(t) = \frac{d}{dw_j^H(t)}\int_0^t k_H(t,s)u_j(s)ds \tag{8.56}$$

and

$$Z_j(t) = \int_0^t k_H(t,s)du_j(s) \tag{8.57}$$

for $j = 1, \ldots, N$. Applying the Girsanov-type formula as given in Chapter 1, it can be shown that the measure generated by the process $(u_1, \ldots, u_N)$ is absolutely continuous with respect to the measure generated by the process $(W_1^H, \ldots, W_N^H)$ and their Radon–Nikodym derivative is given by

$$\exp\left\{-\sum_{j=1}^N \mu_j(\theta)\int_0^T Q_j(s)dZ_j(s) - \sum_{j=1}^N \frac{[\mu_j(\theta)]^2}{2}\int_0^T Q_j^2(s)dw_H(s)\right\}. \tag{8.58}$$

Maximizing this function with respect to the parameter θ, we get the MLE

$$\hat{\theta}_N = -\frac{\sum_{j=1}^N \int_0^T \nu_j Q_j(s)(dZ_j(s) + \rho_j Q_j(s)dw_H(s))}{\sum_{j=1}^N \int_0^T \nu_j^2 Q_j^2(s)dw_H(s)}.$$

Theorem 8.5: Under the assumptions stated in Theorem 8.4, the following conditions are equivalent:

(i) $\sum_{j=1}^{\infty} \nu_j^2/\mu_j(\theta) = \infty$

(ii) $\lim_{N\to\infty} \hat{\theta}_N = \theta$ a.s.

when θ is the true parameter. Furthermore, if the condition (i) holds, then

$$\lim_{N\to\infty} \left(\sum_{j=J}^{N} \frac{\nu_j^2}{\mu_j(\theta)} \right)^{1/2} (\hat{\theta}_N - \theta) \xrightarrow{\mathcal{L}} N(0, \sigma^2) \text{ as } N \to \infty \tag{8.59}$$

for some σ^2 where $J = \min\{j : \mu_i(\theta) = 0 \text{ for all } i \geq j\}$.

Proof of this theorem follows by observing that

$$\lim_{j\to\infty} \mu_j(\theta) \;\; E\left[\int_0^T Q_j^2(s) dw_H(s)\right] = \frac{T}{2} \tag{8.60}$$

$$\lim_{j\to\infty} \mu_j^3(\theta) \;\; Var\left[\int_0^T Q_j^2(s) dw_H(s)\right] = \frac{T}{2} \tag{8.61}$$

and using the strong law of large numbers and the central limit theorem for sums of independent random variables. For a detailed proof, see Cialenco *et al.* (2008). Ergodicity and parameter estimation for infinite-dimensional fractional Ornstein–Uhlenbeck-type processes are investigated in Maslowski and Pospisil (2008).

9

Estimation of self-similarity index

9.1 Introduction

As we discussed in Chapter 1, self-similar processes can be thought of as processes with sample paths that retain the same general appearance irrespective of the scale of measurement or regardless of the distance from which they are observed. Estimation of self-similarity index or Hurst index is of great importance and interest. We have discussed parametric and nonparametric inference for processes driven by fBm with *known* self-similarity index in the earlier chapters. Our aim here is to discuss some methods of estimation of Hurst index H. The estimation of the constant H has been well studied in the case of long-range dependence, that is, when $\frac{1}{2} < H < 1$. This can be seen from the works of Geweke and Porter-Hudak (1983), Taylor and Taylor (1991), Constantine and Hall (1994), Chen *et al*. (1995), Robinson (1995), Abry and Sellan (1996), Comte (1996), McCoy and Walden (1996), Hall *et al*. (1997), Kent and Wood (1997) and Jensen (1998) among others. Beran and Terrin (1994) and Beran (1994) have discussed the problem of estimation of the constant H. Dieker (2004) has given a short review of some methods for estimation of the Hurst index and studied their computational efficiency. Some of these methods are intuitive without a sound theoretical justification. Our short review of these methods here is based on Dieker (2004). Techniques for estimating a constant self-similarity index H are generally based on log-linear regression between suitable variables.

In some modeling applications, treatment of the self-similarity index H as a constant may be right. However, there are many other phenomena which exhibit self-similar behavior but the nature of self-similarity changes as the phenomenon evolves. To model such data, the parameter H should be treated as a function

Statistical Inference for Fractional Diffusion Processes B. L. S. Prakasa Rao

of time and not as a constant. Goncalves and Flandrin (1993) and Flandrin and Goncalves (1994) studied a class of processes which are locally self-similar with time-dependent Hurst index. For such processes, the local Hurst index function $H(t)$ contains information about the behavior of the process. Hence it is important to study methods for estimation of the function $H(t)$. We will describe one such method due to Wang *et al*. (2001).

9.2 Estimation of the Hurst index H when H is a constant and $\frac{1}{2} < H < 1$ for fBm

Let $Y_0, Y_1, \ldots, Y_{N-1}$ be a fBm sample observed at the time points $t_0 = 0$, $t_1 = 1/N, \ldots, t_{N-1} = (N-1)/N$ respectively. Let $X_k = Y_{k+1} - Y_k, k = 0, 1, \ldots, N-2$. Then $Var(X_k) = N^{-2H}$ for $k = 0, 1, \ldots, N-2$. We now describe different methods of estimation based on a 'time domain' approach. Other methods based on a 'spectral domain' approach using a periodogram or wavelets are described in Dieker (2004).

Aggregated variance method

The aggregated variance method is based on the self-similarity of the sample. Let m be a positive integer. Divide the sequence $\{X_k\}$ into blocks of size m. The aggregated process $\{X_k^{(m)}\}$ is defined by the relation

$$X_k^{(m)} = m^{-1}(X_{km} + \cdots + X_{(k+1)m-1}).$$

In view of the self-similarity of the process Y, the process $X^{(m)} = \{X_k^{(m)}, k \geq 0\}$ has the same finite-dimensional distributions as the process $m^{H-1}X$ where $X = \{X_k, k \geq 0\}$. Therefore

$$Var(X_k^{(m)}) = m^{2H-2}Var(X_k) = m^{2H-2}N^{-2H}$$

is the same for every $k \geq 0$. Let $M =$ integer part of N/m. An estimator for $Var(X_k^{(m)})$ is

$$\widehat{Var(X_k^{(m)})} = M^{-1}\sum_{i=0}^{M-1}(X_i^{(m)} - \bar{X}^{(m)})^2 \tag{9.1}$$

where

$$\bar{X}^{(m)} = M^{-1}\sum_{i=0}^{M-1} X_i^{(m)}. \tag{9.2}$$

The estimator of H is obtained by plotting $\widehat{Var(X_k^{(m)})}$ against m on a log–log scale. If the estimates of the variance were equal to their true values, then all the points would fall on a straight line with slope $2H - 2$. The slope is estimated in practice by fitting a line through these points.

Discrete variations method

This method is due to Coeurjolly (2001). Let $\mathbf{a}$ be a filter of length $\ell + 1$ and of order $p \geq 1$, that is, a vector $\mathbf{a} = (a_0, \ldots, a_{p-1})$ of real numbers such that

$$\sum_{q=0}^{\ell} a_q q^r = 0, \quad r = 0, \ldots, p-1 \tag{9.3}$$

and

$$\sum_{q=0}^{\ell} a_q q^r \neq 0, \quad r = p. \tag{9.4}$$

Note that $Y_i = \sum_{k=0}^{i} X_k, \quad i = 0, \ldots, N-1$. Let

$$V_k^{\mathbf{a}} = \sum_{q=0}^{\ell} a_q Y_{k-q}, \quad k = \ell, \ldots, N-1. \tag{9.5}$$

An example of a filter of order 2 is $\mathbf{a} = (0, 1)$ and then $V_k^{\mathbf{a}} = X_k$. Let

$$S(k, \mathbf{a}) = \frac{1}{N - \ell} \sum_{i=\ell}^{N-1} |V_i^{\mathbf{a}}|^k \tag{9.6}$$

for some $k > 0$. Since the sequence $\{X_k, k \geq 1\}$ is a fractional Gaussian noise, it can be shown that

$$E[S(k, \mathbf{a})] = \frac{\Gamma((k+1)/2)}{\sqrt{\pi}} N^{-kH} [2 \;\; Var(N^H V_1^{\mathbf{a}})]^{k/2}. \tag{9.7}$$

This method consists of obtaining an estimator of $E[S(k, \mathbf{a})]$ and equating it to the right hand side of Equation (9.7) to solve for an estimator for H. Observe that $Var(N^H V_1^{\mathbf{a}})$ does not depend on H from the self-similarity of fBm. Coeurjolly (2001) also studied a slightly modified version of this method. Define the sequence of filters $\mathbf{a}^{(m)}, 1 \leq m \leq M,$ for some integer $M \geq 1$, by the relation

$$\begin{aligned} a_i^{(m)} &= a_j \;\text{ for }\; i = jm \\ &= 0 \quad \text{otherwise.} \end{aligned} \tag{9.8}$$

An application of (9.3) and (9.7) for $r = 0$ shows that

$$E[S(k, \mathbf{a}^{(m)})] = m^{Hk} E[S(k, \mathbf{a})]. \tag{9.9}$$

An estimator of H is obtained by linearly regressing $\{\log S(k, \mathbf{a}^{(m)}), 1 \le m \le M\}$ on $\{k \log m, 1 \le m \le M\}$. Couerjolly (2001) proved that this estimator converges at a rate $\sqrt{N}$ to the Gaussian distribution with mean zero and some variance $\sigma^2(k, \mathbf{a})$ depending on k and $\mathbf{a}$. Furthermore, the asymptotic variance is minimal when $k = 2$.

Higuchi method

Following Higuchi (1988), define

$$L(m) = \frac{N-1}{m^3} \sum_{i=1}^{m} \frac{1}{M_i} \sum_{k=1}^{M_i} \left| \sum_{j=i+(k-1)m+1}^{i+km} X_j \right| \tag{9.10}$$

where M_i is the integral part of $(N-i)/m$. The estimator of H is obtained by plotting $L(m)$ in a log–log plot versus m and adding 2 to the slope of the fitted line.

Method using variance of the regression residuals

Following Peng *et al*. (1994) and Cannon *et al*. (1997), the series $\{X_k, 1 \le k \le N-2\}$ is broken into blocks of size m. Within each block, the partial sums are regressed on a line $\hat{\alpha^{(k)}} + i\hat{\beta^{(k)}}$. The residuals of the regression are given by

$$e_i^{(k)} = \sum_{j=km}^{km+i-1} X_j - (\hat{\alpha^{(k)}} + i\hat{\beta^{(k)}}). \tag{9.11}$$

The sample variance of the residuals is then computed for each block. The average of the sample variances over all blocks is proportional to m^{2H} (cf. Taqqu *et al*. (1995)). This fact is used for estimating H.

Method of rescaled range analysis R/S

This method was suggested by Hurst (1951). The series $\{X_j, 1 \le j \le N-2\}$ is divided into K nonoverlapping blocks such that each block contains M elements where M is the integral part of N/K. Let $t_i = M(i-1)$ and

$$R(t_i, r) = \max[W(t_1, 1), \ldots, W(t_i, r)] - \min[W(t_1, 1), \ldots, W(t_i, r)]$$

where

$$W(t_i, k) = \sum_{j=0}^{k-1} X_{t_i+j} - k\left(\frac{1}{r}\sum_{j=0}^{r-1} X_{t_i+j}\right), \ k = 1, \ldots, r.$$

Note that $R(t_i, r) \geq 0$ since $W(t_i, r) = 0$ and the quantity $R(t_i, r)$ can be computed only when $t_i + r \leq N$. Define

$$S^2(t_i, r) = \frac{1}{r}\sum_{j=0}^{r-1} X_{t_i+j}^2 - \left(\frac{1}{r}\sum_{j=0}^{r-1} X_{t_i+j}\right)^2. \tag{9.12}$$

The ratio $R(t_i, r)/S(t_i, r)$ is called the rescaled adjusted range. It is computed for a number of values of r where $t_i = M(i-1)$ is the starting point of the ith block for $i = 1, \ldots, K$. Observe that, for each value of r, we obtain a number of R/S samples. The number of samples decrease as r increases. However, the resulting samples are not independent. It is believed that the R/S-statistic is proportional to r^H as $r \to \infty$ for the fractional Gaussian noise. Assuming this property, we regress $\log(R/S)$ against $\log r$ to obtain an estimator for H. For details, see Bassingthwaighte and Raymond (1994, 1995).

Poggi and Viano method

This method is due to Poggi and Viano (1998) using multi-scale aggregates. Let

$$\Delta_\tau(k) = Y_k - Y_{k-\tau}$$

be the τ-increments of fBm Y. Define

$$m_n(j) = \frac{1}{n}\sum_{k=\tau+n(j-1)}^{\tau+jn-1} \Delta_\tau(k) \tag{9.13}$$

and

$$S_N^2(n) = \frac{1}{[(N-\tau+1)/n]}\sum_{j=1}^{[(N-\tau+1)/n]} m_n^2(j). \tag{9.14}$$

Poggi and Viano (1998) proved the following under the assumption that the process Y is standard fBm:

(i) If $\tau = 1$, then $E[m_n^2(j)] = n^{2H-2}$.

(ii) If $\tau > 1$, then $E[m_n^2(j)] \simeq n^{2H-2}\tau^2$ as $n \to \infty$.

(iii) $S_N^2(n) \to E[m_n^2(j)]$ a.s. and in quadratic mean as $N \to \infty$.

Let $n_s, s = 1, \ldots, \ell$, be different values for n. Let $\hat{H}_N$ be the estimator of H obtained by linearly regressing $\log S_N^2(n_s)$ on $\log n_s, s = 1, \ldots, \ell$. Poggi and Viano (1998) proved that $\hat{H}_N$ is a strongly consistent estimator of H and

$$N^{1/2}(\hat{H}_N - H) \xrightarrow{\mathcal{L}} N(0, \sigma^2) \text{ as } N \to \infty$$

for some σ^2, depending on $n_s, s = 1, \ldots, \ell$, and constants τ and H, which can be explicitly computed.

Kurchenko's method using a Baxter-type theorem

Recall the following notation introduced in Chapter 1 for stating a Baxter-type theorem for fBm due to Kurchenko (2003).

Let $f : (a, b) \to R$ be a function and let k be a positive integer. Let $\Delta_h^{(k)} f(t)$ denote the increment of kth order of the function f in an interval $[t, t+h] \subset (a, b)$ as given, namely,

$$\Delta_h^{(k)} f(t) = \sum_{i=0}^{k} (-1)^i k_{C_i} f\left(t + \frac{i}{k} h\right).$$

Here k_{C_i} denotes the binomial coefficient. For any $m \geq 0$, positive integer $k \geq 1$ and $0 < H < 1$, define

$$V_k(m, H) = \frac{1}{2} \sum_{i,j=0}^{k} (-1)^{i+j+1} k_{C_i} k_{C_j} \left| m + \frac{i-j}{k} \right|^{2H}.$$

It can be checked that $V_1(0, H) = 1$ and $V_2(0, H) = 2^{2-2H} - 1$. Note that

$$\Delta_1^{(2)} f(t) = f(t) - 2f\left(t + \frac{1}{2}\right) + f(t+1).$$

Let $\{a_n, n \geq 1\}$ be a sequence of positive integers and $\{h_n\}$ be a sequence of positive real numbers converging to $0 \leq c \leq \infty$. Kurchenko (2003) stated the following theorem. For the proof, see Kurchenko (2003).

Theorem 9.1: Let Y be fBm with Hurst index $H \in (0, 1)$. Let $k \geq 1$. Suppose that one of the following conditions holds:

(i) $\log n = o(a_n)$ if $k = 1$ and $H \in (0, \frac{1}{2})$;

(ii) $\log n = o(a_n^{2-2H})$ if $k = 1$ and $H \in (\frac{1}{2}, 1)$;

(iii) $\log n = o(a_n)$ if $k \geq 2$.

Then

$$\frac{1}{a_n} \sum_{m=0}^{a_n - 1} \left[\frac{\Delta_{h_n}^{(k)} Y(mh_n)}{h_n^H \sqrt{V_k(0, H)}} \right]^2 \xrightarrow{\text{a.s.}} 1 \text{ as } n \to \infty. \tag{9.15}$$

Remarks: Observe that conditions (i)–(iii) of Theorem 9.1 are satisfied if $a_n = n^\alpha$ where α is a positive integer or if $a_n = \alpha^n$ where α is an integer grater than 1. In particular, for all $0 < H < 1$, and $k \geq 1$,

$$\frac{1}{n} \sum_{m=0}^{n-1} \left[\frac{\Delta_{h_n}^{(k)} Y(mh_n)}{h_n^H \sqrt{V_k(0, H)}} \right]^2 \xrightarrow{\text{a.s.}} 1 \text{ as } n \to \infty. \tag{9.16}$$

Suppose fBm Y is observed at the points

$$\left\{mh_n + \frac{ih_n}{k} : i = 0, \ldots, k-1; m = 0, \ldots, a_n - 1; n \geq 1\right\}$$

for $k = 2$, $a_n = n$ and $h_n = c \in (0, \frac{5}{2})$. From the remarks made above, it follows that

$$\frac{1}{n}\sum_{m=0}^{n-1}\left[\Delta_c^{(2)} Y(mc)\right]^2 \overset{\text{a.s.}}{\rightarrow} c^{2H} V_2(0, H) \text{ as } n \rightarrow \infty. \tag{9.17}$$

Let $\theta(H) \equiv c^{2H} V_2(0, H) = c^{2H}(2^{2-2H} - 1)$. The function $\theta(H)$ is continuous and strictly decreasing in the interval $(0, 1)$ for $c \in (0, \frac{5}{2})$. Furthermore, $\theta(0+) = 3$ and $\theta(1-) = 0$. Hence the function $\theta(H)$ has an inverse and the inverse function $H(\theta)$, defined for $\theta \in (0,3)$, is continuous and strictly decreasing in the interval $(0, 3)$. Hence the problem of estimation of the parameter θ is equivalent to that of estimation of the Hurst index H.

Let

$$\theta_n = \frac{1}{n}\sum_{m=0}^{n-1}[\Delta_c^{(2)} Y(mc)]^2 \tag{9.18}$$

and

$$\begin{aligned}\hat{\theta}_n &= \theta_n && \text{if } \theta_n < 3\\ &= 3\theta_n^n/(1 + \theta_n^n) && \text{if } \theta_n \geq 3.\end{aligned} \tag{9.19}$$

The following theorem is due to Kurchenko (2003). For the proof, see Kurchenko (2003).

Theorem 9.2: For any $c \in (0, \frac{5}{2})$, the estimator $\hat{\theta}_n$ is strongly consistent for θ and $\sqrt{n}(\hat{\theta}_n - \theta)$ is asymptotically normal with mean zero and variance

$$\sigma^2 = \left[2V_2^2(0, H) + 4\sum_{m=1}^{\infty} V_2^2(m, H)\right] c^{4H} \tag{9.20}$$

as $n \rightarrow \infty$ where $H = H(\theta)$.

9.3 Estimation of scaling exponent function $H(.)$ for locally self-similar processes

Let $\{Y(t), -\infty \leq t < \infty\}$ be a stochastic process with $E[Y(t)] = 0$ for every $t \geq 0$ and with covariance

$$R_t(u_1, u_2) = E[Y(t + u_1)Y(t + u_2)].$$

This process is said to be *locally self-similar* if

$$R_t(u_1, u_2) = R_t(0, 0) - C(t)|u_1 - u_2|^{2H(t)}(1 + O(1)) \tag{9.21}$$

as $|u_1| + |u_2| \to 0$, for every $t \geq 0$, where $C(t) > 0$. The function $H(t)$ is called the *local scaling exponent function*.

An example of such a process is

$$\begin{aligned} Y(t) = & \int_{-\infty}^{0} [(t-u)^{H(t)-\frac{1}{2}} - (-u)^{H(t)-\frac{1}{2}}]dW(u) \\ & + \int_{0}^{t} (t-u)^{H(t)-\frac{1}{2}} dW(u),\ t \geq 0 \end{aligned} \tag{9.22}$$

where $\{W(u), -\infty < u < \infty\}$ is standard Brownian motion and the integrals in (9.22) are Wiener integrals. The function $H(t)$ in (9.22) takes values in the interval $(0, 1)$ and is called the *scaling exponent function*. This process is a generalization of fBm and allows the self-similarity index to vary over time. It is called *generalized fractional Brownian motion* (gfBm). For a smooth function $H(t)$, the covariance function of gfBm satisfies (9.21) and hence the process Y is locally self-similar (Wang (1999)). Let ψ denote the Daubechies mother wavelet (Daubechies (1992)) and let $\hat{Y}_a(t)$ denote the wavelet transform of the locally self-similar process Y corresponding to the scale a and location t. Then

$$\begin{aligned} \hat{Y}_a(t) &= a^{-1/2} \int_{-\infty}^{\infty} \psi\left(\frac{u-t}{a}\right) Y(u)du \\ &= a^{1/2} \int_{-\infty}^{\infty} \psi(x)Y(t+ax)dx. \end{aligned} \tag{9.23}$$

Then

$$\begin{aligned} E|\hat{Y}_a(t)|^2 &= \frac{1}{a} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi\left(\frac{u-t}{a}\right)\psi\left(\frac{v-t}{a}\right) E[Y(u)Y(v)]dudv \\ &= a \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi(x)\psi(y)E[Y(a+tx)Y(a+ty)]dxdy \\ &= a \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi(x)\psi(y)R_t(ax, ay)dxdy \\ &\simeq a \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi(x)\psi(y)[R_t(0,0) - C(t)|ax-ay|^{2H(t)}]dxdy \\ &= C_1 a^{1+2H(t)} \end{aligned} \tag{9.24}$$

where

$$C_1 = -C(t)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x-y|^{2H(t)}\psi(x)\psi(y)dxdy. \tag{9.25}$$

Let

$$y_t(a) = \log(|\hat{Y}_a(t)|^2), \tag{9.26}$$

$$C_2 = E[\log(|\hat{Y}_a(t)|^2 / E|\hat{Y}_a(t)|^2)], \tag{9.27}$$

and

$$\epsilon_t(a) = \log(|\hat{Y}_a(t)|^2 / E|\hat{Y}_a(t)|^2) - C_2. \tag{9.28}$$

Then

$$y_t(a) = C_2 + \log(E|\hat{Y}_a(t)|^2) + \epsilon_t(a). \tag{9.29}$$

Equation (9.24) derived above implies the regression model

$$y_t(a) \simeq c + (2\ H(t) + 1) \log a + \epsilon_t(a) \tag{9.30}$$

for small-scale a where $c = \log C_1 + C_2$. This model suggests that the function $H(t)$ can be estimated by the method of least squares using the above regression model. Suppose we select a sequence of scales $a_1 > \cdots > a_k$, say, $a_j = 2^{-j}$, $j = 1, \ldots, k$. Let $x_j = \log a_j$ and $y_j = y_t(a_j)$. Treating (x_j, y_j), $j = 1, \ldots, k$, as a set of bivariate data, the least squares estimator of the function $H(t)$ can be obtained, using the regression model (9.30), as

$$\hat{H}_k(t) = \frac{1}{2}\left[\frac{\sum_{j=1}^{k}(x_j - \bar{x})(y_j - \bar{y})}{\sum_{j=1}^{k}(x_j - \bar{x})^2} - 1\right] \tag{9.31}$$

where $\bar{x} = k^{-1}\sum_{j=1}^{k} x_j$ and $\bar{y} = k^{-1}\sum_{j=1}^{k} y_j$. Wang *et al*. (2001) stated the following theorem.

Theorem 9.3: Suppose Y is a Gaussian process such that the covariance function satisfies the condition (9.21). Then

$$\hat{H}_k(t) \xrightarrow{p} H(t) \quad \text{as} \quad k \to \infty.$$

We omit the proof of this result. For a proof of this theorem and related results, see Wang *et al*. (2001).

10

Filtering and prediction for linear systems driven by fBm

10.1 Introduction

We now study extensions of classical filtering theory to linear systems driven by fBms through the techniques developed by Kleptsyna, Le Breton and others.

10.2 Prediction of fBm

Let $W^H = \{W_t^H, -\infty < t < \infty\}$ be standard fBm with Hurst index $H \in (\frac{1}{2}, 1)$. We now study the problem of prediction of W_s^h for some $s > 0$ on the basis of observations $\{W_t^H, -T < t < 0\}$ for $T > 0$. By the stationarity of the increments of fBm W^H, this problem is equivalent to the problem of predicting the difference $W_{t+u}^H - W_t^H$ on the basis of the observations $\{W_u^H - W_t^H, t-T < u < t\}$ for any t. For $a > 0$ and $T > 0$, let

$$\hat{W}_{a,T}^H = E[W_a^H | W_s^H, s \in (-T, 0)]. \tag{10.1}$$

Since the process W^H is Gaussian, the random variable $\hat{W}_{a,T}^H$ is a linear function of the family $\{W_s^H, s \in (-T, 0)\}$ and it is the optimal predictor under the squared error loss function. Suppose that

$$\hat{W}_{a,T}^H = \int_{-T}^{0} g_T(a, t) dW_t^H \tag{10.2}$$

Statistical Inference for Fractional Diffusion Processes B. L. S. Prakasa Rao

for some real-valued function $g_T(a,t)$. From the fact that the random variable $\hat{W}^H_{a,T}$ has to be an optimal predictor, given $\{W^H_s, s \in (-T, 0)\}$, it follows that

$$E\left[\left(W^H_a - \int_{-T}^{0} g_T(a,t)dW^H_t\right)(W^H_u - W^H_v)\right] = 0, \ -T < v < u \leq 0 \quad (10.3)$$

from the underlying Hilbertian structure of the class of functions g for which the stochastic integral in (10.2) is defined. Using the fact that

$$E\left[\left(\int_{-\infty}^{\infty} f(s)dW^H_s\right)\left(\int_{-\infty}^{\infty} g(s)dW^H_s\right)\right]$$
$$= H(2H-1)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(s)g(t)|s-t|^{2H-2}dtds,$$

Equation (10.3) reduces to

$$(2H-1)\int_0^T g_T(a,-t)|t-s|^{2H-2}dt = (a+s)^{2H-1} - s^{2H-1}, \ 0 \leq s \leq T. \quad (10.4)$$

Furthermore, the function $g_T(a,t)$ has the scaling property

$$g_T(a,t) = g_{T/a}\left(1, \frac{t}{a}\right)$$

from the self-similarity of fBm W^H. Gripenberg and Norros (1996) proved that Equation (10.4) holds if

$$g_T(a,-t) = \frac{\sin(\pi(H-\frac{1}{2}))}{\pi} t^{-H+\frac{1}{2}}(T-t)^{-H+\frac{1}{2}} \int_0^a \frac{\sigma^{H-\frac{1}{2}}(\sigma+T)^{H-\frac{1}{2}}}{\sigma+t}\, d\sigma \quad (10.5)$$

for $T < \infty, t > 0$, and

$$g_\infty(a,-t) = \frac{\sin(\pi(H-\frac{1}{2}))}{\pi} t^{-H+\frac{1}{2}} \int_0^a \frac{\sigma^{H-\frac{1}{2}}}{\sigma+t}\, d\sigma. \quad (10.6)$$

for $t > 0$. For details, see Gripenberg and Norros (1996).

10.3 Filtering in a simple linear system driven by fBm

Suppose the signal is a fixed random variable ξ with Gaussian distribution, with mean zero and variance σ^2, independent of fBm W^H. Further suppose that the observation process Y is defined by the SDE

$$dY(s) = \xi\, a(s)ds + b(s)dW^H_s, \ 0 \leq s \leq t. \quad (10.7)$$

We assume that the functions $a(.)$ and $b(.)$ are known nonrandom real-valued continuous functions with $b(.)$ nonvanishing. The problem is to compute the best estimator of ξ based on the observations $\{Y(s), 0 \leq s \leq t\}$. Since the system (ξ, Y) is Gaussian, it is known that the conditional expectation

$$\hat{\xi} = E[\xi | Y(s), 0 \leq s \leq t] \tag{10.8}$$

is the best estimator under the squared error loss function and it a linear function of $\{Y(s), 0 \leq s \leq t\}$. Consider estimators of the form

$$\hat{\xi}_t(f) = \int_0^t f(s) dY_s. \tag{10.9}$$

The stochastic integral in (10.9) is well defined for functions $f \in L^2((0,t); R)$. We have to find a function $g \in L^2((0,t); R)$ such that

$$E([\hat{\xi}_t(g) - \xi]^2) \leq E([\hat{\xi}_t(f) - \xi]^2)$$

for every $f \in L^2((0,t); R)$. Due to the Hilbertian structure of the space $L^2((0,t); R)$, it follows that the optimal function $g(.)$ satisfies the condition

$$E([\hat{\xi}_t(g) - \xi][Y_u - Y_v]) = 0,\ 0 \leq v < u < t. \tag{10.10}$$

Note that

$$\begin{aligned}
\hat{\xi}_t(f) - \xi &= \int_0^t f(s) dY(s) - \xi \\
&= \int_0^t f(s)[a(s)\xi ds + b(s) dW^H(s)] - \xi \\
&= \left[\int_0^t a(s) f(s)\ \ ds - 1\right]\xi + \int_0^t b(s) f(s) dW^H(s) \qquad (10.11)
\end{aligned}$$

and

$$Y_u - Y_v = \left[\int_v^u a(s) ds\right]\xi + \int_v^u b(s) dW^H(s),\ 0 \leq v \leq u < t. \tag{10.12}$$

Combining Equations (10.10)–(10.12) and using the fact that the random variable ξ is independent of the process W^H, it follows that

$$\begin{aligned}
&E([\hat{\xi}_t(g) - \xi][Y_u - Y_v]) \\
&\quad = \left[\int_0^t a(s) g(s) ds - 1\right]\left[\int_v^u a(s) ds\right] Var(\xi) \\
&\qquad + Cov\left[\int_v^u b(s) dW^H(s), \int_0^t b(s) g(s) dW^H(s)\right]
\end{aligned}$$

$$= \sigma^2 \left[\int_0^t a(s)g(s)ds - 1 \right] \left[\int_v^u a(s)ds \right]$$

$$+ H(2H-1) \int_v^u b(r) \left[\int_0^t b(s)g(s)|s-r|^{2H-2} ds \right] dr \quad (10.13)$$

for $0 \leq v \leq u \leq t$. Suppose we choose $g(.) \in L^2((0,t); R)$ such that

$$\int_0^t b(s)g(s)|s-r|^{2H-2} ds = \frac{\sigma^2}{H(2H-1)} \left[1 - \int_0^t a(s)g(s)ds \right] \frac{a(r)}{b(r)}, \; 0 \leq r \leq t.$$

Then the estimator $\xi_{\hat{t}}(g)$ will satisfy Equation (10.10) and hence the optimal filter. Le Breton (1998) has devised a method to solve the above equation. For details, see Le Breton (1998). Le Breton (1998) has also studied the problem when the observation process Y satisfies a SDE of the type

$$dY(t) = \xi \; a(t)dt + \lambda \; a(t)dW^H(t), \; t \geq 0.$$

10.4 General approach for filtering for linear systems driven by fBms

We now study a general approach developed by Kleptsyna *et al*. (2000b) for filtering for linear systems driven by fBms.

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ be a stochastic basis satisfying the usual conditions and the processes under discussion are $\{\mathcal{F}_t\}$-adapted. We further assume the natural filtration of a process as the P-completion of the filtration generated by the process. For any process $\{\mathcal{F}_t\}$-adapted Z, we denote its natural filtration by $\{\mathcal{Z}_t\}$.

Suppose the processes $X = \{X(t), 0 \leq t \leq T\}$ and $Y = \{Y(t), 0 \leq t \leq T\}$ represent the signal and the observation respectively and they are governed by the linear system of SDEs

$$dX(t) = a(t)X(t)dt + b(t)dW_t^h, \; X_0 = \eta, \; 0 \leq t \leq T$$

$$dY(t) = A(t)X(t)dt + B(t)dW_t^H, \; X_0 = \eta, \; 0 \leq t \leq T \quad (10.14)$$

where W^h and W^H are *independent* standard fBms with Hurst indices h and H respectively with $h \in (\frac{1}{2}, 1)$ and $H \in (\frac{1}{2}, 1)$. Further assume that the coefficients $a(.), b(.), A(.)$ and $B(.)$ are nonrandom (deterministic), bounded and smooth functions with $B(.)$ nonvanishing such that $1/B(.)$ is bounded. In addition, assume that the random variable η is a Gaussian random variable independent of the process (W^h, W^H) with mean m_0 and variance γ_0. Then the system (10.14) has a uniquely defined solution (X, Y) which is Gaussian.

For any process $Z = \{Z_t, 0 \leq t \leq T\}$, such that $E|Z_t| < \infty, 0 \leq t \leq T$, define

$$\pi_t(Z) = E(Z_t|Y_s, 0 \leq s \leq t) \equiv E(Z_t|\mathcal{Y}_t) \tag{10.15}$$

where $\mathcal{Y}_t$ denotes the σ-algebra generated by the process $\{Y_s, 0 \leq s \leq t\}$.

Suppose that the process $\{Y_s, 0 \leq s \leq t\}$ is observed and we wish to estimate X_t. From the discussion in the previous section, it follows that the solution to this problem is the conditional expectation of X_t given the process $\{Y_s, 0 \leq s \leq t\}$ or equivalently the conditional expectation of X_t given $\mathcal{Y}_t$. This is called the *optimal filter* for the problem. Since the system (X, Y) is Gaussian, the optimal filter is a linear function of $\{Y_s, 0 \leq s \leq t\}$ and it is Gaussian, hence it is characterized by its mean $\pi_t(X) = E(X_t|\mathcal{Y}_t)$ and its variance $\gamma_{XX}(t) = E[(X_t - \pi_t(X))^2|\mathcal{Y}_t]$.

Let $C(t) = A(t)X(t)$. Following the notation introduced in Chapter 1, define

$$Q(t) = \frac{d}{dw_t^H} \int_0^t k_H(t,s) \frac{C(s)}{B(s)} ds, \; 0 \leq t \leq T. \tag{10.16}$$

Suppose the sample paths of the process Q belong P-a.s. to $L^2([0,T], dw_s^H)$. Then

$$\int_0^t C(s)ds = \int_0^t K_H^B(t,s)Q(s)dw_s^H \tag{10.17}$$

where

$$K_H^B(t,s) = H(2H-1)\int_s^t B(r)r^{H-\frac{1}{2}}(r-s)^{H-\frac{1}{2}}dr. \tag{10.18}$$

It is interesting to note that

$$\int_0^t B(s)dW_s^H = \int_0^t K_H^B(t,s)dM_s^H \tag{10.19}$$

P-a.s. where M^H is the fundamental Gaussian martingale associated with the process W^H.

The following Lemma is proved in Kleptsyna *et al.* (2000a).

Lemma 10.1: Let X be the process as in (10.14) and Q be the process as defined by (10.16). Let M^h be the fundamental Gaussian martingale associated with fBm W^h. Then the following representations hold:

$$X_t = \eta + \int_0^t a(s)X_s ds + \int_0^t K_h^b(t,s)dM_s^h \tag{10.20}$$

and

$$Q_t = p(t, 0)\eta + \int_0^t a(s)p(t, s)X_s ds + \int_0^t q(t, s)dM_s^h \tag{10.21}$$

where

$$p(t, s) = \frac{d}{dw_t^H} \int_s^t k_H(t, r)\frac{A(r)}{B(r)}dr \tag{10.22}$$

and

$$q(t, s) = \frac{d}{dw_t^H} \int_s^t k_H(t, r)K_h^b(r, s)\frac{A(r)}{B(r)}dr. \tag{10.23}$$

Let

$$Y_t^* = \int_0^t \frac{k_H(t, s)}{B(s)}dY_s,\ 0 \le t \le T. \tag{10.24}$$

Then the process Y^* is a semimartingale with the decomposition

$$Y_t^* = \int_0^t Q(s)dw_s^H + M_t^H,\ 0 \le t \le T \tag{10.25}$$

and the natural filtrations of Y and Y^* coincide. Let $\nu = \{\nu_t, 0 \le t \le T\}$ be defined by

$$\nu_t = Y_t^* - \int_0^t \pi_t(Q)dw_s^H,\ 0 \le t \le T. \tag{10.26}$$

The process ν is a continuous Gaussian $\{\mathcal{Y}_t\}$-martingale with the quadratic variation w^H. Furthermore, if $R = \{R_t, 0 \le t \le T\}$ is a square integrable $\{\mathcal{Y}_t\}$-martingale with $R_0 = 0$, then there exists a $\{\mathcal{Y}_t\}$-adapted process $\Phi = \{\Phi_t, 0 \le t \le T\}$ such that

$$E\left[\int_0^T \Phi_t^2 dw_t^H\right] < \infty$$

and P-a.s.,

$$R_t = \int_0^t \Phi_s d\mu_s,\ 0 \le t \le T. \tag{10.27}$$

This gives a representation of square integrable $\{\mathcal{Y}_t\}$-martingales.

For a detailed proof, see Kleptsyna *et al*. (2000b). The following theorem is due to Kleptsyna *et al*. (2000b).

Theorem 10.2: In addition to the assumptions stated above, let $\zeta = \{\zeta_t, 0 \le t \le T\}$ be a semimartingale with the decomposition

$$\zeta_t = \zeta_0 + \int_0^t \beta_s ds + m_t, \; 0 \le t \le T \tag{10.28}$$

where $E[\zeta_0]^2 < \infty$, $E[\int_0^T \beta_s^2 ds] < \infty$ and $m = \{m_t, 0 \le t \le T\}$ is a square integrable $\{\mathcal{F}_t\}$-martingale with quadratic covariation

$$\langle m, M^H \rangle_t = \int_0^t \lambda_s dw_s^H.$$

Then, P-a.s., for $0 \le t \le T$, the process $\pi_t(\zeta) = E(\zeta|\mathcal{Y}_t)$ satisfies the equation

$$\pi_t(\zeta) = \pi_0(\zeta) + \int_0^t \pi_s(\beta)ds + \int_0^t [\pi_s(\lambda) + \pi_s(\zeta Q) - \pi_s(\zeta)\pi_s(Q)]d\nu_s. \tag{10.29}$$

For any $t \in [0, T]$, define the Gaussian semimartingale $\xi^t = (X^t, Q^t)' = ((X_s^t, Q_s^t)', s \in [0, t])$ given by

$$X_s^t = \eta + \int_0^s a(u)X_u du + \int_0^s K_h^b(t, u)dM_u^h \tag{10.30}$$

and

$$Q_s^t = p(t, 0)\eta + \int_0^s a(u)p(t, u)X_u du + \int_0^s q(t, u)dM_u^h \tag{10.31}$$

from the representation given in Lemma 10.1. Here α' denotes the transpose of the vector α. Kleptsyna *et al*. (2000b) obtained the solution of the filtering problem for the linear system given in (10.14) in terms of Kalman-type filtering equations for the bivariate process ξ. Let

$$\pi_t(\xi)' = (\pi_t(X), \pi_t(Q))$$

and

$$\gamma_{\xi\xi}(t) = E[(\xi_t - \pi_t(\xi))(\xi_t - \pi_t(\xi))'].$$

Note that the mean $\pi_t(X)$ and the variance $\gamma_{XX}(t)$ of the optimum Gaussian filter are the first component of the vector $\pi_t(\xi)$ and the (1,1)th entry of the covariance matrix $\gamma_{\xi\xi}(t)$ respectively. For more details and proofs, see Kleptsyna *et al*. (2000b).

Special case

Suppose the linear system, as specified in (10.14), is given by

$$dX(t) = \theta X(t)dt + dV_t^H, \; X_0 = 0, \; 0 \le t \le T$$
$$dY(t) = \mu X(t)dt + dW_t^H, \; 0 \le t \le T \tag{10.32}$$

where V^H and W^H are *independent* standard fBms with the same Hurst index $H \in (\frac{1}{2}, 1)$ and the parameters θ and $\mu \neq 0$ are fixed constants.

Let

$$Z_t = \int_0^t k_H(t,s)dX_s,\ 0 \le t \le T$$

and

$$Z_t^0 = \int_0^t k_H(t,s)dY_s,\ 0 \le t \le T$$

where $k_H(t,s)$ is as defined in Equation (1.70) of Chapter 1. Let

$$Q(t) = \frac{d}{dw_t^H}\int_0^t k_H(t,s)X(s)ds,\ 0 \le t \le T.$$

Then

$$Z_t = \theta \int_0^t Q(s)dw_s^h + M_t^H,\ 0 \le t \le T \tag{10.33}$$

and

$$Z_t^0 = \theta \int_0^t Q(s)dw_s^h + N_t^H,\ 0 \le t \le T \tag{10.34}$$

where M^H and N^H are independent fundamental Gaussian martingales associated with fBms V^H and W^H respectively. Furthermore, the natural filtrations of the processes X and Z are the same as well as those of Y and Z^0. Also, the following representations hold P-a.s.:

$$X_t = \int_0^t K_H(t,s)dZ_s,\quad Y_t = \int_0^t K_h(t,s)dZ_s^0$$

where

$$K_H(t,s) = H(2H-1)\int_s^t r^{H-\frac{1}{2}}(r-s)^{H-\frac{1}{2}}dr,\ 0 \le s \le t.$$

In particular, it follows that

$$X_t = \theta \int_0^t K_H(t,s)Q(s)dw_s^H + \int_0^t K_H(t,s)dM_s^H. \tag{10.35}$$

Kleptsyna and Le Breton (2002a) have shown that

$$Q(t) = \frac{\lambda_H}{2(2-2H)}\left[t^{2H-1}Z_t + \int_0^t r^{2H-1}dZ_r\right] \tag{10.36}$$

where λ_H is given in Equation (1.71) of Chapter 1. We observe that filtering the signal X from the observations Y is equivalent to filtering the signal X from the observations Z^0. Kleptsyna and Le Breton (2002b) obtained the optimal filter in an explicit form in this special case. See Proposition 3.1 in Kleptsyna and Le Breton (2002b). We do not go into the details here.

Remarks:

(i) Ahmad and Charalambous (2002) studied continuous time filtering for linear multidimensional systems driven by fBm. They derive the optimum linear filter equations which involve a pair of functional–differential equations giving the optimum covariance (matrix-valued) functions and the optimum filter. These equations are the analogues of matrix–Ricatti differential equations in classical Kalman filtering.

(ii) We now discuss an example due to Kleptsyna *et al*. (2000a) illustrating the nonlinear optimal filtering problem. Suppose that the signal is a fixed $\mathcal{F}_0$-measurable random variable η with a probability measure π_0 and the observation process Y is governed by the SDE

$$dY(t) = a(t,\eta)dt + B(t)dW_t^H,\ 0 \leq t \leq T \tag{10.37}$$

where $a(.,.)$ and $B(.)$ are deterministic functions with $B(.)$ nonvanishing. Suppose further that the random variable η and fBm W^H are *independent*. Following the arguments given earlier, it can be shown that solution of the filtering problem up to time t or the optimal filter up to time t in this case is the conditional distribution π_t of η given the σ-algebra generated by the process $\{Y_s, 0 \leq s \leq t\}$. Let

$$Q(t,x) = \frac{d}{dw_t^H}\int_0^t \frac{k_H(t,s)}{B(s)}a(s,x)ds,\ 0 \leq t \leq T,\ x \in R. \tag{10.38}$$

Suppose the process $Q(.,x)$ is well defined and, for all $x \in R$, $Q(.,x) \in L^2([0,T], dw_t^H)$. Let

$$\pi_t(\phi) = E[\phi(\eta)|Y_s, 0 \leq s \leq t]$$

for any bounded continuous real-valued function $\phi(.)$. Applying the Girsanov-type result given in Theorem 1.20 in Chapter 1, we get that

$$\pi_t(\phi) = \frac{\sigma_t(\phi)}{\sigma_t(1)}$$

where

$$\sigma_t(\phi) = \tilde{E}[\phi(\eta)(\Lambda_t)^{-1}|Y_s, 0 \leq s \leq t]$$

and

$$\Lambda_t = \exp\left[-\int_0^t Q(s,\eta)dM_s^H - \frac{1}{2}\int_0^t Q^2(s,\eta)dw_s^H\right].$$

Here $\tilde{E}(.)$ denotes the expectation with respect to the measure $\tilde{P} = \Lambda_T P$. Kleptsyna *et al*. (2000a) derive the equations for the optimal filter.

(iii) Nonlinear filtering problems with fBm as observational noise has been studied by Coutin and Decreusefond (1999), Gawarecki and Mandrekar (2001) and Amirdjanova (2002). Xiong and Zhao (2005) consider the nonlinear filtering problem when the signal process is a Markov diffusion process and the observation process is corrupted by fBm noise. We do not go in to details here.

(iv) Bishwal (2003) considered a partially observed stochastic differential system of the type

$$dY_t = \theta\ f(t, X_t)dt + g(t)dW_t^H,\ Y_0 = \xi,\ 0 \le t \le T$$
$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t,\ X_0 = \eta,\ 0 \le t \le T$$

where $H \in (\frac{1}{2}, 1)$, W^H is fBm with Hurst index H independent of Brownian motion W and the random vector (η, ξ) is independent of the process (W, W^H). The problem is to estimate θ on the basis of observations $\{Y_s, 0 \le s \le T\}$. Under some regularity conditions on the functions $a(.,.)$, $b(.,.)$ and $f(.,.)$, Bishwal (2003) obtained asymptotic properties of the MLE of θ. We omit the details.

References

Abramowitz, M. A. and Stegun, I. A. (1970) *Handbook of Mathematical Functions*, Dover, New York.

Abry, P. and Sellan, F. (1996) The wavelet-based synthesis for fractional Brownian motion proposed by F. Sellan and Y. Meyer: remarks and fast implementation, *Appl. Comput. Harmonic Anal.*, **3**, 377–383.

Acharya, S. K. and Mishra, M. N. (1994) Upper and lower class of functions for the maximum likelihood estimator for nonlinear stationary homogeneous diffusion processes, *Stat. Probab. Lett.*, **12**, 353–366.

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

Alos, E., Mazet, O. and Nualart, D. (2001) Stochastic calculus with respect to Gaussian processes, *Ann. Probab.*, **29**, 766–801.

Amirdjanova, A. (2002) Nonlinear filtering with fractional Brownian motion, *Appl. Math. Optim.*, **46**, 81–88.

Amirdjanova, A. and Linn, M. (2007) Stochastic evolution equations for nonlinear filtering of random fields in the presence of fractional Brownian sheet observation noise, arXiv: 0707.3856v1 [math.PR] 26 July 2007.

Androshchuk, T. (2005) The local asymptotic normality of a family of measures generated by solutions of stochastic differential equations with a small fractional Brownian motion, *Theor. Probab. Math. Stat.*, **71**, 1–15.

Azais, J. M. (1990) Condition for convergence of number of crossings to the local time: application to stable processes with independent increments and to Gaussian processes, *Probab. Math. Stat.*, **11**, 19–36.

Bahadur, R. R. (1960) On the asymptotic efficiency of tests and estimators, *Sankhya*, **22**, 229–252.

Bahadur, R. R. and Ranga Rao, R. (1960) On deviations of the sample mean, *Ann. Math. Stat.*, **31**, 1015–1027.

Baran, S. and Pap, G. (2003) Estimation of the mean of stationary and nonstationary Ornstein-Uhlenbeck processes and sheets, *Comput. Math. Appl.*, **45**, 563–579.

Basawa, I. V. and Prakasa Rao, B. L. S. (1980) *Statistical Inference for Stochastic Processes*, Academic Press, London.

Bassingthwaighte, J. B. and Raymond, G. M. (1994) Evaluating re-scaled range analysis for time series, *Ann. Biomed. Eng.*, **22**, 432–444.

Statistical Inference for Fractional Diffusion Processes B. L. S. Prakasa Rao

Bassingthwaighte, J. B. and Raymond, G. M. (1995) Evaluation of dispersion analysis method for fractal time series, *Ann. Biomed. Eng*., **23**, 491–505.

Bender, C., Sottinen, T. and Valkeila, E. (2006) Arbitrage with fractional Brownian motion?, *Theory Stochastic Process*, **12** (28), 1–12.

Beran, J. (1994) *Statistics for Long Memory Processes*, Chapman and Hall, London.

Beran, J. and Terrin, N. (1994) Estimation of the long-memory parameter based on multivariate central limit theorem, *J. Time Ser. Anal*., **15**, 269–278.

Bercu, B., Coutin, L. and Savy, N. (2008) Sharp large deviations for the fractional Ornstein-Uhlembeck process, arXiv: 0810.449lvl[math.PR] 24 October 2008.

Bertin, K., Torres, S. and Tudor, C. A. (2007) Maximum likelihood estimators and random walks in long memory models, arXiv:0711.0513v1 [mathST] 4 November 2007.

Berzin, C. and Leon, J. R. (2006) Estimation in models driven by fractional Brownian motion, Preprint.

Berzin, C. and Leon, J. R. (2007) Estimating the Hurst parameter, *Stat. Inference Stochastic Process*, **10**, 49–73.

Billingsley, P. (1968) *Convergence of Probability Measures*, John Wiley & Sons, Inc., New York.

Bishwal, J. P. N. (2003) Maximum likelihood estimation in partially observed stochastic differential system driven by a fractional Brownian motion, *Stochastic Anal. Appl*., **21**, 995–1007.

Bishwal, J. P. N. (2008) Large deviations in testing fractional Ornstein-Uhlenbeck models, *Stat. Probab. Lett*., **78**, 953–962.

Bjork, T. and Hult, H. (2005) A note on Wick products and the fractional Black-Scholes model, *Financ. Stochastics*, **9**, 197–209.

Borwanker, J. D., Kallianpur, G. and Prakasa Rao, B. L. S. (1971) The Bernstein-von Mises theorem for Markov processes, *Ann. Math. Stat*., **42**, 1241–1253.

Bose, A. (1986) Berry-Esseen bound for the maximum likelihood estimator in the Ornstein-Uhlenbeck process, *Sankhya Ser. A*, **48**, 181–187.

Buldyrev, S. V., Goldberger, A. L., Havlin, S., Peng, C.-K., Stanley, H. E., Stanley, M. H. R. and Simons, M. (1993) Fractal landscapes and molecular evolution: modelling the Moyosin heavy chain gene family. *Biophys. J*., **65**, 2673–2679.

Cairoli, R. and Walsh, J. B. (1975) Stochastic integrals in the plane, *Acta Math*., **134**, 111–183.

Cannon, M. J., Percival, D. M., Raymond, G. M. and Bassingthwaighte, J. B. (1997) Evaluating windowed variance methods for estimating the Hurst coefficient of time series, *Physica A*, **241**, 606–626.

Chen, G., Hall, P. and Poskitt, D. S. (1995) Periodogram-based estimators of fractal properties. *Ann. Stat*., **23**, 1684–1711.

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

Cheridito, P., Kawaguchi, H. and Maejima, M. (2003) Fractional Ornstein-Uhlenbeck processes, *Electron. J. Probab*., **8**, 1–14.

Cialenco, I., Lototsky, S. G. and Pospisil, J. (2008) Asymptotic properties of the maximum likelihood estimator for stochastic parabolic equations with additive fractional Brownian motion, arXiv:0804.0407v1 [math.PR] 2 April 2008.

Coeurjolly, J.-F. (2001) Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths, *Stat. Inference Stochastic Process*, **4**, 199–207.

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

Constantine, A. G. and Hall, P. (1994) Characterizing surface smoothness via estimation of effective fractal dimension, *J. R. Stat. Soc*., **56**, 97–113.

Coutin, L. and Decreusefond, L. (1999) Abstract nonlinear filtering theory in the presence of fractional Brownian motion, *Ann. Appl. Probab*., **9**, 1058–1090.

Da Prato, G. and Zabczyk, J. (1992) *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, Cambridge.

Dasgupta, A. (1998) Fractional Brownian motion: its properties and applications to stochastic integration, Ph.D. Thesis, University of North Carolina, Chapel Hill, NC.

Daubechies, I. (1992) *Ten Lectures on Wavelets*, Regional Conference Series in Applied Mathematics, SIAM, Philadelphia.

Davis, J. A. (1969) A characterization of the upper and lower classes in terms of the convergence rates. *Ann. Math. Stat*., **40**, 1120–1123.

Decreusefond, L. and Lavaud, N. (1996) Simulation of the fractional Brownian motion and application to the fluid queue, in *Proceedings of the ATNAC'96 Conference*.

Decreusefond, L. and Ustunel, A. S. (1999) Stochastic analysis of the fractional Brownian motion, *Potential Anal*., **10**, 177–214.

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

Dellacherie, C. and Meyer, P. A. (1982) *Probabilities and Potential*, North-Holland, Amsterdam.

Dembo, A. and Zeitouni, O. (1998) *Large Deviation Techniques and Applications*, Second Edition, Springer, New York.

Dieker, T. (2004) *Simulation of fractional Brownian motion*, University of Twente, Enschede.

Doob, J. L. (1942) The Brownian movement and stochastic equations, *Ann. Math*., **43**, 351–369.

Doukhan, P., Oppenheim, G. and Taqqu, M. S. (2003) *Theory of Long-Range Dependence*, Birkhauser, Boston.

Dzhaparidze, K. and van Zanten, H. (2003) Optimality of an explicit series expansion of the fractional Brownian sheet, Preprint.

Dzhaparidze, K. and van Zanten, H. (2004) A series expansion of fractional Brownian motion, *Probab. Theory Relat. Fields*, **130**, 39–55.

Elliott, R. and van der Hock, J. (2003) A general white noise theory and applications to finance, *Math. Financ*., **13**, 301–320.

El-Nouty, C. (2001) On the lower classes of fractional Brownian motion, *Stud. Sci. Math. Hung*., **37**, 363–390.

El-Nouty, C. (2002) Lower classes of fractional Brownian motion under Holder norms, in *Limit Theorems in Probability and Statistics, 1999*, ed. I. Berkes, E. Csaki and M. Csorgo, Janos Bolyai Mathematical Society, Budapest.

El-Nouty, C. (2003a) The fractional mixed fractional Brownian motion, *Stat. Probab. Lett*., **65**, 111–120.

El-Nouty, C. (2003b) A note on the fractional integrated fractional Brownian motion, *Acta Appl. Math*., **78**, 103–114.

Etemadi, N. and Kallianpur, G. (1977) Non-anticipative transformations of the two-parameter Wiener process and a Girsanov theorem, *J. Multivariate Anal*., **7**, 28–49.

Feldman, J. (1958) Equivalence and perpendicularity of Gaussian processes, *Pac. J. Math*., **8**, 699–708; Correction, **9**, 1295–1296.

Feller, W. (1968) *An Introduction to Probability Theory and its Applications*, John Wiley & Sons, Inc., New York.

Feuverger, A., Hall, P. and Wood, A. T. A. (1994) Estimation of the fractal index and fractal dimension of a Gaussian process by counting the number of level crossings, *J. Time Ser. Anal*., **15**, 587–606.

Flandrin, P. and Goncalves, P. (1994) From wavelets to time-scale energy distributions, in *Recent Advances in Wavelet Analysis*, ed. L. L. Schumaker and G. Webb, Academic Press, Boston, pp. 309–334.

Gawarecki, L. and Mandrekar, V. (2001) On the Zakai equation of filtering with Gaussian noise, in *Stochastics in Finite and Infinite Dimensions*, Birkhauser, Boston, pp. 145–151.

Geweke, J. and Porter-Hudak, S. (1983) The estimation and application of long-memory time series models, *J. Time Ser. Anal*., **4**, 221–238.

Gikhman, I. and Skorokhod, A. V. (1972) *Stochastic Differential Equations*, Springer, Berlin.

Goldberger, A. L. and West, B. (1987) Fractals in physiology and medicine, *Yale J. Med. Biol*., **60**, 421–435.

Goncalves, P. and Flandrin, P. (1993) Bilinear time-scale analysis applied to a local scaling exponents estimation, in *Progress in Wavelet Analysis and Applications*, ed. Y. Meyer and S. Roques, Frontiers, Paris, pp. 271–276.

Gripenberg, G. and Norros, I. (1996) On the prediction of fractional Brownian motion, *J. Appl. Probab*., **33**, 400–410.

Gushchin, A. A. and Kuchler, U. (1999) Asymptotic inference for a linear stochastic differential equation with time delay, *Bernoulli*, **5**, 1059–1098.

Hajek, J. (1958) On a property of normal distribution of any stochastic process, *Czech Math. J*., **8**, 610–618.

Hall, P. and Heyde, C. C. (1980) *Martingale Limit Theory and its Applications*, Academic Press, New York.

Hall, P., Koul, H. L. and Turlach, B. A. (1997) Note on convergence rates of semiparametric estimators of dependence index, *Ann. Stat*., **25**, 1725–1739.

Henry, M. and Zafforoni, P. (2003) The long-range dependence paradigm for macroeconomics and finance, in *Theory of Long-Range Dependence*, ed. P. Doukhan, G. Oppenheim and M. S. Taqqu, Birkhauser, Boston, pp. 417–438.

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

Hu, Y. (2001) Prediction and translation of fractional Brownian motion, in *Stochastics in Finite and Infinite Dimensions*, ed. T. Hida, R. L. Karandikar, H. Kunita, B. S. Rajput, S. Watanabe and J. Xiong, Birkhauser, Boston, pp. 153–171.

Hu, Y. and Nualart, D. (2009) Parameter estimation for fractional Ornstein-Uhlenbeck process, arXiv:0901.4925vl [math.PR] 30 January 2009.

Hu, Y. and Oksendal, B. (2003) Fractional white noise calculus and application to finance, *Infinite Dimensional Anal. Quantum Probab. Relat. Top*., **6**, 1–32.

Hu, Y., Xiao, W. and Zhang, W. (2009) Exact maximum likelihood estimator for drift fractional Brownian motion at discrete observation, arXiv:0904.418vl [math.ST] 27 April 2009.

Huang, S. T. and Cambanis, S. (1978) Stochastic and multiple Wiener integrals of Gaussian processes, *Ann. Probab*., **6**, 585–614.

Huebner, M. and Rozovskii, B. L. (1995) On asymptotic properties of maximum likelihood estimators for parabolic stochastic SPDE's, *Probab. Theory Relat. Fields*, **103**, 143–163.

Huebner, M., Khasminski, R. and Rozovskii, B. L. (1993) Two examples of parameter estimation for stochastic partial differential equations, in *Stochastic Processes: A Festschrift in honour of Gopinath Kallianpur*, Springer, New York, pp. 149–160.

Hurst, H. E. (1951) Long term storage capacity of reservoirs (with discussion), *Trans. Am. Soc. Civ. Eng*., **116**, 770–808.

Ibragimov, I. A. and Khasminskii, R. Z. (1981) *Statistical Estimation: Asymptotic Theory*, Springer, New York.

Ikeda, N. and Watanabe, S. (1981) *Stochastic Differential Equations and Diffusion Processes*, North-Holland, Amsterdam.

Ito, K. (1984) *Foundations of Stochastic Differential Equations in Infinite Dimensional Spaces*, Vol. **47**, CBMS Notes, SIAM, Baton Rouge, LA.

Ito, K. and McKean, H. P. (1965) *Diffusion Processes and Their Sample Paths*, Springer, Berlin.

Jain, N. C., Jogdeo, K. and Stout, W. F. (1975) Upper and lower functions for martingales and mixing processes, *Ann. Probab*., **3**, 119–145.

Jeanne, R., Ohley, W. J., Majumdar, S. and Lemineur, G. (2001) Fractal analysis of bone x-ray tomographic microscopy projections, *IEEE Trans. Med. Imaging*, **20**, 443–449.

Jensen, M. (1998) An approximate wavelet MLE of short and long memory parameters, *Stud. Nonlinear Dyn. Econ*., **3**, 239–253.

Jung Lae, C. and Chyi-Chying, L. (1997) Sea clutter rejection in radar image using wavelets and fractals, *Proceedings of the International Conference Image Processing*, Vol. **2**, 354–357.

Kallianpur, G. and Selukar, R. S. (1993) Estimation of Hilbert-space valued parameters by the method of sieves, in *Statistics and Probability*, ed. J. K. Ghosh, S. K. Mitra, K. R. Parthasarthy and B. L. S. Prakasa Rao, Wiley Eastern, New Delhi, pp. 235–347.

Kallianpur, G. and Xiong, J. (1995) *Stochastic Differential Equations in Infinite Dimensions*, IMS Lecture Notes, Vol. **26**, Hayward, CA.

Karatazas, I. and Shreve, S. (1988) *Brownian Motion and Stochastic Calculus*, Springer, Berlin.

Kent, J. T. and Wood, A. T. A. (1997) Estimating the fractal dimension of a locally self-similar Gaussian process by using increments, *J. R. Stat. Soc. B*, **59**, 679–700.

Kleptsyna, M. L. and Le Breton, A. (2002a) Statistical analysis of the fractional Ornstein-Uhlenbeck type process, *Stat. Inference Stochastic Process*, **5**, 229–248.

Kleptsyna, M. L. and Le Breton, A. (2002b) Extension of the Kalman-Bucy filter to elementary linear systems with fractional Brownian noises, *Stat. Inference Stochastic Process*, **5**, 249–271.

Kleptsyna, M. L., Le Breton, A. and Roubaud, M.-C. (2000a) Parameter estimation and optimal filtering for fractional type stochastic systems, *Stat. Inference Stochastic Process*, **3**, 173–182.

Kleptsyna, M. L., Le Breton, A. and Roubaud, M.-C. (2000b) General approach to filtering with fractional Brownian noises: application to linear systems, *Stochastics Stochastic Rep.*, **71**, 119–140.

Klingenhofer, F. and Zahle, M. (1999) Ordinary differential equations with fractal noise, *Proc. Am. Math. Soc.*, **127**, 1021–1028.

Knopov, P. S. (1982) On certain properties of two-parameter stochastic integrals, *Sov. Math. Dokl.*, **25**, 271–273.

Krein, S. G., Petunin, Yu. I. and Semenov, E. M. (1982) *Interpolation of Linear Operators*, Vol. **54**, *Translations of Mathematical Monographs*, American Mathematical Society, Providence, RI.

Kuhn, T. and Linde, W. (2002) Optimal series representation of fractional Brownian sheets, *Bernoulli*, **8**, 669–696.

Kulinich, G. L. (1985) On the law of iterated logarithm for one dimensional diffusion processes. *Theory Probab. Appl.*, **29**, 563–566.

Kurchenko, O. O. (2003) A consistent estimator of the Hurst parameter for a fractional Brownian motion, *Theor. Probab. Math. Stat.*, **67**, 97–106.

Kutoyants, Yu. A. (1994) *Identification of Dynamical Systems with Small Noise*, Kluwer, Dordrecht.

Kutoyants, Yu. A. and Pilibossian, P. (1994) On minimum L_1-norm estimate of the parameter of the Ornstein-Uhlenbeck process, *Stat. Probab. Lett.*, **20**, 117–123.

Lamperti, J. (1962) Semi-stable stochastic processes, *Trans. Am. Math. Soc.*, **104**, 62–78.

Langevin, P. (1908) Sur la theorie du mouvement Brownien, *C.R. Acad. Sci. Paris*, **146**, 530–533.

Le Breton, A. (1998) Filtering and parameter estimation in a simple linear model driven by a fractional Brownian motion, *Stat. Probab. Lett.*, **38**, 263–274.

Leger, S. (2001) Analyse stochastiques des signaux multifractaux et estimation de parametres, Thèse de Doctorat, Université d'Orléans.

Leland, W. E., Taqqu, M. S., Willinger, W. and Wilson, D. V. (1994) On the self-similar nature of Ethernet traffic (extended version), *IEEE/ACM Trans. Networking*, **2**, 1–15.

Lin, S. J. (1995) Stochastic analysis of fractional Brownian motions, *Stochastics Stochastic Rep.*, **55**, 121–140.

Liptser, R. S. (1980) A strong law of large numbers, *Stochastics*, **3**, 217–228.

Liptser, R. S. and Shiryayev, A. N. (1989) *The Theory of Martingales*, Kluwer, Dordrecht.

Liptser, R. S. and Shiryayev, A. N. (2001) *Statistics of Random Processes II: Applications*, Second Edition, Springer, Berlin.

Lyons, T. (1994) Differential equations driven by rough signals, I: an extension of an inequality of L.C. Young, *Math. Res. Lett.*, **1**, 451–464.

Maitre, H. and Pinciroli, M. (1999) Fractal characterization of a hydrological basin using SAR satellite images, *IEEE Trans. Geosci. Remote Sensing*, **37**, 175–181.

Mandelbrot, B. B. (1971) A fast fractional Gaussian noise generator, *Water Resour. Res.*, **7**, 543–553.

Mandelbrot, B. B. (1982) *The Fractal Geometry of Nature*, W.H. Freeman, San Francisco.

Mandelbrot, B. B. and Van Ness, J. (1968) Fractional Brownian motions, fractional noises and applications, *SIAM Rev*., **10**, 422–437.

Mandelbrot, B. B and Wallis, J. R. (1969) Computer experiments with fractional Gaussian noise, *Water Resour. Res*., **5**, 228–267.

Maslowski, B. and Pospisil, J. (2008) Ergodicity and parameter estimates for infinite-dimensional fractional Ornstein-Uhlenbeck process, *Appl. Math. Optim*., **57**, 401–429.

McCoy, E. J. and Walden, A. T. (1996) Wavelet analysis and synthesis of stationary long memory processes, *J. Comput. Graphical Stat*., **5**, 26–56.

Memin, J., Mishura, Y. and Valkeila, E. (2001) Inequalities for the moments of Wiener integrals with respect to a fractional Brownian motion, *Stat. Probab. Lett*., **51**, 197–206.

Miao, Y., Ren, W. and Ren, Z. (2008) On the fractional mixed fractional Brownian motion, *Appl. Math. Sci*., **2**, 1729–1738.

Michel, R. and Pfanzagl, J. (1971) The accuracy of the normal approximation for minimum contrast estimate, *Z. Wahrs. verw Gebiete*, **18**, 73–84.

Millar, P. W. (1984) A general approach to the optimality of the minimum distance estimators, *Trans. Am. Math. Soc*. **286**, 249–272.

Mishra, M. N. and Prakasa Rao, B. L. S. (2006) Large deviation probabilities for MLE and BE for a parameter for fractional Ornstein-Uhlenbeck type process, *Bull. Inf. Cybern*., **38**, 71–83.

Mishra, M. N. and Prakasa Rao, B. L. S. (2008) ϵ-upper and lower functions for maximum likelihood estimator for processes driven by fractional Brownian motion, *Random Oper. Stochastic Equations*, **16**, 165–180.

Mishra, M. N. and Prakasa Rao, B. L. S. (2009a) Nonparametric estimation of trend for stochastic differential equations driven by fractional Brownian motion, Preprint, University of Hyderabad.

Mishra, M. N. and Prakasa Rao, B. L. S. (2009b) Estimation for fractional diffusion processes under misspecified models, Preprint, University of Hyderabad.

Mishura, Y. (2008) *Stochastic Calculus for Fractional Brownian Motion and Related Processes*, Lecture Notes in Mathematics No. 1929, Springer, Berlin.

Mishura, Y. and Rudomino-Dusyatska, N. (2004) Consistency of drift parameter estimates in fractional Brownian diffusion models, Research paper No. 04-2004, Acturial Education and Reference Centre of the Ukrainian Actuarial Society, Lviv, Ukraine.

Mohammed, S.-E. A. and Scheutzow, M. K. R. (1990) Lyapunov exponents and stationary solutions for affine stochastic delay equations, *Stochastics Stochastic Rep*., **29**, 259–283.

Molchan, G. (1969) Gaussian processes with spectra which are asymptotically equivalent to a power of λ. *Theory Probab. Appl*., **14**, 530–532.

Montanari, A. (2003) Long-range dependence in hydrology, in *Theory of Long-Range Dependence*, ed. P. Doukhan, G. Oppenheim and M. S. Taqqu, Birkhauser, Boston, pp. 461–472.

Necula, C. (2002) Option pricing in a fractional Brownian motion environment, Preprint, DOFIN, Academy of Economic Studies, Bucharest.

Nguyen, H. T. and Pham, T. D. (1982) Identification of non-stationary diffusion model by the method of sieves, *SIAM J. Control Optim*., **20**, 603–611.

Norros, I. (2003) Large deviations of queues with long-range dependent input, in *Theory of Long-Range Dependence*, ed. P. Doukhan, G. Oppenheim and M. S. Taqqu, Birkhauser, Boston, pp. 409–415.

Norros, I., Valkeila, E. and Virtamo, J. (1999) An elementary approach to a Girsanov type formula and other analytical results on fractional Brownian motion, *Bernoulli*, **5**, 571–587.

Novikov, A. A. (1972) Sequential estimation of the parameters of diffusion processes, *Math. Notes*, **12**, 812–818.

Novikov, A. A. and Valkeila, E. (1999) On some maximal inequalities for fractional Brownian motion, *Stat. Probab. Lett*., **44**, 47–54.

Nualart, D. and Ortiz, S. (2008) Central limit theorems for multiple stochastic integrals and Malliavin calculus, *Stochastic Process. Appl*., **118**, 614–628.

Nualart, D. and Rascanu, A. (2002) Differential equations driven by fractional Brownian motion, *Collect. Math*., **53**, 55–81.

Nuzman, C. J. and Poor, H. V. (2000) Linear estimation of self-similar processes via Lamperti's transformation, *J. Appl. Probab*., **37**, 429–452.

Nuzman, C. J. and Poor, H. V. (2001) Reproducing kernel Hilbert space methods for wide-sense self-similar processes, *Ann. Appl. Probab*., **11**, 1199–1219.

Ossandik, S. M., Buldyrev, S. V., Goldberger, A. L., Havlin, S., Mantegna, R. N., Peng, C.-K., Simons, M. and Stanley, H. E. (1994) Correlation approach to identify coding regions in DNA sequences, *Biophys. J*., **67**, 64–70.

Peng, C.-K., Buldyrev, S. V., Goldberger, A. L., Havlin, S., Sciortino, F., Simons, M. and Stanley, H. E. (1992) Long-range correlation in nucleotide sequences, *Nature*, **356**, 168–170.

Peng, C.-K., Buldyrev, S. V., Simons, M., Stanley, H. E. and Goldberger, A. L. (1994) Mosaic organization of DNA nucleotides, *Phys. Rev. E*, **49**, 1685–1689.

Peng, C.-K., Hausdorff, J. M., Mietus, J. E., Havlin, S., Stanley, H. E. and Goldberger, A. L. (1995a) Fractals in physiological control from heartbeat to gait, in *Levy Flights and Related Phenomena in Physics*, ed. M. F. Shlesinger, G. M. Zaslavsky and U. Frisch, *Proceedings of the 1994 International Conference on Levy Flights*, Springer, Berlin, pp. 315–330.

Peng, C.-K., Havlin, S., Stanley, H. E. and Goldberger, A. L. (1995b) Quantification of scaling exponents and crossover phenomena in non-stationary heartbeat time series, *Chaos*, **5**, 82–87.

Percival, D. P. and Guttorp, P. (1994) Long-memory processes, the Allan variance and wavelets, in *Wavelets in Geophysics*, ed. E. Foufoula-Georgiou and P. Kumar, Academic Press, New York, pp. 325–357.

Pesquet-Popescu, B. and Levy Vehel, J. (2002) Stochastic fractal models for image processing models, *IEEE Signal Process Mage*, September 2002, pp. 48–62.

Pipiras, V. and Taqqu, M. S. (2000) Integration questions related to fractional Brownian motion, *Probab. Theory Relat. Fields*, **118**, 121–291.

Pipiras, V. and Taqqu, M. S. (2002) Deconvolution of fractional Brownian motion, *J. Time Ser. Anal*., **23**, 487–501.

Poggi, J.-M. and Viano, M.-C. (1998) An estimate of the fractal index using multi-scale aggregates, *J. Time Ser. Anal*., **19**, 221–233.

Prakasa Rao, B. L. S. (1981) The Bernstein-von Mises theorem for a class of diffusion processes, *Teor. Sluch. Proc.*, **9**, 95–101 (in Russian).

Prakasa Rao, B. L. S. (1984) On Bayes estimation for diffusion fields, in *Statistics: Applications and New Directions*, ed. J. K. Ghosh and J. Roy, Indian Statistical Institute, Calcutta, pp. 504–511.

Prakasa Rao, B. L. S. (1987) *Asymptotic Theory of Statistical Inference*, John Wiley & Sons, Inc., New York.

Prakasa Rao, B. L. S. (1988) Statistical inference from sampled data for stochastic processes, in *Statistical Inference from Stochastic Processes*, Contemporary Mathematics, ed. N. U. Prabhu, Vol. **80**, AMS, Providence, RI, pp. 249–284.

Prakasa Rao, B. L. S. (1999a) *Statistical Inference for Diffusion Type Processes*, Arnold, London and Oxford University Press, New York.

Prakasa Rao, B. L. S. (1999b) *Semimartingales and Their Statistical Inference*, CRC Press, Boca Raton, FL and Chapman and Hall, London.

Prakasa Rao, B. L. S. (2000) Bayes estimation for stochastic partial differential equations, *J. Stat. Plann. Inference*, **91**, 511–524.

Prakasa Rao, B. L. S. (2001) Statistical inference for stochastic partial differential equations, in *Selected Proceedings of the Symposium on Inference for Stochastic Processes*, ed. I. V. Basawa, C. C. Heyde and R. L. Taylor, IMS Monograph Series, Vol. **37**, Institute of Mathematical Statistics, Beachwood, OH, pp. 47–70.

Prakasa Rao, B. L. S. (2002) On some problems of estimation for some stochastic partial differential equations, in *Uncertainty and Optimality*, ed. J. C. Mishra, World Scientific, Singapore, pp. 71–153.

Prakasa Rao, B. L. S. (2003) Parameter estimation for linear stochastic differential equations driven by fractional Brownian motion, *Random Oper. Stochastic Equations*, **11**, 229–242.

Prakasa Rao, B. L. S. (2004a) Minimum L_1-norm estimation for fractional Ornstein-Uhlenbeck type process, *Theor. Probab. Math. Stat.*, **71**, 160–168.

Prakasa Rao, B. L. S. (2004b) Sequential estimation for fractional Ornstein-Uhlenbeck type process, *Sequential Anal.*, **23**, 33–44.

Prakasa Rao, B. L. S. (2004c) Identification for linear stochastic systems driven by fractional Brownian motion, *Stochastic Anal. Appl.*, **22**, 1487–1509.

Prakasa Rao, B. L. S. (2004d) Self-similar processes, fractional Brownian motion and statistical inference, *Festschrift for Herman Rubin*, ed. A. Das Gupta, Lecture Notes and Monograph Series, Vol. **45**, Institute of Mathematical Statistics, Beachwood, OH, pp. 98–125.

Prakasa Rao, B. L. S. (2004e) Parameter estimation for some stochastic partial differential equations driven by infinite dimensional fractional Brownian motion, *Theory Stochastic Process*, **10**, (26), 116–125.

Prakasa Rao, B. L. S. (2005a) Berry-Esseen bound for MLE for linear stochastic differential equations driven by fractional Brownian motion, *J. Korean Stat. Soc.*, **34**, 281–295.

Prakasa Rao, B. L. S. (2005b) Sequential testing for simple hypotheses for processes driven by fractional Brownian motion, *Sequential Anal.*, **24**, 189–203.

Prakasa Rao, B. L. S. (2005c) Estimation for translation of a process driven by fractional Brownian motion, *Stochastic Anal. Appl.*, **23**, 1199–1212.

Prakasa Rao, B. L. S. (2007) Instrumental variable estimation for linear stochastic differential equations driven by fractional Brownian motion, *Stochastic Anal. Appl.*, **25**, 1203–1215.

Prakasa Rao, B. L. S. (2008a) Parametric estimation for linear system of stochastic differential equations driven by fractional Brownian motions with different Hurst indices, *Theor. Probab. Math. Stat.*, **79**, 129–137.

Prakasa Rao, B. L. S. (2008b) Parameter estimation for linear stochastic delay differential equations driven by fractional Brownian motion, *Random Oper. Stochastic Equations*, **16**, 27–38.

Prakasa Rao, B. L. S. (2008c) Singularity of fractional motions with different Hurst indices, *Stochastic Anal. Appl.*, **26**, 334–337.

Prakasa Rao, B. L. S. (2009) Estimation for stochastic differential equations driven by mixed fractional Brownian motions, *Calcutta Statistical Association Bulletin*, **61** (Special 6th Triennial Proceedings Volume), pp. 141–153.

Rao, M. M. (2000) *Stochastic Processes: Inference Theory*, Kluwer, Dordrecht.

Revuz, D. and Yor, M. (1991) *Continuous Martingales and Brownian Motion*, Springer, Berlin.

Robinson, P. (1995) Gaussian semiparametric estimation of long range dependence, *Ann. Stat.*, **23**, 1630–1661.

Rogers, L. C. G. (1997) Arbitrage with fractional Brownian motion, *Math. Financ.*, **7**, 95–105.

Rozovskii, B. L. (1990) *Stochastic Evolution Systems*, Kluwer, Dordrecht.

Rudomino-Dusyatska, N. (2003) Properties of maximum likelihood estimates in diffusion and fractional-Brownian models, *Theor. Probability and Math. Stat.*, **68**, 139–146.

Samko, S. G., Kilbas, A. A. and Marichev, O. I. (1993) *Fractional Integrals and Derivatives: Theory and Applications*, Gordon and Breach Science, Yverdon.

Samorodnitsky, G. and Taqqu, M. S. (1994) *Stable Non-Gaussian Processes: Stochastic Models with Infinite Variance*, Chapman and Hall, London.

Skorokhod, A. V. (1965) *Studies in the Theory of Random Processes*, Addison-Wesley, Reading, MA.

Slominski, L. and Ziemkiewicz, B. (2005) Inequalities for the L^p norms of integrals with respect to a fractional Brownian motion, *Stat. Probab. Lett.*, **73**, 79–90.

Smith, H. F. (1938) An empirical law describing heterogeneity in the yields of agricultural crops, *J. Agric. Sci.*, **28**, 1–23.

Sorensen, M. (1991) Likelihood methods for diffusions with jumps, *Statistical Inference in stochastic Processes*, ed. N. U. Prabhu and I. V. Basawa, Marcel Dekker, New York, pp. 67–105.

Sottinen, T (2001) Fractional Brownian motion, random walks and binary market models, *Financ. Stochasties*, **5**, 343–355.

Sottinen, T. and Tudor, C. (2008) Parameter estimation for stochastic equations with additive fractional Brownian sheet, *Stat. Inference Stochastic Process*, **11**, 221–236.

Stein, E. M. (1971) *Singular Integrals and Differentiability*, Princeton University Press, Princeton, NJ.

Stewart, C. V., Moghaddam, B., Hintz, K. J. and Novak, L. M. (1993) Fractional Brownian motion models for synthetic aperture radar imagery scene segmentation, *Proc. IEEE*, **81**, 1511–1521.

Taqqu, M. S. (1975) Weak convergence to fractional Brownian motion and to the Rosenblatt process, *Z. Wahrsch. Verw. Geb*., **31**, 287–302.

Taqqu, M. S., Teverovsky, V. and Willinger, W. (1995) Estimators for long-range dependence: an empirical study, *Fractals*, **3**, 785–798.

Taylor, C. C. and Taylor, S. J. (1991) Estimating the dimension of a fractal, *J. R. Stat. Soc. B*, **53**, 353–364.

Tindel, S., Tudor, C. and Viens, F. (2003) Stochastic evolution equations with fractional Brownian motion, *Probab. Theory Relat. Fields*, **127**, 186–204.

Tudor, M. and Tudor, C. (2002) On the multiparameter fractional Brownian motion, Preprint.

Uhlenbeck, G. E. and Ornstein, L. S. (1930) On the theory of the Brownian motion, *Phys. Rev*., **36**, 823–841.

Wang, Y. (1999) Generalized fractional Brownian motion, Technical Report No. 99–11, Department of Statistics, University of Connecticut.

Wang, Y., Cavanaugh, J. E. and Song, C. (2001) Self-similarity index estimation via wavelets for locally self-similar processes, *J. Stat. Plann. Inference*, **99**, 91–110.

Watson, G. N. (1995) *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, Cambridge.

Willinger, W., Taqqu, M. S., Sherman, R. and Wilson, D. (1997) Self-similarity through high variability: statistical analysis of Ethernet LAN traffic at the source level, *IEEE/ACM Trans. Networking*, **5**, 71–86.

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

Willinger, W., Paxson, V., Riedi, R. H. and Taqqu, M. S. (2003) Long-range dependence and data network traffic, in *Theory of Long-Range Dependence*, ed. P. Doukhan, G. Oppenheim and M. S. Taqqu, Birkhauser, Boston, pp. 373–407.

Wong, E. and Zakai, M. (1976) Weak martingales and stochastic integrals in the plane, *Ann. Probab*., **4**, 570–586.

Xiong, J. and Zhao, X. (2005) Nonlinear filtering with fractional Brownian motion noise, *Stochastic Anal. Appl*., **23**, 55–67.

Yazici, B. and Kashyap, R. L. (1997) A class of second-order stationary self-similar processes for $1/f$ phenomena, *IEEE Trans. Signal Process*, **45**, 396–410.

Zahle, M. (1998) Integration with respect to fractal functions and stochastic calculus I, *Prob. Theory Relat. Fields*, **111**, 333–374.

Zili, M. (2006) On the mixed fractional Brownian motion, *J. Appl. Math. Stochastic Anal*., DOI 10.1155/JAMSA/2006/32435.

Index

Statistical Inference for Fractional Diffusion Processes B. L. S. Prakasa Rao

WILEY SERIES IN PROBABILITY AND STATISTICS

Established by WALTER A. SHEWHART and SAMUEL S. WILKS

The Wiley Series in Probability and Statistics is well established and authoritative. It covers many topics of current research interest in both pure and applied statistics and probability theory. Written by leading statisticians and institutions, the titles span both state-of-the-art developments in the field and classical methods.

Reflecting the wide range of current research in statistics, the series encompasses applied, methodological and theoretical statistics, ranging from applications and new techniques made possible by advances in computerized practice to rigorous treatment of theoretical approaches.

This series provides essential and invaluable reading for all statisticians, whether in academia, industry, government, or research.

ABRAHAM and LEDOLTER · Statistical Methods for Forecasting
AGRESTI · Analysis of Ordinal Categorical Data
AGRESTI · An Introduction to Categorical Data Analysis
AGRESTI · Categorical Data Analysis, Second Edition
ALTMAN, GILL and McDONALD · Numerical Issues in Statistical Computing for the Social Scientist
AMARATUNGA and CABRERA · Exploration and Analysis of DNA Microarray and Protein Array Data
ANDĚL · Mathematics of Chance
ANDERSON · An Introduction to Multivariate Statistical Analysis, Third Edition
*ANDERSON · The Statistical Analysis of Time Series
ANDERSON, AUQUIER, HAUCK, OAKES, VANDAELE and WEISBERG · Statistical Methods for Comparative Studies
ANDERSON and LOYNES · The Teaching of Practical Statistics
ARMITAGE and DAVID (editors) · Advances in Biometry
ARNOLD, BALAKRISHNAN and NAGARAJA · Records
*ARTHANARI and DODGE · Mathematical Programming in Statistics
*BAILEY · The Elements of Stochastic Processes with Applications to the Natural Sciences
BALAKRISHNAN and KOUTRAS · Runs and Scans with Applications
BALAKRISHNAN and NG · Precedence-Type Tests and Applications
BARNETT · Comparative Statistical Inference, Third Edition
BARNETT · Environmental Statistics: Methods & Applications

*Now available in a lower priced paperback edition in the Wiley Classics Library.

BARNETT and LEWIS · Outliers in Statistical Data, Third Edition
BARTOSZYNSKI and NIEWIADOMSKA-BUGAJ · Probability and Statistical Inference
BASILEVSKY · Statistical Factor Analysis and Related Methods: Theory and Applications
BASU and RIGDON · Statistical Methods for the Reliability of Repairable Systems
BATES and WATTS · Nonlinear Regression Analysis and Its Applications
BECHHOFER, SANTNER and GOLDSMAN · Design and Analysis of Experiments for Statistical Selection, Screening and Multiple Comparisons
BEIRLANT, GOEGEBEUR, SEGERS, TEUGELS and DE WAAL · Statistics of Extremes: Theory and Applications
BELSLEY · Conditioning Diagnostics: Collinearity and Weak Data in Regression
BELSLEY, KUH and WELSCH · Regression Diagnostics: Identifying Influential Data and Sources of Collinearity
BENDAT and PIERSOL · Random Data: Analysis and Measurement Procedures, Third Edition
BERNARDO and SMITH · Bayesian Theory
BERRY, CHALONER and GEWEKE · Bayesian Analysis in Statistics and Econometrics: Essays in Honor of Arnold Zellner
BHAT and MILLER · Elements of Applied Stochastic Processes, Third Edition
BHATTACHARYA and JOHNSON · Statistical Concepts and Methods
BHATTACHARYA and WAYMIRE · Stochastic Processes with Applications
BIEMER, GROVES, LYBERG, MATHIOWETZ and SUDMAN · Measurement Errors in Surveys
BILLINGSLEY · Convergence of Probability Measures, Second Edition
BILLINGSLEY · Probability and Measure, Third Edition
BIRKES and DODGE · Alternative Methods of Regression
BISWAS, DATTA, FINE and SEGAL · Statistical Advances in the Biomedical Sciences: Clinical Trials, Epidemiology, Survival Analysis, and Bioinformatics
BLISCHKE and MURTHY (editors) · Case Studies in Reliability and Maintenance
BLISCHKE and MURTHY · Reliability: Modeling, Prediction and Optimization
BLOOMFIELD · Fourier Analysis of Time Series: An Introduction, Second Edition
BOLLEN · Structural Equations with Latent Variables
BOLLEN and CURRAN · Latent Curve Models: A Structural Equation Perspective
BOROVKOV · Ergodicity and Stability of Stochastic Processes
BOSQ and BLANKE · Inference and Prediction in Large Dimensions
BOULEAU · Numerical Methods for Stochastic Processes
BOX · Bayesian Inference in Statistical Analysis
BOX · R. A. Fisher, the Life of a Scientist
BOX and DRAPER · Empirical Model-Building and Response Surfaces
*BOX and DRAPER · Evolutionary Operation: A Statistical Method for Process Improvement
BOX · Improving Almost Anything *Revised Edition*
BOX, HUNTER and HUNTER · Statistics for Experimenters: An Introduction to Design, Data Analysis and Model Building

*Now available in a lower priced paperback edition in the Wiley Classics Library.

BOX, HUNTER and HUNTER · Statistics for Experimenters: Design, Innovation and Discovery, Second Edition
BOX and LUCEÑO · Statistical Control by Monitoring and Feedback Adjustment
BRANDIMARTE · Numerical Methods in Finance: A MATLAB-Based Introduction
BROWN and HOLLANDER · Statistics: A Biomedical Introduction
BRUNNER, DOMHOF and LANGER · Nonparametric Analysis of Longitudinal Data in Factorial Experiments
BUCKLEW · Large Deviation Techniques in Decision, Simulation and Estimation
CAIROLI and DALANG · Sequential Stochastic Optimization
CASTILLO, HADI, BALAKRISHNAN and SARABIA · Extreme Value and Related Models with Applications in Engineering and Science
CHAN · Time Series: Applications to Finance
CHARALAMBIDES · Combinatorial Methods in Discrete Distributions
CHATTERJEE and HADI · Regression Analysis by Example, Fourth Edition
CHATTERJEE and HADI · Sensitivity Analysis in Linear Regression
CHERNICK · Bootstrap Methods: A Practitioner's Guide
CHERNICK and FRIIS · Introductory Biostatistics for the Health Sciences
CHILÉS and DELFINER · Geostatistics: Modeling Spatial Uncertainty
CHOW and LIU · Design and Analysis of Clinical Trials: Concepts and Methodologies, Second Edition
CLARKE · Linear Models: The Theory and Application of Analysis of Variance
CLARKE and DISNEY · Probability and Random Processes: A First Course with Applications, Second Edition
*COCHRAN and COX · Experimental Designs, Second Edition
CONGDON · Applied Bayesian Modelling
CONGDON · Bayesian Models for Categorical Data
CONGDON · Bayesian Statistical Modelling, Second Edition
CONOVER · Practical Nonparametric Statistics, Second Edition
COOK · Regression Graphics
COOK and WEISBERG · An Introduction to Regression Graphics
COOK and WEISBERG · Applied Regression Including Computing and Graphics
CORNELL · Experiments with Mixtures, Designs, Models and the Analysis of Mixture Data, Third Edition
COVER and THOMAS · Elements of Information Theory
COX · A Handbook of Introductory Statistical Methods
*COX · Planning of Experiments
CRESSIE · Statistics for Spatial Data, Revised Edition
CSÖRGÖ and HORVÁTH · Limit Theorems in Change Point Analysis
DANIEL · Applications of Statistics to Industrial Experimentation
DANIEL · Biostatistics: A Foundation for Analysis in the Health Sciences, Sixth Edition
*DANIEL · Fitting Equations to Data: Computer Analysis of Multifactor Data, Second Edition
DASU and JOHNSON · Exploratory Data Mining and Data Cleaning
DAVID and NAGARAJA · Order Statistics, Third Edition
*DEGROOT, FIENBERG and KADANE · Statistics and the Law

*Now available in a lower priced paperback edition in the Wiley Classics Library.

DEL CASTILLO · Statistical Process Adjustment for Quality Control
DEMARIS · Regression with Social Data: Modeling Continuous and Limited Response Variables
DEMIDENKO · Mixed Models: Theory and Applications
DENISON, HOLMES, MALLICK and SMITH · Bayesian Methods for Nonlinear Classification and Regression
DETTE and STUDDEN · The Theory of Canonical Moments with Applications in Statistics, Probability and Analysis
DEY and MUKERJEE · Fractional Factorial Plans
DILLON and GOLDSTEIN · Multivariate Analysis: Methods and Applications
DODGE · Alternative Methods of Regression
*DODGE and ROMIG · Sampling Inspection Tables, Second Edition
*DOOB · Stochastic Processes
DOWDY, WEARDEN and CHILKO · Statistics for Research, Third Edition
DRAPER and SMITH · Applied Regression Analysis, Third Edition
DRYDEN and MARDIA · Statistical Shape Analysis
DUDEWICZ and MISHRA · Modern Mathematical Statistics
DUNN and CLARK · Applied Statistics: Analysis of Variance and Regression, Second Edition
DUNN and CLARK · Basic Statistics: A Primer for the Biomedical Sciences, Third Edition
DUPUIS and ELLIS · A Weak Convergence Approach to the Theory of Large Deviations
EDLER and KITSOS (editors) · Recent Advances in Quantitative Methods in Cancer and Human Health Risk Assessment
*ELANDT-JOHNSON and JOHNSON · Survival Models and Data Analysis
ENDERS · Applied Econometric Time Series
ETHIER and KURTZ · Markov Processes: Characterization and Convergence
EVANS, HASTINGS and PEACOCK · Statistical Distribution, Third Edition
FELLER · An Introduction to Probability Theory and Its Applications, Volume I, Third Edition, Revised; Volume II, Second Edition
FISHER and VAN BELLE · Biostatistics: A Methodology for the Health Sciences
FITZMAURICE, LAIRD and WARE · Applied Longitudinal Analysis
*FLEISS · The Design and Analysis of Clinical Experiments
FLEISS · Statistical Methods for Rates and Proportions, Second Edition
FLEMING and HARRINGTON · Counting Processes and Survival Analysis
FULLER · Introduction to Statistical Time Series, Second Edition
FULLER · Measurement Error Models
GALLANT · Nonlinear Statistical Models.
GEISSER · Modes of Parametric Statistical Inference
GELMAN and MENG (editors) · Applied Bayesian Modeling and Casual Inference from Incomplete-data Perspectives
GEWEKE · Contemporary Bayesian Econometrics and Statistics
GHOSH, MUKHOPADHYAY and SEN · Sequential Estimation
GIESBRECHT and GUMPERTZ · Planning, Construction and Statistical Analysis of Comparative Experiments

*Now available in a lower priced paperback edition in the Wiley Classics Library.

GIFI · Nonlinear Multivariate Analysis
GIVENS and HOETING · Computational Statistics
GLASSERMAN and YAO · Monotone Structure in Discrete-Event Systems
GNANADESIKAN · Methods for Statistical Data Analysis of Multivariate Observations, Second Edition
GOLDSTEIN and WOLFF · Bayes Linear Statistics, Theory & Methods
GOLDSTEIN and LEWIS · Assessment: Problems, Development and Statistical Issues
GREENWOOD and NIKULIN · A Guide to Chi-Squared Testing
GROSS, SHORTLE, THOMPSON and HARRIS · Fundamentals of Queueing Theory, fourth Edition
GROSS, SHORTLE, THOMPSON and HARRIS · Solutions Manual to Accompany Fundamentals of Queueing Theory, fourth Edition
*HAHN and SHAPIRO · Statistical Models in Engineering
HAHN and MEEKER · Statistical Intervals: A Guide for Practitioners
HALD · A History of Probability and Statistics and their Applications Before 1750
HALD · A History of Mathematical Statistics from 1750 to 1930
HAMPEL · Robust Statistics: The Approach Based on Influence Functions
HANNAN and DEISTLER · The Statistical Theory of Linear Systems
HARTUNG, KNAPP and SINHA · Statistical Meta-Analysis with Applications
HEIBERGER · Computation for the Analysis of Designed Experiments
HEDAYAT and SINHA · Design and Inference in Finite Population Sampling
HEDEKER and GIBBONS · Longitudinal Data Analysis
HELLER · MACSYMA for Statisticians
HERITIER, CANTONI, COPT and VICTORIA-FESER · Robust Methods in Biostatistics
HINKELMANN and KEMPTHORNE · Design and Analysis of Experiments, Volume 1: Introduction to Experimental Design
HINKELMANN and KEMPTHORNE · Design and analysis of experiments, Volume 2: Advanced Experimental Design
HOAGLIN, MOSTELLER and TUKEY · Exploratory Approach to Analysis of Variance
HOAGLIN, MOSTELLER and TUKEY · Exploring Data Tables, Trends and Shapes
*HOAGLIN, MOSTELLER and TUKEY · Understanding Robust and Exploratory Data Analysis
HOCHBERG and TAMHANE · Multiple Comparison Procedures
HOCKING · Methods and Applications of Linear Models: Regression and the Analysis of Variance, Second Edition
HOEL · Introduction to Mathematical Statistics, Fifth Edition
HOGG and KLUGMAN · Loss Distributions
HOLLANDER and WOLFE · Nonparametric Statistical Methods, Second Edition
HOSMER and LEMESHOW · Applied Logistic Regression, Second Edition
HOSMER and LEMESHOW · Applied Survival Analysis: Regression Modeling of Time to Event Data
HUBER · Robust Statistics
HUBERTY · Applied Discriminant Analysis
HUBERY and OLEJNKI · Applied MANOVA and Discriminant Analysis, 2nd Edition

*Now available in a lower priced paperback edition in the Wiley Classics Library.

HUNT and KENNEDY · Financial Derivatives in Theory and Practice, Revised Edition
HURD and MIAMEE · Periodically Correlated Random Sequences: Spectral Theory and Practice
HUSKOVA, BERAN and DUPAC · Collected Works of Jaroslav Hajek—with Commentary
HUZURBAZAR · Flowgraph Models for Multistate Time-to-Event Data
IMAN and CONOVER · A Modern Approach to Statistics
JACKMAN · Bayesian Analysis for the Social Sciences
JACKSON · A User's Guide to Principle Components
JOHN · Statistical Methods in Engineering and Quality Assurance
JOHNSON · Multivariate Statistical Simulation
JOHNSON and BALAKRISHNAN · Advances in the Theory and Practice of Statistics: A Volume in Honor of Samuel Kotz
JOHNSON and BHATTACHARYYA · Statistics: Principles and Methods, Fifth Edition
JOHNSON and KOTZ · Distributions in Statistics
JOHNSON and KOTZ (editors) · Leading Personalities in Statistical Sciences: From the Seventeenth Century to the Present
JOHNSON, KOTZ and BALAKRISHNAN · Continuous Univariate Distributions, Volume 1, Second Edition
JOHNSON, KOTZ and BALAKRISHNAN · Continuous Univariate Distributions, Volume 2, Second Edition
JOHNSON, KOTZ and BALAKRISHNAN · Discrete Multivariate Distributions
JOHNSON, KOTZ and KEMP · Univariate Discrete Distributions, Second Edition
JUDGE, GRIFFITHS, HILL, LU TKEPOHL and LEE · The Theory and Practice of Econometrics, Second Edition
JUREEKOVÁ and SEN · Robust Statistical Procedures: Asymptotics and Interrelations
JUREK and MASON · Operator-Limit Distributions in Probability Theory
KADANE · Bayesian Methods and Ethics in a Clinical Trial Design
KADANE and SCHUM · A Probabilistic Analysis of the Sacco and Vanzetti Evidence
KALBFLEISCH and PRENTICE · The Statistical Analysis of Failure Time Data, Second Edition
KARIYA and KURATA · Generalized Least Squares
KASS and VOS · Geometrical Foundations of Asymptotic Inference
KAUFMAN and ROUSSEEUW · Finding Groups in Data: An Introduction to Cluster Analysis
KEDEM and FOKIANOS · Regression Models for Time Series Analysis
KENDALL, BARDEN, CARNE and LE · Shape and Shape Theory
KHURI · Advanced Calculus with Applications in Statistics, Second Edition
KHURI, MATHEW and SINHA · Statistical Tests for Mixed Linear Models
*KISH · Statistical Design for Research
KLEIBER and KOTZ · Statistical Size Distributions in Economics and Actuarial Sciences
KLUGMAN, PANJER and WILLMOT · Loss Models: From Data to Decisions
KLUGMAN, PANJER and WILLMOT · Solutions Manual to Accompany Loss Models: From Data to Decisions

*Now available in a lower priced paperback edition in the Wiley—Interscience Paperback Series.

KOSKI and NOBLE · Bayesian Networks: An Introduction
KOTZ, BALAKRISHNAN and JOHNSON · Continuous Multivariate Distributions, Volume 1, Second Edition
KOTZ and JOHNSON (editors) · Encyclopedia of Statistical Sciences: Volumes 1 to 9 with Index
KOTZ and JOHNSON (editors) · Encyclopedia of Statistical Sciences: Supplement Volume
KOTZ, READ and BANKS (editors) · Encyclopedia of Statistical Sciences: Update Volume 1
KOTZ, READ and BANKS (editors) · Encyclopedia of Statistical Sciences: Update Volume 2
KOVALENKO, KUZNETZOV and PEGG · Mathematical Theory of Reliability of Time-Dependent Systems with Practical Applications
KOWALSI and TU · Modern Applied U-Statistics
KROONENBERG · Applied Multiway Data Analysis
KULINSKAYA, MORGENTHALER and STAUDTE · Meta Analysis: A Guide to Calibrating and Combining Statistical Evidence
KUROWICKA and COOKE · Uncertainty Analysis with High Dimensional Dependence Modelling
KVAM and VIDAKOVIC · Nonparametric Statistics with Applications to Science and Engineering
LACHIN · Biostatistical Methods: The Assessment of Relative Risks
LAD · Operational Subjective Statistical Methods: A Mathematical, Philosophical and Historical Introduction
LAMPERTI · Probability: A Survey of the Mathematical Theory, Second Edition
LANGE, RYAN, BILLARD, BRILLINGER, CONQUEST and GREENHOUSE · Case Studies in Biometry
LARSON · Introduction to Probability Theory and Statistical Inference, Third Edition
LAWLESS · Statistical Models and Methods for Lifetime Data, Second Edition
LAWSON · Statistical Methods in Spatial Epidemiology, Second Edition
LE · Applied Categorical Data Analysis
LE · Applied Survival Analysis
LEE · Structural Equation Modelling: A Bayesian Approach
LEE and WANG · Statistical Methods for Survival Data Analysis, Third Edition
LEPAGE and BILLARD · Exploring the Limits of Bootstrap
LEYLAND and GOLDSTEIN (editors) · Multilevel Modelling of Health Statistics
LIAO · Statistical Group Comparison
LINDVALL · Lectures on the Coupling Method
LIN · Introductory Stochastic Analysis for Finance and Insurance
LINHART and ZUCCHINI · Model Selection
LITTLE and RUBIN · Statistical Analysis with Missing Data, Second Edition
LLOYD · The Statistical Analysis of Categorical Data
LOWEN and TEICH · Fractal-Based Point Processes
MAGNUS and NEUDECKER · Matrix Differential Calculus with Applications in Statistics and Econometrics, Revised Edition
MALLER and ZHOU · Survival Analysis with Long Term Survivors
MALLOWS · Design, Data and Analysis by Some Friends of Cuthbert Daniel
MANN, SCHAFER and SINGPURWALLA · Methods for Statistical Analysis of Reliability and Life Data

MANTON, WOODBURY and TOLLEY · Statistical Applications Using Fuzzy Sets
MARCHETTE · Random Graphs for Statistical Pattern Recognition
MARKOVICH · Nonparametric Analysis of Univariate Heavy-Tailed Data: Research and practice
MARDIA and JUPP · Directional Statistics
MARKOVICH · Nonparametric Analysis of Univariate Heavy-Tailed Data: Research and Practice
MARONNA, MARTIN and YOHAI · Robust Statistics: Theory and Methods
MASON, GUNST and HESS · Statistical Design and Analysis of Experiments with Applications to Engineering and Science, Second Edition
MCCULLOCH and SERLE · Generalized, Linear and Mixed Models
MCFADDEN · Management of Data in Clinical Trials
MCLACHLAN · Discriminant Analysis and Statistical Pattern Recognition
MCLACHLAN, DO and AMBROISE · Analyzing Microarray Gene Expression Data
MCLACHLAN and KRISHNAN · The EM Algorithm and Extensions
MCLACHLAN and PEEL · Finite Mixture Models
MCNEIL · Epidemiological Research Methods
MEEKER and ESCOBAR · Statistical Methods for Reliability Data
MEERSCHAERT and SCHEFFLER · Limit Distributions for Sums of Independent Random Vectors: Heavy Tails in Theory and Practice
MICKEY, DUNN and CLARK · Applied Statistics: Analysis of Variance and Regression, Third Edition
*MILLER · Survival Analysis, Second Edition
MONTGOMERY, JENNINGS and KULAHCI · Introduction to Time Series Analysis and Forecasting Solutions Set
MONTGOMERY, PECK and VINING · Introduction to Linear Regression Analysis, Fourth Edition
MORGENTHALER and TUKEY · Configural Polysampling: A Route to Practical Robustness
MUIRHEAD · Aspects of Multivariate Statistical Theory
MULLER and STEWART · Linear Model Theory: Univariate, Multivariate and Mixed Models
MURRAY · X-STAT 2.0 Statistical Experimentation, Design Data Analysis and Nonlinear Optimization
MURTHY, XIE and JIANG · Weibull Models
MYERS and MONTGOMERY · Response Surface Methodology: Process and Product Optimization Using Designed Experiments, Second Edition
MYERS, MONTGOMERY and VINING · Generalized Linear Models. With Applications in Engineering and the Sciences
†NELSON · Accelerated Testing, Statistical Models, Test Plans and Data Analysis
†NELSON · Applied Life Data Analysis
NEWMAN · Biostatistical Methods in Epidemiology
OCHI · Applied Probability and Stochastic Processes in Engineering and Physical Sciences

*Now available in a lower priced paperback edition in the Wiley Classics Library.
†Now available in a lower priced paperback edition in the Wiley - Interscience Paperback Series.

OKABE, BOOTS, SUGIHARA and CHIU · Spatial Tesselations: Concepts and Applications of Voronoi Diagrams, Second Edition
OLIVER and SMITH · Influence Diagrams, Belief Nets and Decision Analysis
PALTA · Quantitative Methods in Population Health: Extentions of Ordinary Regression
PANJER · Operational Risks: Modeling Analytics
PANKRATZ · Forecasting with Dynamic Regression Models
PANKRATZ · Forecasting with Univariate Box-Jenkins Models: Concepts and Cases
PARDOUX · Markov Processes and Applications: Algorithms, Networks, Genome and Finance
PARMIGIANI and INOUE · Decision Theory: Principles and Approaches
*PARZEN · Modern Probability Theory and Its Applications
PEÑA, TIAO and TSAY · A Course in Time Series Analysis
PESARIN and SALMASO · Permutation Tests for Complex Data: Theory, Applications and Software
PIANTADOSI · Clinical Trials: A Methodologic Perspective
PORT · Theoretical Probability for Applications
POURAHMADI · Foundations of Time Series Analysis and Prediction Theory
POWELL · Approximate Dynamic Programming: Solving the Curses of Dimensionality
PRAKASA RAO · Statistical Inference for Fractional Diffusion Processes
PRESS · Bayesian Statistics: Principles, Models and Applications
PRESS · Subjective and Objective Bayesian Statistics, Second Edition
PRESS and TANUR · The Subjectivity of Scientists and the Bayesian Approach
PUKELSHEIM · Optimal Experimental Design
PURI, VILAPLANA and WERTZ · New Perspectives in Theoretical and Applied Statistics
PUTERMAN · Markov Decision Processes: Discrete Stochastic Dynamic Programming
QIU · Image Processing and Jump Regression Analysis
RAO · Linear Statistical Inference and its Applications, Second Edition
RAUSAND and HØYLAND · System Reliability Theory: Models, Statistical Methods and Applications, Second Edition
RENCHER · Linear Models in Statistics
RENCHER · Methods of Multivariate Analysis, Second Edition
RENCHER · Multivariate Statistical Inference with Applications
RIPLEY · Spatial Statistics
RIPLEY · Stochastic Simulation
ROBINSON · Practical Strategies for Experimenting
ROHATGI and SALEH · An Introduction to Probability and Statistics, Second Edition
ROLSKI, SCHMIDLI, SCHMIDT and TEUGELS · Stochastic Processes for Insurance and Finance
ROSENBERGER and LACHIN · Randomization in Clinical Trials: Theory and Practice
ROSS · Introduction to Probability and Statistics for Engineers and Scientists
ROSSI, ALLENBY and MCCULLOCH · Bayesian Statistics and Marketing
ROUSSEEUW and LEROY · Robust Regression and Outline Detection
ROYSTON and SAUERBREI · Multivariable Model - Building: A Pragmatic Approach to Regression Anaylsis based on Fractional Polynomials for Modelling Continuous Variables

*Now available in a lower priced paperback edition in the Wiley Classics Library.

RUBIN · Multiple Imputation for Nonresponse in Surveys
RUBINSTEIN · Simulation and the Monte Carlo Method, Second Edition
RUBINSTEIN and MELAMED · Modern Simulation and Modeling
RYAN · Modern Engineering Statistics
RYAN · Modern Experimental Design
RYAN · Modern Regression Methods
RYAN · Statistical Methods for Quality Improvement, Second Edition
SALEH · Theory of Preliminary Test and Stein-Type Estimation with Applications
SALTELLI, CHAN and SCOTT (editors) · Sensitivity Analysis
*SCHEFFE · The Analysis of Variance
SCHIMEK · Smoothing and Regression: Approaches, Computation and Application
SCHOTT · Matrix Analysis for Statistics
SCHOUTENS · Levy Processes in Finance: Pricing Financial Derivatives
SCHUSS · Theory and Applications of Stochastic Differential Equations
SCOTT · Multivariate Density Estimation: Theory, Practice and Visualization
*SEARLE · Linear Models
SEARLE · Linear Models for Unbalanced Data
SEARLE · Matrix Algebra Useful for Statistics
SEARLE and WILLETT · Matrix Algebra for Applied Economics
SEBER · Multivariate Observations
SEBER and LEE · Linear Regression Analysis, Second Edition
SEBER and WILD · Nonlinear Regression
SENNOTT · Stochastic Dynamic Programming and the Control of Queueing Systems
*SERFLING · Approximation Theorems of Mathematical Statistics
SHAFER and VOVK · Probability and Finance: Its Only a Game!
SILVAPULLE and SEN · Constrained Statistical Inference: Inequality, Order and Shape Restrictions
SINGPURWALLA · Reliability and Risk: A Bayesian Perspective
SMALL and MCLEISH · Hilbert Space Methods in Probability and Statistical Inference
SRIVASTAVA · Methods of Multivariate Statistics
STAPLETON · Linear Statistical Models
STAUDTE and SHEATHER · Robust Estimation and Testing
STOYAN, KENDALL and MECKE · Stochastic Geometry and Its Applications, Second Edition
STOYAN and STOYAN · Fractals, Random and Point Fields: Methods of Geometrical Statistics
STREET and BURGESS · The Construction of Optimal Stated Choice Experiments: Theory and Methods
STYAN · The Collected Papers of T. W. Anderson: 1943–1985
SUTTON, ABRAMS, JONES, SHELDON and SONG · Methods for Meta-Analysis in Medical Research
TAKEZAWA · Introduction to Nonparametric Regression
TAMHANE · Statistical Analysis of Designed Experiments: Theory and Applications
TANAKA · Time Series Analysis: Nonstationary and Noninvertible Distribution Theory
THOMPSON · Empirical Model Building
THOMPSON · Sampling, Second Edition

*Now available in a lower priced paperback edition in the Wiley Classics Library.

THOMPSON · Simulation: A Modeler's Approach
THOMPSON and SEBER · Adaptive Sampling
THOMPSON, WILLIAMS and FINDLAY · Models for Investors in Real World Markets
TIAO, BISGAARD, HILL, PEÑA and STIGLER (editors) · Box on Quality and Discovery: with Design, Control and Robustness
TIERNEY · LISP-STAT: An Object-Oriented Environment for Statistical Computing and Dynamic Graphics
TSAY · Analysis of Financial Time Series
UPTON and FINGLETON · Spatial Data Analysis by Example, Volume II: Categorical and Directional Data
VAN BELLE · Statistical Rules of Thumb
VAN BELLE, FISHER, HEAGERTY and LUMLEY · Biostatistics: A Methodology for the Health Sciences, Second Edition
VESTRUP · The Theory of Measures and Integration
VIDAKOVIC · Statistical Modeling by Wavelets
VINOD and REAGLE · Preparing for the Worst: Incorporating Downside Risk in Stock Market Investments
WALLER and GOTWAY · Applied Spatial Statistics for Public Health Data
WEERAHANDI · Generalized Inference in Repeated Measures: Exact Methods in MANOVA and Mixed Models
WEISBERG · Applied Linear Regression, Second Edition
WELSH · Aspects of Statistical Inference
WESTFALL and YOUNG · Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment
WHITTAKER · Graphical Models in Applied Multivariate Statistics
WINKER · Optimization Heuristics in Economics: Applications of Threshold Accepting
WONNACOTT and WONNACOTT · Econometrics, Second Edition
WOODING · Planning Pharmaceutical Clinical Trials: Basic Statistical Principles
WOODWORTH · Biostatistics: A Bayesian Introduction
WOOLSON and CLARKE · Statistical Methods for the Analysis of Biomedical Data, Second Edition
WU and HAMADA · Experiments: Planning, Analysis and Parameter Design Optimization
WU and ZHANG · Nonparametric Regression Methods for Longitudinal Data Analysis: Mixed-Effects Modeling Approaches
YANG · The Construction Theory of Denumerable Markov Processes
YOUNG, VALERO-MORA and FRIENDLY · Visual Statistics: Seeing Data with Dynamic Interactive Graphics
ZACKS · Stage-Wise Adaptive Designs
*ZELLNER · An Introduction to Bayesian Inference in Econometrics
ZELTERMAN · Discrete Distributions: Applications in the Health Sciences
ZHOU, OBUCHOWSKI and McCLISH · Statistical Methods in Diagnostic Medicine

*Now available in a lower priced paperback edition in the Wiley Classics Library.