CRRAO Advanced Institute of Mathematics, Statistics and Computer Science (AIMSCS)

Research Report

Author (s): **M.N.Mishra & B.L.S. Prakasa Rao**

Title of the Report: **Local asymptotic normality and estimation via Kalman-Bucy lter for a linear system with signal driven by a fractional Brownian motion and observation driven by a Brownian motion**

Research Report No.: **RR2016-07**

Date: **September 26, 2016**

Prof. C R Rao Road, University of Hyderabad Campus,
Gachibowli, Hyderabad-500046, INDIA.
www.crraoaimscs.org

Local asymptotic normality and estimation via
Kalman-Bucy filter for a linear system
with signal driven by a fractional Brownian motion
and observation driven by a Brownian motion

M.N. Mishra
Institute of Mathematics and its Applications, Bhubaneswar, India
e-mail: mnmeeshraw@yahoo.com
B.L.S. Prakasa Rao
CR Rao Advanced Institute of Mathematics, Statistics
and Computer Science, Hyderabad, India
e-mail: blsprao@gmail.com

Abstract

We study the local asymptotic normality and estimation for drift parameter obtained through Kalman-Bucy filter for a linear system when the signal is driven by a fractional Brownian motion and the observation is driven by a Brownian motion.

Keywords: Brownian motion; Estimation of drift parameter; Fractional Brownian motion; Innovation process; Kalman-Bucy filter; Linear systems; Optimal filtering.

2010 MS Subject Classification: Primary 62M20, 60G22, 60G35; Secondary 93E11.

1 Introduction

Suppose $X = \{X_t, 0 \leq t \leq T\}$ and $Y = \{Y_t, 0 \leq t \leq T\}$ are real-valued stochastic proceses, representing the signal and the observation respectively, governed by the following homogeneous linear system of stochastic differential equations

(1. 1)
$$\begin{aligned} dX_t &= \theta X_t dt + \epsilon\, dV_t^h, 0 \leq t \leq T, X_0 = x_0 \neq 0, \\ dY_t &= \theta X_t dt + \epsilon\, dW_t, 0 \leq t \leq T, Y_0 = 0. \end{aligned}$$

Here the processes $V^h = \{V_t^h, 0 \leq t \leq T\}$ is a fractional Brownian motion with Hurst index $h \in [\frac{1}{2}, 1)$ and $W = \{W_t, 0 \leq t \leq T\}$ is a standard Brownian motion independent of V^h and

$\theta \in \Theta$ open in R. Suppose the component $Y = \{Y_t, 0 \leq t \leq T\}$ is observed and the problem is to estimate the unknown parameter θ based on the observation $Y = \{Y_t, 0 \leq t \leq T\}$ and study its asymptotic properties as $\epsilon \to 0$. The system (1.1) has a unique solution (X, Y) which is a Gaussian process . Suppose that we observe the process Y alone but would like to have information about the process X at time t. This problem is known as *filtering* the signal X at time t from the observation of Y up to time t. The solution to this problem is the conditional expectation of X_t given the σ-algebra generated by the process $\{Y_s, 0 \leq s \leq t\}$. Since the processes (X, Y) is jointly Gaussian, the conditional expectation of X_t given $\{Y_s, 0 \leq s \leq t\}$ is linear in $\{Y_s, 0 \leq s \leq t\}$. It is also the *optimal filter* in the sense of minimizing the mean square error. The problem of finding the optimal filter reduces to finding the conditional mean $\pi_t(\theta, X) = E_\theta(X_t|Y_s, 0 \leq s \leq t)$. This problem leads to Kalman-Bucy filter if $h = \frac{1}{2}$. Le Breton (1998) and Kleptsyna and Le Breton (2002b) and Kleptsyna et al. (2000a,b) studied this problem of filtering for $h \in (\frac{1}{2}, 1)$. For $h = 1/2$, this problem has been solved by Kutoyants (1994). For optimal filtering for fractional stochastic systems, see Kleptsyna, Kloden and Ahn (1998). Asymptotic properties of maximum likelihood estimator of the drift parameter for partially observed fractional diffusion systems are investigated in Brouste and Kleptsyna (2010). Kallianpur and Selukar (1991,1993) have studied parameter estimation and local asymptotic normality in linear filtering for linear systems driven by Brownian motions. They have also obtained a large deviation inequality for the maximum likelihood estimator (MLE) of the parameter. Mishra and Prakasa Rao (2016) investigated local asymptotic normality and estimation via Kalman-Bucy filter for linear systems when both the signal and the observation are driven by independent fractional Brownian motions with the same Hurst index subject to a technical condition. Our results, in the special case discussed here, do not need any extra technical condition.

We obtain the asymptotic properties of the maximum likelihood estimator (MLE) of the parameter θ by studying the asymptotic properties of the log-likelihood ratio process with index ϵ as $\epsilon \to 0$. We follow the techniques used by Prakasa Rao (1968), Ibragimov and Khasminskii (1981) and others. We prove the weak convergence of the appropriately normalized log-likelihood ratio random process and appeal to the continuous mapping theorem to study the asymptotic behaviour of the MLE of the parameter θ as $\epsilon \to 0$.

We now state the main result of this paper. Let θ denote the true parameter. Let $\hat{\theta}_\epsilon$ denote the maximum likelihood estimator of θ based on the observation of the process Y over the interval $[0, T]$ satisfying the stochastic differential system defined by (1.1). Then, as

$\epsilon \to 0$, the random variable

$$\epsilon^{-1}(\hat{\theta}_\epsilon - \theta)$$

converges in distribution to the Gaussian distribution with mean zero and variance $[\sigma^2(\theta)]^{-1}$ where $\sigma^2(\theta)$ will be specified later.

2 Preliminaries

We now introduce some notation and some basic results. 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 a standard fractional Brownian motion 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

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

Let us consider a stochastic process $J = \{J_t, t \geq 0\}$ governed by the stochastic integral equation

$$J_t = \int_0^t C(s)ds + \int_0^t B(s)dW_s^h, t \geq 0 \tag{2. 2}$$

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

$$dJ_t = C(t)dt + B(t)dW_t^h, t \geq 0; J_0 = 0 \tag{2. 3}$$

driven by the fractional Brownian motion W^h. Even though the process J 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 $(\mathcal{J}_t)$ of the process J (Kleptsyna et al. (2000a)). Define, for $0 < s < t$,

$$k_h = 2h\,\Gamma(\frac{3}{2} - h)\Gamma(h + \frac{1}{2}), \tag{2. 4}$$

$$\kappa_h(t,s) = k_h^{-1} s^{\frac{1}{2}-h}(t-s)^{\frac{1}{2}-h}, \tag{2. 5}$$

$$\lambda_h = \frac{2h\,\Gamma(3-2h)\Gamma(h+\frac{1}{2})}{\Gamma(\frac{3}{2}-h)}, \tag{2. 6}$$

$$w_t^h = \lambda_h^{-1} t^{2-2h}, \tag{2. 7}$$

and

$$M_t^h = \int_0^t \kappa_h(t,s)dW_s^h, t \geq 0. \tag{2. 8}$$

The process M^h is a Gaussian martingale, called the *fundamental martingale* and its quadratic variation $< M_t^h >= w_t^h$. Further more the natural filtration of the martingale M^h coincides with the natural filtration of the fBm W^h.

Suppose the sample paths of the process $\{\frac{C(t)}{B(t)}, t \geq 0\}$ are smooth so that

$$Q_h(t) = \frac{d}{dw_t^h}\int_0^t \kappa_h(t,s)\frac{C(s)}{B(s)}ds, t \in [0,T] \tag{2. 9}$$

is well-defined where the functions w^h and $k_h(t,s)$ are as defined in (2.7) and (2.5) 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 J such that the natural filtration $(\mathcal{Z}_t)$ of Z coincides with the natural filtration $(\mathcal{J}_t)$ of J.

Theorem 2.1: *Suppose the sample paths of the process* Q_h *belong to* $L^2([0,T], dw^h)$ *a.s. Let the process* $Z = (Z_t, t \in [0,T])$ *be defined by*

$$Z_t = \int_0^t \kappa_h(t,s)B^{-1}(s)dJ_s. \tag{2. 10}$$

Then the following results hold:

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

$$Z_t = \int_0^t Q_h(s)dw_s^h + M_t^h \tag{2. 11}$$

where M^h *is the fundamental martingale defined above.*

(ii) the natural filtrations $(\mathcal{Z}_t)$ *and* $(\mathcal{J}_t)$ *coincide.*

For more details on properties of fractional diffusion processes, see Prakasa Rao (2010).

Suppose that $\{\eta_t, 0 \leq t \leq T\}$ is a random process adapted to the filtration $(\mathcal{F}_t)$ such that $E_\theta|\eta_t| < \infty$ on the underlying probability space $(\Omega, \mathcal{F}, P)$. Let $\pi_t(\theta, \eta)$ denote the conditional expectation of η_t given $\{Y_s, 0 \leq s \leq t\}$ when θ is the true parameter. Let $(\mathcal{Y}_t)$ denote the filtration generated by the process Y. Let

$$\epsilon\nu_t = Y_t - \theta\int_0^t \pi_s(\theta, X)ds, 0 \leq t \leq T \tag{2. 12}$$

where $\pi_t(\theta, X) = E_\theta(X(t)|Y_s, 0 \le s \le t)$. The process $\nu = \{\nu_t, 0 \le t \le T\}$ is the *innovation type process* and is a Wiener process in the present problem. Furthermore, if $N = \{N_t, 0 \le t \le T\}$ is a square integrable $(\mathcal{Y}_t)$-martingale, $N_0 = 0$, then there exists a $(\mathcal{Y}_t)$-adapted process $\alpha = \{\alpha_t, 0 \le t \le T\}$ such that

$$E(\int_0^T \alpha_s^2 ds) < \infty$$

and P-a.s

$$N_t = \int_0^t \alpha_s d\nu_s, 0 \le t \le T.$$

3 Main Results

Consider the linear system described by (1.1). Let $x_t(\theta) = x_0 e^{\theta t}, 0 \le t \le T$ denote the solution of the corresponding ordinary differential equation

$$\frac{dx_t}{dt} = \theta x_t$$

with $x_t(\theta) = x_0$ at $t = 0$. Suppose $\theta \in \Theta$ open in R. It can be checked that the processes Q and $\pi(\theta, Q)$, as defined in Section 2, are given by the relations

$$Q_t = \frac{\theta}{\epsilon} X(t), \pi_t(\theta, Q) = \frac{\theta}{\epsilon} \pi_t(\theta, X)$$

for this problem (cf. Kleptsyna et al. (2000b), p. 129). Let

$$p(t, s) = \frac{\theta}{\epsilon} \frac{d}{dw_t^h} \int_s^t \kappa_h(t, r) dr.$$

An application of Theorem 4 of Kleptsyna et al. (2000b) to the process X shows that

$$\pi_t(\theta, X) = x_0 + \theta \int_0^t \pi_s(\theta, X) ds + \int_0^t c(\theta, \epsilon, t, s) d\nu_s, 0 \le t \le T$$

for some non-random function $c(\theta, \epsilon, t, s)$. Following equation (6.10) in Kutoyants (1994), p.194, the representation given above can also be expressed as a stochastic differential equation

$$d\pi_t(\theta, X) = \theta \pi_t(\theta, X) dt + \theta \epsilon z_t(\theta) d\nu_t, \pi_0(\theta, X) = x_0$$

where $z_t(\theta) = \gamma_t(\theta) \epsilon^{-2}$ and $\gamma_t(\theta) = E_\theta[\pi_t(\theta, X) - X_t]^2$. Let

$$J_t(\theta) = \theta - z_t(\theta) \theta^2$$

and

$$M_t(\theta) = \theta z_t(\theta).$$

Following again the derivations in obtaining the equation (6.11) in Kutoyants (1994), p.194, we obtain that

$$d\bar{\pi}_t(\theta, X) = [\bar{J}_t(\theta)\pi_t(\theta, X) + J_t(\theta)\bar{\pi}_t(\theta, X)]dt + \bar{M}_t(\theta)d\nu_t, \bar{\pi}_0(\theta, X) = 0$$

where $\bar{V}(\theta, X)$ denotes the derivative of $V(\theta, X)$ with respect to θ in L_2-mean. Let

$$\zeta_t(\theta, X) = \pi_t(\theta, X) + \theta\bar{\pi}_t(\theta, X) \tag{3. 1}$$

and

$$\zeta_t(\theta, x) = x_0 e^{\theta t} + \theta^2 x_0 e^{\theta t}. \tag{3. 2}$$

Note that

$$E_\theta[\zeta_t(\theta, X) - \zeta_t(\theta, x)]^2 \leq 2(E_\theta[(\pi_t(\theta, X) - x_0 e^{\theta t})^2] + E_\theta[(\theta\bar{\pi}_t(\theta, X) - x_0\theta^2 e^{\theta t})^2]). \tag{3. 3}$$

Following Lemma 4.0 proved below, this inequality, in turn, implies that the random variable $\zeta_t(\theta, X)$ converges in L_2-mean to $\zeta_t(\theta, x)$ as $\epsilon \to 0$ when θ is the true parameter.

Fix $\theta \in \Theta \in R$. Let

$$\Delta_t = (\theta + \epsilon u)\pi_t(\theta + \epsilon u, X) - \theta\ \pi_t(\theta, X).$$

For convenience, we denote $\theta + \epsilon u_1 = \beta_1$ and $\theta + \epsilon u_2 = \beta_2$.

Let

$$\sigma^2(\theta) = \int_0^T [\zeta_t(\theta, x)]^2 dt \tag{3. 4}$$

and

$$L_0(u) = u\xi - \frac{1}{2}u^2\sigma^2(\theta), u \in R \tag{3. 5}$$

where ξ is a Gaussian random variable with mean zero and variance $\sigma^2(\theta)$ and the function $\zeta_t(\theta, x)$ is as specified in (3.2).

We now state the main result of this paper.

Theorem 3.1: Let θ denote the true parameter. Let $\hat{\theta}_\epsilon$ denote the maximum likelihood estimator of θ based on the observation of the process Y over the interval $[0, T]$ satisfying the stochastic differential system defined by (1.1). Then, as $\epsilon \to 0$, the normalized random vector

$$\epsilon^{-1}(\hat{\theta}_\epsilon - \theta)$$

converges to the Gaussian distribution with mean zero and variance $[\sigma^2(\theta)]^{-1}$.

Local asymptotic normality: Let P_θ be the probability measure generated by the process Y on the space $C[-g, g]$ associated with the uniform topology when θ is the true parameter. Here $C[-g, g]$ is the space of continuous real-valued functions on the interval $[-g, g]$ where $g > 0$. Equation (26) of Kleptsyna et al. (2000b) implies that, for any θ_1 and θ_2 in Θ,

$$\frac{dP_{\theta_2}}{dP_{\theta_1}} = \exp\{\frac{1}{\epsilon}\int_0^T [\theta_2\pi_s(\theta_2, X) - \theta_1\pi_s(\theta_1, X)]d\nu_s - \frac{1}{2\epsilon^2}\int_0^T [\theta_2\pi_s(\theta_2, X) - \theta_1\pi_s(\theta_1, X)]^2 ds\}.$$

Consider the log-likelihood ratio process

$$L_\epsilon(u) = \log \frac{dP_{\theta+\epsilon u}}{dP_\theta}$$

for fixed u such that $\theta, \theta + \epsilon u \in \Theta$.

Let K denote a compact subset of Θ such that $\theta \in K$ and $\theta + \epsilon u \in K$. Let C_K denote the space of continuous functions defined on the compact set K. Let $K_\theta = \{u : \theta \in K \text{ and } \theta + \epsilon u \in K\}$.

Theorem 3.2: *The family of probability measures, generated by the log-likelihood ratio random process $\{L_\epsilon(u), u \in K_\theta\}$ on C_{K_θ} associated with the uniform norm topology is locally asymptotically normal and converge weakly to the probability measure generated by the random process $\{L_0(u), u \in K_\theta\}$ on C_{K_θ} as $\epsilon \to 0$.*

From the general theory of weak convergence of probability measures on the space C_{K_θ} associated with the uniform norm topology (cf. Billingsley (1968), Parthasarathy (1967), Prakasa Rao (1975)), in order to prove Theorem 3.2, it is sufficient to prove that the finite dimensional distributions of the random field $\{L_\epsilon(u), u \in K_\theta\}$ converge to the corresponding finite dimensional distributions of the random field $\{L_0(u), u \in K_\theta\}$ and the family of probability measures generated by the random fields $\{L_\epsilon(u), u \in K_\theta\}$ for different ϵ is tight.

4 Proofs of Theorems 3.1 and 3.2

Before we give proofs of Theorem 3.1 and Theorem 3.2, we prove some related lemmas.

Lemma 4.0: *Let $\theta \in \Theta$. There exists a neighbourhood $N_\theta = \{\theta' : |\theta' - \theta| < \epsilon u\}$ of θ contained in Θ and a constant $c_t > 0$ depending on θ such that*

$$(i) \sup_{\theta' \in N_\theta} \sup_{0 \leq s \leq t} E_\theta |\pi_s(\theta', X) - x_s(\theta')|^2 \leq c_t \epsilon^2 t$$

and

$$(ii) \sup_{\theta' \in N_\theta} \sup_{0 \leq s \leq t} E_\theta |\pi_s(\theta', Q) - \epsilon^{-1}\theta' x_s(\theta')|^2 \leq c_t t.$$

Proof : An application of the Grownwall's inequality (cf. Kutoyants (1994), Lemma 1.13) shows that

$$\sup_{\theta' \in N_\theta} \sup_{0 \leq s \leq t} |\pi_s(\theta', X) - x_s(\theta')| \leq c_0 \epsilon \sup_{0 \leq s \leq t} |\nu_s|$$

and hence

$$\sup_{\theta' \in N_\theta} \sup_{0 \leq s \leq t} E_\theta [|\pi_s(\theta', X) - x_s(\theta')|^2] \leq c_t \epsilon^2 t. \tag{4. 1}$$

The second inequality follows from the first inequality following the representation for the process $\pi_t(\theta, Q)$ given above.

Following the arguments in Kutoyants (1994), p.194-195, it follows that the process $\{\theta \pi_t(\theta, X), 0 \leq t \leq T\}$ is L_2-differentiable with respect to θ and

$$E_\theta[||(\theta + \epsilon u)\pi(\theta + \epsilon u, X) - \theta \pi(\theta, X) - \epsilon u \, \zeta(\theta, X)||^2] \leq C \epsilon^4 u^4. \tag{4. 2}$$

Lemma 4.1: *The finite dimensional distributions of the random process $\{L_\epsilon(u), u \in K_\theta\}$ converge to the corresponding finite dimensional distributions of the random process $\{L_0(u), u \in K_\theta\}$ as $\epsilon \to 0$.*

Proof: We will first investigate the convergence of the one-dimensional marginal distributions of the random process $L_\epsilon(u)$ as $\epsilon \to 0$. The convergence of other classes of finite-dimensional distributions follows from the Cramer-Wold device. From the equation (26) in Kleptsyna et al. (2000b), it follows that

$$\begin{aligned} L_\epsilon(u) &= \frac{1}{\epsilon} \int_0^T \Delta_t d\nu_t - \frac{1}{2\epsilon^2} \int_0^T \Delta_t^2 dt \\ &= \frac{1}{\epsilon} \int_0^T (\Delta_t - \epsilon u \, \zeta_t(\theta, x)) d\nu_t + \frac{1}{\epsilon} \int_0^T \epsilon u \, \zeta_t(\theta, x) d\nu_t \end{aligned}$$

$$
\begin{aligned}
& -\frac{1}{2\epsilon^2}\int_0^T \Delta_t^2 dt \\
= \; & I_1 + I_2 + I_3 \;\; \text{(say)}.
\end{aligned}
$$

Note that the process $\{\nu(t), 0 \le t \le T\}$ is the innovation process which is a Wiener process. Observe that

$$
E(I_1^2) = \frac{1}{\epsilon^2}\int_0^T E_\theta[\Delta_t - \epsilon u\, \zeta_t(\theta, x)]^2 dt = o(1) \tag{4. 3}
$$

as $\epsilon \to 0$ by equations (3.1) to (3.3),(4.1) and (4.2) and hence $I_1 = o_p(1)$. Note that

$$
I_2 = \int_0^T u\zeta_t(\theta, x) d\nu_t
$$

is a Gaussian random variable with mean zero and variance $\int_0^T u^2[\zeta_t(\theta, x)]^2\, dt$. Furthermore

$$
\begin{aligned}
I_3 &= -\frac{1}{2\epsilon^2}\int_0^T \Delta_t^2 dt \\
&= -\frac{1}{2\epsilon^2}\int_0^T (\Delta_t - \epsilon u\, \zeta_t(\theta, x) + \epsilon u\, \zeta_t(\theta, x))^2 dt \\
&= -\frac{1}{2\epsilon^2}\int_0^T [(\Delta_t - \epsilon u\, \zeta_t(\theta, x))^2 + (\epsilon u\, \zeta_t(\theta, x))^2 + 2(\Delta_t - \epsilon u\, \zeta_t(\theta, x))\epsilon u\, \zeta_t(\theta, x)] dt \\
&= -\frac{1}{2\epsilon^2}\int_0^T (\epsilon u\, \zeta_t(\theta, x))^2 dt + o_p(1).
\end{aligned}
$$

As a consequence of the above computations, we observe that, as $\epsilon \to 0$,

$$
\begin{aligned}
\frac{1}{\epsilon^2}\int_0^T \Delta_t^2 dt &= \frac{1}{\epsilon^2}\int_0^T [(\theta + \epsilon u)\pi_t(\theta + \epsilon u, X) - \theta\pi_t(\theta, X)]^2 dt \\
&= u^2 \int_0^T [\zeta_t(\theta, x)]^2 dt + o_p(1)
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{1}{\epsilon}\int_0^T \Delta_t d\nu_t &= \frac{1}{\epsilon}\int_0^T [(\theta + \epsilon u)\pi_t(\theta + \epsilon u, X) - \theta\pi_t(\theta, X)] d\nu_t \\
&= u\int_0^T \zeta_t(\theta, x) d\nu_t + o_p(1) \\
&= u\psi + o_p(1)
\end{aligned}
$$

as $\epsilon \to 0$ where ψ is a Gaussian random variable with mean zero and variance $\sigma^2(\theta)$. Hence the random variable $L_\epsilon(u)$ is asymptotically Gaussian with mean $-(1/2)\sigma^2(\theta)u^2$ and variance $\sigma^2(\theta)u^2$.

We have proved the convergence of the univariate distributions of the random process $\{L_\epsilon(u), u \in K_\theta\}$ as $\epsilon \to 0$, after proper scaling. Convergence of all the other finite dimensional distributions of the random field $\{L_\epsilon(u), u \in K_\theta\}$, after proper scaling, as $\epsilon \to 0$, follows by an application of the Cramer-Wold device. In order to prove that a sequence of k-dimensional random vectors $\mathbf{X}_n$ converge in law to a k-dimensional random vector $\mathbf{X}$ as $n \to \infty$, it is sufficient to prove that the sequence of random variables $\lambda'\mathbf{X}_n$ converges in law to the random variable $\lambda'\mathbf{X}$ for all $\lambda \in R^k$. This is known as the Cramer-Wold technique for converting the problem of the finite dimensional convergence to convergence of one-dimensional random variables. Similar ideas have been applied earlier in proving the weak convergence of the processes. See, for instance, Fokianos and Newmann (2013)). We can use this technique to prove the convergence of the finite-dimensional distributions to complete the proof of the lemma.

We now state two lemmas which will be used in the following computations. For proofs of these lemmas, see Lemmas 5.2 and 5.3 in Mishra and Prakasa Rao (2014).

Lemma 4.2: *Let* $\{D_t, 0 \leq t \leq T\}$ *be a random process such that* $\sup_{0\leq t\leq T} E(D_t^4) \leq \gamma < \infty$. *Then, for* $0 \leq \theta_2 \leq \theta_1 \leq T$,

$$E([\int_{\theta_2}^{\theta_1} D_t dt]^4) \leq |\theta_1 - \theta_2|^3 \int_{\theta_2}^{\theta_1} E[D_t^4] dt \leq \gamma |\theta_1 - \theta_2|^4.$$

The next lemma gives an inequality for the 4-th moment of a stochastic integral with respect to a martingale.

Lemma 4.3: *Let the process* $\{f_t, 0 \leq t \leq T\}$ *be a random process adapted to a square integrable martingale* $\{M_t, \mathcal{F}_t, t \geq 0\}$ *with the quadratic variation* $< M >_t$ *such that*

$$\int_0^T E(f_s^4) d < M >_s < \infty.$$

Then

$$E((\int_0^T f_t dM_t)^4) \leq 36 < M >_T \int_0^T E(f_t^4) d < M >_t .$$

and, in general, for $0 \leq \theta_2 \leq \theta_1 \leq T$,

$$E[(\int_{\theta_2}^{\theta_1} f_t dM_t)^4] \leq 36(< M >_{\theta_1} - < M >_{\theta_2}) \int_{\theta_2}^{\theta_1} E[f_t^4] d < M >_t .$$

Lemma 4.4: *Let $\Gamma_\epsilon(u) = \exp\{L_\epsilon(u)\}$. Then, for any $R > 0$, there exist a constant $C > 0$ such that*

$$E_\theta \left|\Gamma_\epsilon^{\frac{1}{4}}(u_2) - \Gamma_\epsilon^{\frac{1}{4}}(u_1)\right|^4 \leq C(u_1 - u_2)^4, |u_i| \leq R, i = 1, 2.$$

Proof : Let $-R \leq u_1, u_2 \leq R$ for some $R > 0$. Let

$$\delta_t = (\theta + \epsilon u_1)\pi_t(\theta + \epsilon u_1, X) - (\theta + \epsilon u_2)\pi_t(\theta + \epsilon u_2, X)$$

and

$$\bar{\delta}_t = \epsilon(u_1 - u_2)\bar{\zeta}_t(\theta, x).$$

Recall the notation $\theta + \epsilon u_1 = \beta_1, \theta + \epsilon u_2 = \beta_2$ used earlier. Let

$$R_t = \exp[\frac{1}{4\epsilon}\int_0^t \delta_s d\nu_s - \frac{1}{8\epsilon^2}\int_0^t \delta_s^2 ds], R_0 = 1.$$

Note that the process R_t is the process $\left(\frac{dP_{\beta_1}}{dP_{\beta_2}}(X)\right)^{\frac{1}{4}}$ and, by the Ito formula, we have

$$dR_t = -\frac{3}{(32)\epsilon^2}\delta_t^2 R_t dt + \frac{1}{4\epsilon}\delta_t R_t d\nu_t.$$

Hence

$$R_t = 1 - \frac{3}{(32)\ \epsilon^2}\int_0^t \delta_s^2 R_s ds + \frac{1}{4\ \epsilon}\int_0^t \delta_s R_s d\nu_s, 0 \leq s, t \leq T$$

Note that

$$\begin{aligned} & E_\theta \left|\Gamma_\epsilon^{\frac{1}{4}}(u_2) - \Gamma_\epsilon^{\frac{1}{4}}(u_1)\right|^4 \\ = \ & E_\theta(\frac{dP_{\beta_2}}{dP_\theta}|1 - R_T|^4) = E_{\beta_2}(|1 - R_T|^4) \\ \leq \ & C\frac{1}{\epsilon^8}E_{\beta_2}\left|\int_0^T \delta_t^2 R_t dt\right|^4 + C\frac{1}{\epsilon^4}E_{\beta_2}\left|\int_0^T \delta_t R_t d\nu_t\right|^4 \end{aligned}$$

where C is an absolute constant. In order to get the bounds for the expectations of the integrals in the above inequality, we now use the Lemmas 4.2 and 4.3.

Let us now estimate the term

$$E_{\beta_2}\left|\int_0^T \delta_t^2 R_t dt\right|^4.$$

Note that

$$\begin{aligned} I_1 &\equiv E_{\beta_2}\left|\int_0^T \delta_t^2 R_t dt\right|^4 \\ &\leq cT^3\int_0^T E_{\beta_2}|\delta_t^2 R_t|^4 dt \\ &\leq cT^3\int_0^T E_{\beta_1}|\delta_t^2|^4 dt \\ &\leq cT^{8-8H}\sup_{\theta,0\leq t\leq T} E_\theta[\delta_t^8] \\ &\leq c\epsilon^8(u_2-u_1)^8. \end{aligned}$$

Let us now estimate the term

$$I_2 \equiv E_{\beta_2}|\int_0^T \delta_t^2 R_t d\nu_t|^4.$$

Observe that

$$\begin{aligned} I_2 &\leq cw_t^H\int_0^T E_{\beta_2}|\delta_t R_t|^4 dt \\ &\leq cw_t^H\int_0^T E_{\beta_2}|\delta_t R_t|^4 dt \\ &\leq cT^{2-2H}\int_0^T E_{\beta_1}|\delta_t|^4 dt \\ &\leq c(u_1-u_2)^4\epsilon^4. \end{aligned}$$

Combining the above estimates, we obtain that

$$\sup_{|u_i|\leq R,|v_i|\leq R}(u_1-u_2)^{-4}E_\theta|\Gamma_\epsilon^{1/4}(u_2)-\Gamma_\epsilon^{1/4}(u_1)|^4 < c < \infty$$

which proves the tightness from the results in Prakasa Rao (1975) or Neuhaus (1971).

As a consequence of Lemma 4.4, it follows that the family of probability measures generated by the processes $\{\Gamma_\epsilon^{\frac{1}{4}}(u), u \in K_\theta\}$ on C_{K_θ} with uniform topology is tight from the results in Billingsley (1968) (cf. Prakasa Rao (1987)) and hence the family of probability measures generated by the processes $\{L_\epsilon(u), u \in K_\theta\}$ on C_{K_θ} is tight.

Lemmas 4.1 and 4.4 together imply that that the family of probability measures generated by the processes $\{L_\epsilon(u, u \in K_\theta\}$ on C_{K_θ} converge weakly to the probability measure generated by the processes $\{L_0(u), u \in K_\theta\}$ on C_{K_θ} from the general theory of weak convergence of probability measures on complete separable metric spaces(cf. Billingsley (1968), Parthasarathy (1967), Prakasa Rao (1987) and Ibragimov and Khasminskii (1981)). This completes the proof of Theorem 3.2.

The following maximal inequality is proved in Lemma 5.6 in Mishra and Prakasa Rao (2014) using the Slepian's lemma (cf. Leadbetter et al. (1983) and Matsui and Shieh (2009)). We will use it in the sequel.

Lemma 4.5: *Let W^H be a fractional Brownian motion with Hurst index H. For any $\lambda > 0$,*

$$E[\exp\{\lambda \max_{0\leq t\leq T} |W_t^H|\}] \leq 1 + \lambda\sqrt{2\pi T^{2H}} \exp\{\frac{\lambda^2 T^{2H}}{2}\}.$$

We now apply Lemma 4.5 to get the following result.

Lemma 4.6: *Let $\Gamma_\epsilon(u) = \exp\{L_\epsilon(u)\}, u \in R$. Then, for any compact set $K \subset \Theta$, and for any $0 < p < 1$, there exists a positive constant C such that*

$$\sup_{\theta\in K} E_\theta[(\Gamma_\epsilon(u))^p] \leq e^{-C\,u^2} \tag{4. 4}$$

for all $u \in R$.

Proof: Now, for any $0 < p < 1$, we will now estimate $E_{\theta,\tau}(\Gamma_\epsilon(u))^p$. For convenience, let $u \in R$ and $v > 0$ and let

$$F_1 \equiv \int_0^T \Delta_t d\nu_t$$

and

$$F_2 \equiv \int_0^T \Delta_t^2 dt.$$

Let q be such that $p^2 < q < p$. Then

$$\begin{aligned} E_\theta[(\Gamma_\epsilon(u))^p] &= E_\tau[\exp\{\frac{p}{\epsilon}F_1 - \frac{p}{2\epsilon^2}F_2\}] \\ &= E_\tau[\exp\{\frac{p}{\epsilon}F_1 - \frac{q}{2\epsilon^2}F_2 - \frac{(p-q)}{2\epsilon^2}F_2\}]. \end{aligned}$$

Let

$$G_1 = \exp\{-\frac{(p-q)}{2\epsilon^2}F_2\}$$

and

$$G_2 = \exp\{\frac{p}{\epsilon}F_1 - \frac{q}{2\epsilon^2}F_2\}.$$

Then

$$\begin{aligned} E_\theta[(\Gamma_\epsilon(u))^p] &= E_\theta[G_1 G_2] \\ &\leq (E_\theta[G_1^{p_1}])^{1/p_1}(E_\theta[G_2^{p_2}])^{1/p_2} \end{aligned}$$

by the Holder inequality for any p_1 and p_2 such that $p_2 > 1$ and $\frac{1}{p_1} + \frac{1}{p_2} = 1$. Choose $p_2 = \frac{q}{p^2} > 1$. Then $p_1 = \frac{q}{q-p^2}$. Observe that

$$\begin{aligned} E_\theta[G_2^{p_2}] &= E_\theta[\exp\{p_2(\frac{p}{\epsilon}F_1 - \frac{q}{2\epsilon^2}F_2)\}] \\ &= E_\theta[\exp\{\frac{q}{p^2}(\frac{p}{\epsilon}F_1 - \frac{q}{2\epsilon^2}F_2)\}] \\ &= E_\theta[\exp\{\frac{1}{\epsilon}\frac{q}{p}F_1 - \frac{1}{2\epsilon^2}\frac{q^2}{p^2}F_2\}]. \end{aligned}$$

The random variable, under the expectation sign in the last line, is the Radon-Nikodym derivative of two probability measures which are absolutely continuous with respect to each other by the Girsanov's theorem for martingales. Hence the expectation is equal to one. Hence

$$\begin{aligned} E_\theta[(\Gamma_\epsilon(u))^p] &\leq (E[\exp\{-\frac{p_1(p-q)}{2\epsilon^2}F_2\}])^{1/p_1} \\ &= (E[\exp\{-\gamma\epsilon^{-2}F_2\}])^{1/p_1}. \end{aligned}$$

where $\gamma = \frac{q(p-q)}{2(q-p^2)} > 0$. Let us now estimate $E_\theta[e^{-\gamma\epsilon^{-2}F_2}]$. Let

$$\bar{\Delta}_t = (\theta + \epsilon u)x_t(\theta + \epsilon u) - \theta x_t(\theta), 0 \leq t \leq T.$$

Applying the inequality

$$a^2 \geq b^2 - 2|b(a-b)|,$$

it follows that

$$E_\theta[e^{-\gamma\epsilon^{-2}F_2}] \leq \exp\{-\gamma\epsilon^{-2}\int_0^T \bar{\Delta}_t^2 dt\} \times$$

$$\times E_\theta[\exp\{2\gamma\epsilon^{-2}(\int_0^T (|(\pi_t(\theta+\epsilon u,Q)-\epsilon^{-1}(\theta+\epsilon u)x_t(\theta+\epsilon u)|+$$
$$+|(\pi_t(\theta,Q)-\epsilon^{-1}\theta x_t(\theta)|)\epsilon^{-1}|(\theta+\epsilon u)x_t(\theta+\epsilon u)-\theta x_t(\theta)|dt\}].$$

We now get an upper bound on the term under the expectation sign on the right side of the above inequality. Observe that there exists a a constant $c > 0$, such that,

$$\begin{aligned}
&\int_0^T [\pi_t(\theta,Q)-\epsilon^{-1}\theta x_t(\theta)]^2\,dt \\
&\leq c\epsilon^2[\int_0^T dt]\sup_{0\leq t\leq T}|\nu_t|^2 \\
&\leq c\epsilon^2 T\sup_{0\leq t\leq T}|\nu_t|^2
\end{aligned}$$

for some constant $c > 0$ possibly depending on T and Θ where $\{\nu_t, 0 \leq t \leq T\}$ is the innovation Wiener process. An application of the Cauchy-Schwartz inequality implies that

$$\sup_{\theta,\theta',\theta+\epsilon u\in\Theta,0<\epsilon<\epsilon_0}[\int_0^T \epsilon^{-1}|(\theta+\epsilon u)x_t(\theta+\epsilon u)-\theta x_t(\theta)||\pi_t(\theta',Q)-\epsilon^{-1}\theta' x_t(\theta')|dt]^2$$
$$\leq C_0\epsilon^4 u^2 T\sup_{0\leq t\leq T}|\nu_t|^2$$

for some constant $C_0 > 0$. Hence

$$\sup_{\theta,\theta'=\theta+\epsilon u\in\Theta,0<\epsilon<\epsilon_0}[\int_0^T |\epsilon^{-1}(\theta+\epsilon u)x_t(\theta+\epsilon u)-\epsilon^{-1}\theta x_t(\theta)||\pi_t(\theta',Q)-\epsilon^{-1}\theta' x_t(\theta')|dt]$$
$$\leq C_1\epsilon^2|u|\sup_{0\leq t\leq T}|\nu_t|.$$

for some constant $C_1 > 0$. Therefore

$$\begin{aligned}
&\sup_{\theta,\theta+\epsilon u\in\Theta,0<\epsilon<\epsilon_0} E_\theta[\exp\{2\gamma\epsilon^{-2}(\int_0^T |\pi_t(\theta+\epsilon u,Q)-\epsilon^{-1}(\theta+\epsilon u)x_t(\theta+\epsilon u)| \\
&\quad +|(\pi_t(\theta,Q)-\epsilon^{-1}\theta x_t(\theta))\epsilon^{-1}((\theta+\epsilon u)x_t(\theta+\epsilon u)-\theta x_t(\theta))|dt\}] \\
&\leq E_\theta[\exp\{C\gamma|u|\sup_{0\leq t\leq T}|\nu_t|\}] \\
&\leq 1+\gamma C|u|\sqrt{2\pi T}\exp\{\frac{c\gamma^2 T u^2}{2}\}
\end{aligned}$$

for some positive constants C and c depending on T and the set Θ by Lemma 6.5. Applying arguments similar to those in Lemma 2.4 in Kutoyants (1994), we get that

$$\sup_{\theta\in K, 0<\epsilon<\epsilon_0} E_\theta[\Gamma_\epsilon^p(u)] \leq e^{-C\,u^2}$$

for some positive constant $C > 0$ depending on T and Θ.

An application of Lemma 4.6, proved above, shows that the maximum likelihood estimator $\hat{\theta}_\epsilon$ will lie in a compact set K with probability tending to one as $\epsilon \to 0$ from Theorem 5.1 in Chapter 1, p.42 of Ibragimov and Khasminskii (1981).

We now give a proof of Theorem 3.1 stated above.

Proof of Theorem 3.1: Let C_K denote the family of continuous functions defined on a compact set K in R. In view of Theorem 3.2, it follows that the family of probability measures generated by the random processes $\{L_\epsilon(u), u \in K\}, \epsilon > 0$ on C_K converge weakly to the probability measure generated by the random process $\{L_0(u), u \in K\}$ on C_K as $\epsilon \to 0$. Let $\hat{u}_\epsilon$ denote the infimum of the points of the maxima of the random field $\{L_\epsilon(u), u \in K\}, \epsilon > 0$ on C_K. Let u_0 denote the location of the maxima of the process $\{L_0(u), u \in K\}$ on C_K. The location u_0 of the maxima is unique almost surely by the property of Gaussian random processes. Since the random process $\{L_\epsilon(u), u \in K\}, \epsilon > 0$ on C_K converge weakly to the random field $\{L_0(u), u \in K\}$ on C_K as $\epsilon \to 0$, by the continuous mapping theorem, it follows that the distribution of $\hat{\theta}_\epsilon$ appropriately normalized converges in law to the distribution of u_0 by the continuous mapping theorem (cf. Billingsley (1968)). Lemma 4.6 implies that the random variable $\hat{u}_\epsilon = \epsilon^{-1}(\hat{\theta}_\epsilon - \theta) \in K$ with probability tending to one as $\epsilon \to 0$. Applying arguments similar to those in Theorem 10.1 in Chapter II, p.103 of Ibragimov and Khasminskii (1981) (cf. Prakasa Rao (1968)), we obtain the following result. Let θ be the true parameter. As a consequence of the arguments and the discussion given above, it follows that the random variable

$$\hat{u}_\epsilon = \epsilon^{-1}(\hat{\theta}_\epsilon - \theta)$$

converges in law to the distribution of the random variable u_0, the location of the maximum of the random field $\{L_0(u), -\infty < u, v < \infty\}$, as $\epsilon \to 0$, which is the Gaussian distribution with mean zero and variance σ^{-2}.

Acknowledgement

Work of B.L.S. Prakasa Rao was supported under the scheme "Ramanujan Chair Professor" at the CR Rao Advanced Institute for Mathematics, Statistics and Computer Science, Hyderabad, India.

References:

1 Billingsley, P. (1968) *Convergence of Probability Measures*, Wiley, New York.

2 Brouste, A. and Kleptsyna, M.L. (2010) Asymptotic properties of MLE for partially observed fractional diffusion system, *Statistical Inference for Stochastic Processes*, **13**, 1-13.

3 Fokianos, K and Newmann, M . (2013) A goodness-of-fit test for Poisson count processes, *Electronic Journal of Statistics*, **7** , 793-819.

4 Ibragimov, I.A. and Has'minskii, R.Z. (1981) *Statistical Estimation: Asymptotic Theory*, Springer, New York.

5 Kallianpur, G. and Selukar, R.S. (1991) Parameter estimation in linear filtering, *Journal of Multivariate Analysis*, **39**, 284-304.

6 Kallianpur, G. and Selukar, R.S. (1993) Estimation of Hilbert space valued parameters by the method of sieves, *Statistics and Probability: A Raghu Raj Bahadur Festschrift*, Ed. J.K. Ghosh, S.K. Mitra, K.R. Parthasarathy and B.L.S. Prakasa Rao, Wiley Eastern, New Delhi, pp. 325-348.

7 Kleptsyna, M.L., Kloden, P.E., and Ahn, V.V. (1998) Linear filtering with fractional Brownian motion, *Stochastic Analysis and Applications*, **16**, 907-914.

8 Kleptsyna, M.L. and Le Breton, A. (2002a) Statistical analysis of the fractional Ornstein-Uhlenbeck type process, *Statistical Inference for Stochastic Processes*, **5**, 249-271.

9 Kleptsyna, M.L. and Le Breton, A. (2002b) Extension of the Kalman-Bucy filter to elementary linear systems with fractional noise, *Statistical Inferrnce for Stochastic Processes*, **5**, 229-248.

10 Kleptsyna,M.L., Le Breton, A., and Roubaud, M.-C. (2000a) Parameter estimation and optimal filtering for fractional type stochastic systems, *Statistical Inference for Stochastic Processes*, **3**, 173-182.

11 Kleptsyna,M.L., Le Breton, A., and Roubaud, M.-C. (2000b) General approach to filtering with fractional Brownian noises-Application to linear systems., *Stochastics and Stochastics Reports*, **71**, 119-140.

12 Kutoyants, Yu. A. (1994) *Identification of Dynamical Systems with Small Noise*, Kluwer, Dordrecht.

13 Leadbetter, M.R., Lindgren, G. and Rootzen, H. (1983) *Extremes and Related Properties of Random Sequences and Processes*, Springer, New York.

14 Le Breton, A. (1998) Filtering and parameter estimation in a simple linear model driven by a fractional Brownian motion, *Statistics and . Probability Letters*, **38**, 263-274.

15 Matsui, Muneya and Shieh, Narn-Rueih. (2009) On the exponentials of fractional Ornstein-Uhlenbeck processes, *Electronic Journal of Probability*, **14**, 594-611.

16 Mishra, M.N. and Prakasa Rao, B.L.S. (2014) Estimation of change point for switching fractional diffusion processes, *Stochastics: An International Journal of Probability and Stochastic Processes*, **86**, 429-449.

17 Mishra, M.N. and Prakasa Rao, B.L.S. (2016) Local asymptotic normality and estimation via Kalman-Bucy filter for linear systems driven by fractional Brownian motions, *Stochastic Analysis and Applications*, **34**, 707-721.

18 Neuhaus, G. (1971) On weak convergence of stochastic processes with multi-dimensional time parameter, *Annals of Mathematical Statistics*, **42**, 1285-1294.

19 Parthasarathy, K.R. (1967) *Probability Measures on Metric Spaces*, Academic Press, London.

20 Prakasa Rao, B.L.S. (1968) Estimation of the location of the cusp of a continuous density, *Annals of Mathematical Statistics*, **39**, 76-87.

21 Prakasa Rao, B.L.S. (1975) Tightness of probability measures generated by stochastic processes on metric spaces, *Bulletin of the Institute of Mathematics, Academia Sinica*, **3**, 353-367.

22 Prakasa Rao, B.L.S. (1987) *Asymptotic Theory of Statistical Inference*, Wiley, New York.

23 Prakasa Rao, B.L.S. (1999) *Semimartingales and their Statistical Inference*, Chapman and Hall/CRC, Boca Raton,.

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