DOI: 10.28919/ejma.2025.5.11

Eur. J. Math. Appl. (2025)5:11
URL: http://ejma.euap.org


# ANALYSIS OF THE FRACTIONAL COX-INGERSOLL-ROSS MODEL BASED ON OPTIMAL STOPPING RULES

JAYA P. N. BISHWAL

Abstract. In this paper, based on optimal stopping rules, we study the inverse problem of sequential inference of the unknown mean reversion parameters in the fractional Cox-Ingersoll-Ross model which has been the main building block for interest rate and stochastic volatility models. For forward problem, this type of observations are used in the pricing of American options. We observe the process both continuously and discretely in time. Continuous observation has theoretical interest and discrete observations have practical interest. We have a unified theory for the subcritical, critical and supercritical cases. We discuss several stopping rules based on barrier, threshold and observed Fisher information. We also consider processes with jumps and long-memory.

## 1. Introduction

There are close connections between some models in biology and finance. Feller (1951) reached at the square-root process as the weak limit of Galton-Watson branching process with immigration while studying a problem in genetics. Using the Feller's square-root process, Cox *et al.* (1985) studied the theory of term structure of interest rates and the model is now known as the Cox-Ingersoll-Ross model. Overbeck and Ryden (1997) studied asymptotics of conditional least squares estimators of Cox-Ingersoll-Ross process from discrete observations using an auto-regressive type representation of the model with non-Gaussian error. Dehtiar *et al.* (2021) studied strong consistency for the maximum likelihood method and an alternative method of estimation of the drift parameters of the Cox-Ingersoll-Ross process based on continuous observations. Mishura and Yurchenko-Tytarenko (2018) studied hitting probability of fractional Cox-Ingersoll-Ross model which involves long memory. Mackevicius (2015) used stochastic Verhulst model as an alternative to CIR model for modeling interest rate as both processes have similar behavior. Mackevicius (2011) studied weak approximation of CIR equation by discrete random variables. Lenkasas and Mackevicius (2015) obtained a second order weak approximation of Heston model by discrete random variables. Lileika and Mackevicius (2020) studied weak approximation of CKLS and CEV process (cf. Cox (1996)) by discrete random variables. The Cox-Ingersoll-Ross (CIR) model is extensively used as a short rate mean reverting model

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

*E-mail address*: J.Bishwal@charlotte.edu.

*Key words and phrases.* Itô stochastic differential equation; fractional Cox-Ingersoll-Ross model; branching diffusion; fractional Brownian motion; fractional Poisson process; fractional Levy process, sequential estimator; continuous observation; discrete observations; local asymptotic quadraticity; functional local asymptotic normality; local asymptotic minimaxity; completeness, sufficiency; efficiency; jumps, long memory; interest rate; stochastic volatility; American option.

in term structure of interest rates and a stochastic volatility process in the Heston model, see Bishwal (2022). In view of this, it becomes necessary to estimate the unknown parameters in the model from discrete data. See Bishwal (2008) for asymptotic results on approximate likelihood asymptotics and approximate Bayes asymptotics for drift estimation of discretely observed diffusions based on high frequency data.

Li and Linetsky (2014) studied time-changed Ornstein-Uhlenbeck processes and their applications in commodity derivative models. The Cox-Ingersoll-Ross model, also called the square-root process, is used as a short rate mean reverting model in term structure of interest rates and in membrane potential evolution in single neurons in the nervous system, both based on high frequency discrete data. In view of this, it becomes necessary to estimate the unknown parameters in the model. Ditlevsen and Samson (2014) studied the estimation in the partially observed stochastic Morris-Lecar neuronal model with particle filter and stochastic approximation method. The long time asymptotics of the maximum likelihood estimator (MLE) and the Bayes estimators (BEs) of the drift parameter in the nonlinear nonhomogeneous Markov diffusion processes was studied by Borkar and Bagchi (1982), Levanony, Shwarz and Zeitouni (1994) and Kutoyants (1984b, 2003). Dietz (1989) asymptotic properties of MLE as the intensity of noise $\epsilon \to 0$ or the observation time $T \to \infty$ when the model satisfies the LAMN condition. Dietz (1992) studied the properties of MLE in a concrete example of diffusion type process which is an exponential memory nonhomogeneous process, a non-Markovian alternative to the Ornstein-Uhlenbeck process as the observation time $T \to \infty$. Kutoyants (1984a, 1994) also studied the asymptotic properties of MLE as the diffusion coefficient $\epsilon \to 0$ (for fixed $T$). For a linear stochastic differential equation with time delay, Gushchin and Küchler (1998) showed that the MLE shows eleven different behaviors for eleven parts of the parameter space. See Bishwal (2008, 2022) for large time asymptotic results on approximate likelihood estimators and approximate Bayes estimators of the unknown drift parameter of discretely observed diffusions based on high frequency data. We estimate the drift parameter by the sequential maximum likelihood (SML) method and study the asymptotic minimaxity of the resulting estimators.

For first order efficiency (in the sense of C.R. Rao) of estimators for stochastic processes, see Hall and Heyde (1980). We prove the local asymptotic minimaxity of the SMLE in the Hajek-Le Cam sense. Roughly speaking an estimator is said to be locally asymptotically minimax if it attains the lower bound in Hajek-Le Cam minimax theorem (see Jeganathan (1982)), i.e. if it attains the lower bound to the local asymptotic minimax risk of the normalized error an estimator. The minimum requirement for Hajek's minimax theorem is that the model should satisfy the LAN (locally asymptotically normal) or LAMN (locally asymptotically mixed normal) condition. When these conditions are satisfied the lower bound is attained only if the estimator is *asymptotically centering* (AC) (see Jeganathan (1982) for a definition). But there are situations where either of the above two conditions may not be attained. Consider, for example, the Ornstein-Uhlenbeck process with drift coefficient $\theta X_t$. This process exhibits qualitatively different behaviour for different values of the parameter $\theta$. For $\theta < 0$, the model satisfies the LAN condition and for $\theta > 0$ the model satisfies the LAMN condition. The point $\theta = 0$ is critical. At $\theta = 0$, it satisfies neither the LAN condition nor the LAMN condition, but it satisfies the LABF (locally asymptotically Brownian functional) condition. The model satisfies LAQ (locally asymptotically quadraticity) (see Le Cam and Yang (1990), Jeganathan

(1995)) for all $\theta$. Similar situations occur in its discrete time counterpart : the Gaussian autoregressive process of first order and other proceses like the Galton-Watson branching processes, pure birth processes etc. (see Guschin (1995)). Greenwood and Shiryayev (1992) proved the uniform local asymptotic minimaxity of the SMLE of the parameter in the first order Gaussian autoregressive process by studying the uniform weak convergence of statistical experiments using the convergence of the associated Hellinger processes. Under the LAQ condition, Hajek's minimax theorem is available, but the AC estimators do not attain the lower bound, i.e., they will not be locally asymptotically minimax. Recently Greenwood and Wefelmeyer (1993) showed that local asymptotic minimax bound is attained by asymptotically centering estimators even at critical points, which requires sequential sampling. Höpfner, R. (1993a) studied the statistics of Markov processes with representation of log-likelihood ratio processes in filtered local models. Höpfner, R. (1993b) studied asymptotic inference for Markov step processes suing observation up to a random time. Li and Linetsky (2015) studied discretely monitored first passage problems and barrier options using an eigenfunction expansion approach. Lipton and Kaushansky (2020) studied the first hitting time density for a reducible diffusion process. Löcherbach (2002b) studied LAN and LAMN for systems of interacting diffusions with branching and immigration. Alili *et al.* (2005) studied the representations of first hitting time density of an Ornstein-Uhlenbeck process.

For the linear diffusion model where the drift coefficient is $b(\theta, t, x) = \theta a(t, x)$ and the diffusion coefficient is $\sigma(t, x) = 1$, Novikov (1972) (see also Liptser and Shiryayev (1978)) proved that the SMLE of $\theta$ is unbiased and exactly normally distributed for all values of the parameter in the parameter space $\Theta \subset \mathbb{R}$. Further, he showed that SMLE is optimal in the mean square sense and is more efficient in the sense of having less mean square error than the ordinary MLE based on fixed time observation, under the assumption that the mean durations of observation time in both the sampling plans are the same. Tikhov (1978) proved that for the case $b(\theta, t, x) = \theta a(t)$ and $\sigma(t, x) = 1$, the SMLE is optimal relative to the power loss function $L_\alpha(|\delta(x) - \theta|) = |\delta(x) - \theta|^\alpha, \alpha \geq 1$. Tikhov (1980) generalised this to the case $b(\theta, t, x) = \theta b_1(t, x) + b_0(t, x)$. Sørensen (1983) gave a review of sequential maximum likelihood estimation in linearly parametrized diffusion type processes. Sørensen (1986) (see also Küchler and Sørensen (1997)) studied similar properties of SMLE for exponential families of stochastic processes. Musiela (1977, 1979) studied sequential ML estimation in a linear diffusion model. Le Breton and Musiela (1981) studied sequential estimation of parameters of continuous Gaussian Markov processes. Le Breton and Musiela (1985) studied similar properties of SMLE in linear homogeneous multidimensional SDE. Rozanskii (1989) extended the work of Novikov (1972) to a linear homogeneous diffusion field. Melnikov and Novikov (1988) extended the work of Novikov (1972) from linear diffusion model to a linear semimartingale model. Brown and Hewitt (1975) studied the properties of sequential maximum likelihood estimator for diffusion branching process using a different type of stopping rule. Löcherbach (2002a) studied likelihood ratio processes for Markovian particle systems with killing and jumps.

We estimate the unknown parameter by means of observation of the process until the observed Fisher information exceeds a predetermined level of precision. This idea of using observed Fisher information to define a stopping rule dates back to Anscombe (1952) (see also Grambsch (1983) and Ghosh *et al.* (1997). This type of stopping rule was used for autoregressive parameter estimation in Lai and Siegmund (1985), Dimitrinko and Konev (1994), Dimitrinko *et al.* (1997), Malinovskii (1993), Pergamenshchikov (1991) among others.

## 2. Continuous Observation

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a stochastic basis on which is defined the Cox-Ingersoll-Ross process $\{X_t\}$ satisfying the Itô stochastic differential equation

$$dX_t = (\alpha - \beta\ X_t)\ dt + \ 2\sigma\sqrt{X}_t\ dW_t,\ t \geq 0, \quad X_0 = x_0 \tag{2.1}$$

where $\{W_t\}$ is a standard Wiener process with the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ and $\alpha, \beta, \sigma > 0$ are the unknown parameters to be estimated on the basis of continuous and discrete observations of the process.

Let the continuous realization $\{X_t, 0 \leq t \leq T\}$ be denoted by $X_0^T$. First we consider the case where $\alpha = 0$ and $2\sigma = 1$, that is the diffusion branching process. Let $P_\beta^T$ be the measure generated on the space $(C_T, B_T)$ of continuous functions on $[0, T]$ with the associated Borel $\sigma$-algebra $B_T$ generated under the supremum norm by the process $X_0^T$ and let $P_0^T$ be the standard Wiener measure. It is well known that when $\beta$ is the true value of the parameter $P_\beta^T$ is absolutely continuous with respect to $P_0^T$ and the Radon-Nikodym derivative (likelihood) of $P_\beta^T$ with respect to $P_0^T$ based on $X_0^T$ is given by

$$L_T(\beta) := \frac{dP_\beta^T}{dP_0^T}(X_0^T) = \exp\left\{\beta\int_0^T dX_t - \frac{\beta^2}{2}\int_0^T X_t dt\right\}. \tag{2.2}$$

Consider the score function, the derivative of the log-likelihood function $l_T(\beta)$, which is given by

$$l_T'(\beta) := \int_0^T dX_t - \beta\int_0^T X_t dt. \tag{2.3}$$

A solution of the estimating equation $l_T'(\beta) = 0$ provides the maximum likelihood estimate (MLE)

$$\hat{\beta}_T := \frac{X_T - X_0}{\int_0^T X_t dt}. \tag{2.4}$$

Let us consider the distribution of the continuous energy process $I_T$ which is a sufficient statistic for estimation of $\beta$, the other sufficient statistic being $X_T - X_0$. The asymptotic distribution of the $I_T$ process is closely related to the inverse Gaussian distribution with parameters $(T/2,\ T/2)$.

The maximum likelihood estimator of $\beta$ in the diffusion branching process (2.1) is known to have different limit distributions for the three cases $\beta < 0$ (subcritical), $\beta = 0$ (critical) and $\beta > 0$ (supercritical). This paper is concerned with the two cases $\beta = 0$ and $\beta > 0$. It will be useful to develop a unified approach to estimation which does not require the prior information as to whether $\beta = 0$ or $\beta > 0$. We propose to use sequential approach to achieve this.

Overbeck (1998) studied large sample asymptotics for the maximum likelihood estimation in diffusion branching process. Brown and Hewitt (1975) studied sequential estimation of the diffusion branching process. They used a one barrier stopping rule.

Define the stopping rule

$$\tau_h := \inf\{t \geq 0 : \int_0^t X_s ds \geq h\} \tag{2.5}$$

where $h$ is the known precision. This stopping time is inspired by observed Fisher information. $X_\tau$ and $\int_0^\tau X_s ds$ are *complete sufficient statistics* for $\alpha$ and $\beta$.

Define the stopping rule

$$\tau := \inf\{t \geq 0 : X_t = C\}. \tag{2.6}$$

If one uses this stopping rule, $X_\tau$ is fixed and $\int_0^\tau X_s ds$ is alone is a sufficient statistic. In fact the form of the likelihood function will show that $\int_0^\tau X_s ds$ is a complete sufficient statistic. The distribution of the sufficient statistic is inverse Gaussian.

Both OU and CIR models have non-oscillatory natural boundaries. Hence the spectrum is purely discrete. Hitting time densities for CIR and OU diffusions was studied in Linetsky (2004) in terms of Sturm-Liouville eigenfunction expansions. See equation (40) in that paper.

Define the first hitting time as

$$\tau := \inf\{t \geq 0 : X_t = C\} \tag{2.7}$$

where $C$ is the mean reversion level. The process is positive recurrent:

$$P(\tau < \infty) = 1, \quad E_x(\tau) < \infty. \tag{2.8}$$

For the O-U process, the density of the first hitting time is given by

$$f_\tau(t) = \frac{|x - C|}{\sqrt{2\pi}\sigma} \left(\frac{\kappa}{\sinh(\kappa t)}\right) \exp\left(\frac{\kappa t}{2} - \frac{\kappa(x - C)^2 e^{-\kappa t}}{2\sigma^2 \sinh(\kappa t)}\right) \tag{2.9}$$

where $\kappa$ is the mean-reversion speed and $C$ is the mean-reversion level. See Ricciardi and Sato (1988). They also calculated moments of the distribution. See also Going-Jaeschke and Yor (2003) who used Laplace transform of Bessel process.

The closed form solution is inverse Gaussian if $x < 0$ and $C = 0$ (first hitting time to 0):

$$f_\tau(t) = \frac{2|x|}{\sqrt{2\pi}}(e^{2t} - 1)^{-3/2} e^{2t} \exp\left(-\frac{x^2}{2(e^{2t} - 1)}\right). \tag{2.10}$$

Consider the random time $\rho_t = \int_0^t X_s ds$ and the associated process $Y_{\rho_t} = X_t$. The process $Y$ is a Brownian motion with drift $\beta$, acceleration 1 and $Y_0 = 0$.

Define the stopping rule

$$\tau := \inf\{t \geq 0 : Y_t = A\}$$

which is the first passage time of $Y_t$ reaching a barrier $A$. The distribution of $\tau$ is inverse Gaussian. Note that $\tau$ is a complete sufficient statistic for $\beta$ and $A\tau^{-1} - A^{-1}$ is a minimum variance unbiased estimator (MVUE) for $\beta$.

The sequential MLE is given by

$$\tilde{\beta}_\tau := \frac{1}{h}\left[\int_0^\tau X_s dX_s - (X_\tau - X_0)\bar{X}_\tau\right], \tag{2.11}$$

$$\tilde{\alpha}_\tau := -\bar{X}_\tau \tilde{\beta}_\tau + \frac{1}{\tau}(X_\tau - X_0). \tag{2.12}$$

*Threshold Based Estimator*

We study discrete sequential estimator based on observations of first hitting time.

Consider the general SDE

$$dX_t = \mu(X_t,\theta)dt + \sigma(X_t)dW_t, X_0 = x \tag{2.13}$$

$$\tau_h = \inf\{t \geq 0 : X_t = A\} = \inf\{t \geq 0 : M_t = h\} \tag{2.14}$$

where $M_t = \sup_{s\leq t} X_s$ is the record process.

The truncated jump process is defined as

$$\hat{\tau}_h^\eta = \sum_{x\leq u<h} \Delta T_u \ I_{(\Delta T_u \geq \eta)}. \tag{2.15}$$

The contrast function is given by

$$U_h(\theta) = \frac{1}{2}\int_{[x,h)} \frac{\mu^2(u,\theta)-\mu^2(u,\theta_0)}{\sigma^2(u)} d\tau_u - \int_x^h \frac{(\mu(u,\theta)-\mu(u,\theta_0)}{\sigma^2(u)} du. \tag{2.16}$$

The log-likelihood is given by

$$\log L_{\tau_h}(\theta) = \int_0^{\tau_h} \frac{\mu(X_s,\theta)-\mu(X_s,\theta_0)}{\sigma^2(X_s)} dX_s - \frac{1}{2}\int_0^{\tau_h} \frac{\mu^2(X_s,\theta)-\mu^2(X_s,\theta_0)}{\sigma^2(X_s)} ds. \tag{2.17}$$

Observe that

$$-\log L_{\tau_h}(\theta) = U_h(\theta) + o_P(\Delta). \tag{2.18}$$

$U_h(\theta)$ converges in probability to

$$K(\theta_0,\theta) = \frac{1}{2}\int_x^h \frac{(\mu(u,\theta)-\mu(u,\theta_0))^2}{\sigma^2(u)\mu(u,\theta_0)} du. \tag{2.19}$$

the Kullback-Lieiber information.

$$\int_{[x,h)} \phi(u) d\tau_u \to^P \int_x^h \frac{\phi(u)}{\mu(u,\theta_0)} du. \tag{2.20}$$

Converting $L_{\tau_h}$ by Itô formula, we have

$$\begin{aligned} \log L_{\tau_h} \;&=\; \int_x^h \frac{(\mu(u,\theta)-\mu(u,\theta_0)}{\sigma^2(u)} du \\ &\quad -\frac{1}{2}\int_0^{\tau_A} \left[ \frac{(\mu^2(X_s,\theta)-\mu^2(X_s,\theta_0)}{\sigma^2(X_s)} + \sigma^2(X_s)\frac{\partial}{\partial X_s}\frac{\mu(X_s,\theta)-\mu(X_s,\theta_0)}{\sigma^2(X_s)} \right] ds. \end{aligned} \tag{2.21}$$

We know that for $0 \leq t \leq \tau_h$ and $x \leq a \leq h$, $(M_t \geq h) = (\tau_h \leq t)$. Hence

$$\int_0^{\tau_h} \phi(M_s) ds = \int_{[x,h)} \phi(u) d\tau_u. \tag{2.22}$$

*Examples*

1) For WD (Wiener process with drift) case, $\mu(u,\theta) = \theta$,

$$\widehat{\theta}_h = \frac{h-x}{\tau_h}. \tag{2.23}$$

2) For the OU (Ornstein-Uhlenbeck) case, $\mu(u,\theta) = \theta u$,

$$\widehat{\theta}_h = \frac{\int_x^h u du}{\int_{[x,h)} u^2 d\tau_u} = \frac{h^2 - x^2}{2\int_{[x,h)} u^2 d\tau_u} \tag{2.24}$$

3) For the CIR (Cox-Ingersoll-Ross) case, $\mu(u,\theta) = a + bu,\ a > 0,\ b > 0$

$$\widehat{b}_h = \frac{\log(h/x)\int_{[x,h)} ud\tau_u - \tau_h(h-x)}{\int_{[x,h)} u^{-1}d\tau_u \int_{[x,h)} ud\tau_u - \tau_h^2}, \tag{2.25}$$

$$\widehat{a}_h = \frac{(h-x)\int_{[x,h)} u^{-1}d\tau_u - \tau_h \log(h/x)}{\int_{[x,h)} u^{-1}d\tau_u \int_{[x,h)} ud\tau_u - \tau_h^2}. \tag{2.26}$$

# 3. Discrete Observations

Discretization of stopping time is a delicate problem. Li and Linetsky (2013) studied discretization of stopping time and optimality. Let $0 = t_0 < t_1 < \cdots t_n = T$ and $t_i = i\Delta,\ i = 1, 2\cdots, n$ with $h$ fixed. Let $\{X_t\}$ be observed at times $0 = t_0 < t_1 < \cdots t_n = T$ with $t_i - t_{i-1} = \frac{T}{n} = \Delta, i = 1, 2\cdots, n$.

Define the one barrier stopping time

$$\gamma_{n,B} := \inf\{n \geq i \geq 1 : X_{t_i} \geq B\} \tag{3.1}$$

where $B > 0$. Define the sequential estimator as

$$\hat{\beta}_{\gamma_{n,B}} = \frac{B-1}{\sum_{i=1}^{\gamma} X_{t_{i-1}}(t_i - t_{i-1})} - \frac{1}{B-1}. \tag{3.2}$$

The statistic $\sum_{i=1}^{\gamma} X_{t_{i-1}}(t_i - t_{i-1})$ is a complete sufficient statistic. Thus $\hat{\beta}_\gamma$ is the minimum variance unbiased estimator (MVUE).

Let $Y_t := \sqrt{X_t}$. By Itô formula

$$dY_t = \left(-\frac{1}{8Y_t} - \frac{1}{2}Y_t\right)dt + \frac{1}{2}dW_t,\ t \geq 0. \tag{3.3}$$

This is a Bessel process. Using Itô formula, one can derive the SDE for $R_t = 4/(\sigma^2 Y_t^{-2})$ which is given by

$$dR_t = -\alpha(R_t - \beta)R_t dt + \sigma R_t^{3/2}\, dW_t. \tag{3.4}$$

This process , known as 3/2 model, was proposed by Cox *et al.* (1985, p. 402, Equation 50) as a model for the inflation rate in their three-factor inflation model.

We will approach the CIR model as a weak limit of branching processes. Let $Z_n = \{Z_{n,k} : k = 0, 1, 2, \ldots\}$ be a sequence of simple branching processes whose offspring distributions have means $m_n$, variances $\sigma_n$ ; and are uniformly square integrable in $n$.

$$Z_{n,[nt]} = \sum_{k=1}^{Z_{n-1,[(n-1)t]}} \xi_{n-1,k} + \zeta_n \tag{3.5}$$

where $\xi_{n-1,k}$ is the number of offsprings of the $k$th individual belonging to the $(n-1)$th generation and $\zeta_n$ denotes the number of immigrants in the $n$th generation. Suppose that $\{\xi_{n-1,k}\}, k = 0, 1, 2, \ldots, n = 1, 2, \ldots$, and $\{\zeta_n\}, n = 1, 2, \ldots$ are two independent sequences of independent and identically distributed (i.i.d.) random variables. The initial state $Z_0$ is a random variable which is independent of $\{\xi_{n-1,k}\}$ and $\{\zeta_n\}$ has an arbitrary distribution. The offspring and the immigration distributions are assumed to be unspecified with means $m_n$ and $\lambda_n$ and variances $\sigma_n$ respectively.

Define

$$X_{n,t} := \frac{1}{n} Z_{n,[nt]}. \tag{3.6}$$

Then $X_n \to X$ weakly in the sense of convergence of the corresponding probability measures on $D[0, \infty)$ if

$$\frac{1}{n} Z_{n,0} \to \nu, \ n(m_n - 1) \to \beta, \ \sigma_n \to 2 \text{ as } n \to \infty. \tag{3.7}$$

By the Levy result for quadratic variation of Brownian motion

$$\lim_{n\to\infty} \sum_{i=1}^{n} [X_{t_i} - X_{t_{i-1}}]^2 = 2 \int_0^t X_s ds \ \text{ a.s.} \tag{3.8}$$

$$\hat{\beta}_{n,T} = \frac{X_T - X_0}{\sum_{i=1}^{n} [X_{t_i} - X_{t_{i-1}}]^2} \to \beta_T = \frac{X_T - X_0}{\int_0^T X_t dt} \ \text{ a.s.} \tag{3.9}$$

$$\hat{\beta}_{n,\tau} = \frac{X_\tau - X_0}{\sum_{i=1}^{n\wedge[n\tau]} [X_{t_i} - X_{t_{i-1}}]^2} \to \beta_\tau = \frac{X_\tau - X_0}{\int_0^\tau X_t dt} \ \text{ a.s.} \tag{3.10}$$

Define the stopping rule

$$N := \inf\{n \geq 1 : \sum_{i=1}^{n} (X_i - \bar{X}_n)^2 \geq h\}. \tag{3.11}$$

Define the sequential estimators

$$\widehat{\beta}_{N,h} := \frac{1}{\Delta} \log \frac{\sum_{i=1}^{N} (X_i - \bar{X}_N)(X_{i-1} - \bar{X}'_N)}{\sum_{i=1}^{N} (X_i - \bar{X}_N)^2}, \quad \widehat{\alpha}_{N,h} := \frac{\bar{X}_N - e^{\hat{\beta}_{N,h}\Delta} \bar{X}'_N}{e^{\hat{\beta}_{N,h}\Delta} - 1} \hat{\beta}_{N,h} \tag{3.12}$$

where

$$\bar{X}_N := \sum_{i=1}^{N} X_i, \quad \bar{X}'_N := \sum_{i=1}^{N} X_{i-1}. \tag{3.13}$$

### Discretization of Continuous Stopping Time

The discrete stopping time is given by

$$N_h := \inf\{n \in \mathbb{N} : n\Delta \geq \tau_h\}. \tag{3.14}$$

We assume that $P(\tau_a < \infty | X_0 = x) = P(N < \infty | X_0 = x) = 1$.

Now let us introduce the killed process as follows: Let $X_t^k = X_t$ when $X_t < h$ and $X_t = C$ (the coffin state) when $X_t \geq h$. Let the discretization of $X_t^k$ be $X_i^k$ which is defined as: $X_i^k = X_i$ for $i < N_H$ and $X_i^k = C$ for $i \geq N_h$.

Observing the original discretized process $X_i$ up to the stopping time $N - 1$ is equivalent to observing the killed process $X_i^k$ infinitely, since from the first visit to $C$ no more information is gained. In both cases, we can interpret each observed trajectory as a realization of a single random variable.

The discrete sequential estimators are given by

$$\widehat{b}_{N_h,h} = -\frac{1}{\Delta} \log \left( \frac{N_h \sum_{i=1}^{N_h} X_i X_{i-1}^{-1} - \sum_{i=1}^{N_h} X_i \sum_{i=1}^{N_h} X_{i-1}^{-1}}{{N_h}^2 - \sum_{i=1}^{N_h} X_i \sum_{i=1}^{N_h} X_{i-1}^{-1}} \right), \tag{3.15}$$

$$\widehat{a}_{N_h,h} = \frac{1}{N_h} \sum_{i=1}^{N_h} X_i + \frac{e^{-\hat{b}_{N_h,h}\Delta}(X_{N_h} - x)}{N_h \hat{b}_{N_h,h}(1 - e^{-\hat{b}_{N_h,h}\Delta})}. \tag{3.16}$$

## 4. Sequential Estimation for Fractional Levy CIR Process

For direct pricing problem, random observation period and optimal stopping problem are related to American options. Valuation of real options of American type under persistent shocks (color noise) was studied in Bishwal (2017). We focus on the inverse problem.

Recall that in the singular case for zero innovation mean locally asymptotically Brownian functional (LABF) condition holds while for nonzero innovation mean, local asymptotic normality (LAN) condition holds, see Bishwal (2018). Sequential estimation unifies the ergodic, nonergodic and singular case and gives asymptotic normality in all three cases, see Bishwal (2018). Sequential maximum likelihood estimation in semimartingales was studied in Bishwal (2006). Sequential estimation in Hilbert space valued stochastic differential equations was studied in Bishwal (1999). Sequential maximum likelihood estimation for reflected Ornstein-Uhlenbeck process was studied in Lee *et al.* (2012).

First consider the FLCIR model

$$dY_t = a(b - Y_t)dt + \sigma\sqrt{Y_t}dV_t^H \tag{4.1}$$

where $V_t^H$ is a fractional Levy (FL) process with Hurst parameter $H > 1/2$.

Fractional Levy Process (FLP) is defined as

$$V_t^H = V_{H,t} = \frac{1}{\Gamma(H+\frac{1}{2})}\int_{\mathbb{R}}[(t-s)_+^{H-1/2} - (-s)_+^{H-1/2}]d\widetilde{V}_s, \quad t \in \mathbb{R}$$

where $\{\widetilde{V}_t, t \in \mathbb{R}\}$ is a Levy process on $\mathbb{R}$ with $E(\widetilde{V}_1) = 0,\ E(\widetilde{V}_1^2) < \infty$.

Here are some properties of the fractional Levy process:

1) The covariance of the process is given by

$$\mathrm{cov}(V_{H,t}, V_{H,s}) = \frac{E(\widetilde{V}_1^2)}{2\Gamma(2H+1)\sin(\pi H)}[|t|^{2H} + |s|^{2H} - |t-s|^{2H}].$$

2) $V_H$ is not a martingale. For a large class of Levy processes, $V_H$ is neither a semimartingale.

3)$V_H$ is Hölder continuous of any order $\beta$ less than $H - \frac{1}{2}$.

4) $V_H$ has stationary increments.

5) $V_H$ is symmetric.

6) $\widetilde{V}$ is self-similar, but $V_H$ is not self-similar.

7) $V_H$ has infinite total variation on compacts.

Thus FLP is a generalization and a natural counterpart of FBM. Fractional stable motion (FSM) is a special case of FLP. Fractional Poisson process (FPP) is a special case of FLP.

A fractional Poisson process $\{N^H(t), t > 0\}$ with Hurst parameter $H \in (1/2, 1)$ is defined as

$$N^H(t) = \frac{1}{\Gamma(H-\frac{1}{2})}\int_0^t u^{\frac{1}{2}-H}\left(\int_u^t \tau^{H-\frac{1}{2}}(\tau-u)^{H-\frac{3}{2}}d\tau\right)dq(u)$$

where $q(u) = \frac{N(u)}{\sqrt{\lambda}} - \sqrt{\lambda}u$ and $N(u)$ is a homogeneous Poisson process with intensity $\lambda > 0$.

The covariance of $N^H$ is given by

$$E(N^H(t)N^H(s)) = \frac{R_H^2}{2}(t^{2H} + s^{2H} - |t-s|^{2H})$$

where

$$R_H^2 := -\frac{\Gamma(2-2H)\cos(\pi H)}{(2H-1)\pi H}.$$

The process is self-similar in the wide sense, has wide sense stationary increments, has fat-tailed non-Gaussian distribution, and exhibits long range dependence. The process converges to fractional Brownian motion in distribution. The process is self-similar in the asymptotic sense.

Then by Proposition 5.7 of Buchmann and Kluppelberg (2006), we have

$$Y_t = f(X_t) \tag{4.2}$$

where

$$dX_t = a(b - X_t)dt + dV_t^H, \quad X_0 = f^{-1}(Y_0), \quad t \in [0,T] \tag{4.3}$$

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

Let $b = 0,\ \sigma = 1$ and $a = \theta_1$. Then $X_t$ is described by the Ornstein-Uhlenbeck SDE

$$dX_t = -\theta_1 X_t dt + dV_t^H(\theta_2),\ X_0 = f^{-1}(Y_0) \tag{4.4}$$

where $\theta_1 \in \Theta_1 = \mathbb{R}$ and $V(\theta_2)$ is a process with stationary independent increments with $V_0(\theta_2) = 0$ and Levy characteristics $(b(\theta_2)t,\ ct,\ L(\theta_2)t)$ depending on a parameter $\theta_2$ from an arbitrary set $\Theta_2$. We assume that the trajectories of $V(\theta_2)$ are right continuous with left limits. This model provides the natural analogue of the discrete time of AR(1) models with i.i.d. innovations.

For $\theta = (\theta_1, \theta_2) \in \Theta = \Theta_1 \times \Theta_2$, let $P_{\theta,t}$ be the distribution of $X$ when observed up to time $t$. For $\theta \in \Theta$, let $P_\Theta$ denote the distribution of the unique solution of the SDE on $\Omega = D(\mathbb{R}_+, \mathbb{R})$ equipped with the coordinate process $X = (X_t)_{t\geq 0}$ and the $\sigma$-algebra $\mathcal{F} = \sigma(X_t : t \geq 0)$ where $D(\mathbb{R}_+, \mathbb{R})$ is the space of real valued functions on $\mathbb{R}_+$ which are right continuous with left limits. Let $(\mathcal{F}_t)_{t\geq 0}$ denote the right continuous filtration generated by $X$.

Under $P_\theta$, the process $V^H(\theta_1)$ defined by $V_t^H(\theta_1) = X_t + \theta_1 \int_0^t X_s ds$ has dependent stationary increments and the driver $\widetilde{V}_t$ has the Levy characteristics $(b(\theta_2)t, ct, L(\theta_2)t)$ relative to some fixed continuous bounded truncation function $h : \mathbb{R} \to \mathbb{R}$ with compact support satisfying $g(x) = x$ in a neighborhood of zero, that is,

$$E_\theta \exp(iz\widetilde{V}_t(\theta_1)) = \exp\left(t\left[izb(\theta_2) - \frac{1}{2}cz^2 + \int (\exp(izx) - 1 - izg(x))L(\theta_2)(dx)\right]\right), \quad z \in \mathbb{R}, \tag{4.5}$$

where $b(\theta_2) \in \mathbb{R}, c > 0$, and $L(\theta_2)$ is a Levy measure on $\mathbb{R}$ satisfying $L(\theta_2)(\{0\}) = 0$ and $\int (x^2 \wedge 1)L(\theta_2)(dx) < \infty$ ( See Jacod and Shiryayev (1987) II.4.19, III.2.26). Thus $X$ is a solution of the above SDE with respect to $V(\theta_1)$. Assume for simplicity that $c = 1$.

Now fix $\tau, \theta \in \Theta$ such that $P_\tau \neq P_\theta$. Let $m(\tau_2) = E_\tau(\widetilde{V}_1(\tau_1))$ and $\sigma^2(\tau_2) = \text{Var}_\tau \widetilde{V}_1(\tau_1)$.

Then

$$m(\tau_2) = b(\tau_2) + \int (x - g(x))L(\tau_2)(dx), \quad \sigma^2(\tau_2) = 1 + \int x^2 L(\tau_2)dx. \tag{4.6}$$

We call $(m(\tau_2)t,\ t,\ L(\tau_2)t)$ Levy characteristics of $\widetilde{V}(\tau_1)$ without truncation.

The basic regularity conditions are the following: We assume the conditions (A1) – (A4).

**(A1)** The Levy measures $L(\tau_2)$ and $L(\theta_2)$ are mutually absolutely continuous and $L(\theta_2)$ is homogeneous and $L^2$ differentiable, i.e.,

$$\int (f(\tau_2,\theta_2))^{1/2}-1)^2 dL(\theta_2) < \infty \text{ where } f(\tau_2,\theta_2) = dL(\tau_2)/dL(\theta_2).$$

There exists $\dot{f}(\theta_2) \in L^2(L(\theta_2))$ such that

$$\int [f(\theta_2+z)^{1/2}-1-\frac{1}{2}z\dot{f}(\theta_2)]^2 dL(\theta_2) = o(z^2) \text{ as } z \to 0 \text{ and } \int \dot{f}(\theta_2)dL(\theta_2) > 0.$$

The assumption (A1) implies that $\int |(f(\tau_2,\theta_2)-1)g|dL(\theta_2) < \infty$. Hence $a(\tau_2,\theta_2) \in \mathbb{R}$. Note also that $a(\tau_2,\theta_2) = -a(\tau_2,\theta_2)$. Moreover, $a(\tau_2,\theta_2)$ does not depend on the choice of the function $g$.

Assume

$$b(\tau_2) - b(\theta_2) = \int (f(\tau_2,\theta_2)-1)g dL(\theta_2).$$

Define the Kullback- Leibler information of Levy measures by

$$K(L(\tau_2,\theta_2)) := \int (f(\theta_2,\tau_2) - 1 - \log f(\theta_2,\tau_2))dL(\tau_2).$$

**(A2)** $K((L(\tau_2), L(\theta_2)) < \infty$.
**(A3)** $\int (\log f(\tau_2,\theta_2))^2 dL(\tau_2) < \infty$.
**(A4)** $\int x^2 L(\tau_2)(dx) < \infty$.

The assumption (A4) is equivalent to $E_\tau V_t^2(\tau_1) < \infty$ for every $t \geq 0$. Let $m(\tau_2) = E_\tau \widetilde{V}_t(\tau_1)$. Then $m(\tau_2) = b(\tau_2) + \int (x - g(x))L(\tau_2)(dx)$ and $\sigma^2(\tau_2) = 1 + \int x^2 L(\tau_2)(dx)$.

Under (A1), $P_{\tau,t}$ and $P_{\tau,t}$ are mutually absolutely continuous and he log-likelihood ratio $\Lambda_T(\tau,\theta) = \log(dP_{\tau,t}/dP_{\theta,t})$ admits the representation

$$\Lambda_T(\tau,\theta) = -\Lambda_T(\theta,\tau) = \int_0^T [(\theta_1-\tau_1)X_{s-} + a(\tau_2,\theta_2)]dX_s^c(\tau) + \frac{1}{2}\int_0^T [(\theta_1-\tau_1)X_s + a(\tau_2,\theta_2)]^2 ds$$
$$+ \int (1 - f(\theta_2,\tau_2))d(\mu - \nu(\tau_2)) + \int_0^T \int (f(\theta_2,\tau_2) - 1 - \log f(\tau_2,\theta_2))d\mu \quad (4.7)$$

where $\mu$ is the Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}$ associated with the jumps of $X$ by $\mu = \sum_{t\geq 0} \varepsilon_{(t,\Delta X_t)} I_{\{\Delta X_t \neq 0\}}$ with with $\Delta X_t = X_t - X_{t-}$, $\Delta X_0 = 0, \nu(\tau_2, dt, dx) = dt \otimes L(\tau_2)dx$ and $d\nu = \nu(\omega, dt, dx)$ is the compensator. Note that $X^c(\tau)$ is a standard Wiener process. Further

$$\sum_{0<s\leq t} I_{(|\Delta X_s|>1)} = \int_0^t \int_{|x|>1} x d\mu. \quad (4.8)$$

This gives

$$\Lambda_T(\tau,\theta) = Y_T + \frac{1}{2}\int_0^T [(\theta_1-\tau_1)X_s + a(\tau_2,\theta_2)]^2 ds + K(L(\tau_2), L(\theta_2))T \quad (4.9)$$

where

$$Y_T = \int_0^T [(\theta_1-\tau_1)X_{s-} + a(\tau_2,\theta_2)]dX_s^c(\tau) + \int_0^T \int \log f(\tau_2,\theta_2)d(\mu - \nu(\tau_2)). \quad (4.10)$$

We assume

$$\int |x| dL(\theta_3, dx) < \infty \quad \text{and} \quad E_\theta V(\theta) = \int x dL(\theta_3) \quad (4.11)$$

for every $\theta$.

The maximum likelihood estimator (MLE) based on the observations in $[0,T]$ is given by

$$\hat{\theta}_{1,T} := \frac{T - X_T^2 + \sum_{s\le T}\Delta X_s^2}{2\int_0^T X_s^2 ds}.$$

Bishwal(2011) studied Berry-Esseen inequalities for the maximum likelihood estimator in Ornstein-Uhlenbeck driven by Gamma process in the ergodic case. Here we study the singular case.

The MLE based on the observations in $[0,T]$ for the Vasicek model

$$dX_t = (\theta_2 - \theta_1 X_t)dt + dV_t^H(\theta_2),\ X_0 = 0$$

are given by

$$\hat{\theta}_{1,T} := \frac{TJ_{1,T} + \int_0^T X_s ds J_{2,T}}{T\int_0^T X_s^2 ds - (\int_0^T X_s ds)^2},$$

$$\hat{\theta}_{2,T} := \frac{\int_0^T X_s ds J_{1,T} + \int_0^T X_s^2 ds J_{2,T}}{T\int_0^T X_s^2 ds - (\int_0^T X_s ds)^2},$$

where

$$J_{1,T} := (\sum_{s\le T}\Delta X_s^2 + T - X_T^2)/2,\quad J_{2,T} := X_T - \sum_{s\le T}\Delta X_s. \tag{4.12}$$

Suppose $V^H(\theta) - X^c(\theta)$ is a gamma process with parameter $1/\theta_3$ under $P_\theta$, with Lebesgue density $\theta_3^{-t}x^{t-1}e^{-x/\theta_3}I_{(0,\infty)}(x)/\Gamma(t)$, Levy measure $L(\theta)$ has the Lebesgue density $x^{-1}e^{-x/\theta_3}I_{(0,\infty)}(x)$, $E(V(\theta)) = \int x dL(\theta_3) = \theta_3, \int x^2 L(\theta_3, dx) = \theta^3$ or a Poisson process with intensity $\theta_3$, $E(V(\theta)) = 1/\theta_3$. Then the MLE of $\theta_3$ is given by

$$\hat{\theta}_{3,T} := T^{-1}\sum_{s\le T}\Delta X_s.$$

This estimator is regular and efficient.

Let $\nu_i(ds,dx) = ds\otimes L_i(dx),\ i = 1,2$. The density of $Y$ depends on $t$ and $K_t(\nu_0,\nu_1) = tK(L_0,L_1)$. where $K(L_0,L_1)$ is the Kullback-Leibler distance between $L_0$ and $L_1$. Let the first passage time be given by

$$\tau_h := \inf\{s\ge 0 : \frac{1}{2}\langle M\rangle_s + sK(L_0,L_1) > h\}. \tag{4.13}$$

Observation of the process $X$ up to time $\tau_h$ corresponds to $\mathcal{F}_{\tau_h}$.

We have

$$\Lambda_t = \widetilde{N}_t + \frac{1}{2}\langle M\rangle_t + K_t(\nu_0,\nu_1),\quad \widetilde{N}_{\tau_h} = M_{\tau_h} + \int_0^{\tau_h}\int \log Y d(\mu - \nu_0). \tag{4.14}$$

The log-likelihood is given by

$$\log\frac{dP_0}{dP_1}|\mathcal{F}_{\tau_h} = \Lambda_{\tau_h} = N_h + h \tag{4.15}$$

where $N_h = \widetilde{N}_{\tau_h}$. The process $N$ is a local square-integrable martingale under the filtration $\mathbb{F}_{\tau_h} = (\mathcal{F}_{\tau_h\wedge t})_{t\ge 0}$.

Let

$$r_1(h) := \frac{1}{2}\langle M\rangle_h + K_h(\nu_0,\nu_1) = K_h(P_0,P_1),\ \ r_2(h) := (\langle M\rangle_h + \int_0^h\int(\log Y)^2 d\nu_0)^{1/2} = (Var\,\Lambda_h)^{1/2}.$$

Assume that
$$\frac{r_1(h)}{r_2(h)} \to \infty.$$

**Theorem 4.1** Assume $\frac{\tau_h}{r_2^2(h)} \to \gamma$ in $P_0$-probability, $\frac{h}{r_2^2(h)} \to \delta$, where $\gamma, \delta \in [0, \infty)$ and $\frac{h}{r_2(h)} \to \infty$. Then
$$\mathcal{L}(r_2^{-1}(h)(\Lambda_{\tau_h} - h)|P_0) \to \mathcal{N}(0, \widetilde{F})$$
where
$$\widetilde{F} := 2\delta + \gamma \left[ \int (\log Y)^2 dL_0 - 2K(L_0, L_1) \right].$$

**Proof.** We have
$$\langle \widetilde{N} \rangle_h = 2 \left[ \frac{1}{2} \langle M \rangle_h + hK(L_0, L_1) \right] + h \left[ \int (\log Y)^2 dL_0 - 2K(L_0, L_1) \right]. \tag{4.16}$$
Therefore,
$$\langle N \rangle_h = \langle \widetilde{N} \rangle_{\tau_h} = 2h + \tau_h \left[ \int (\log Y)^2 dL_0 - 2K(L_0, L_1) \right]. \tag{4.17}$$
This implies
$$\frac{\langle N \rangle_h}{r_2^2(h)} \to^{P_0} \widetilde{F} \tag{4.18}$$
in $P_0$-probability. Furthermore,
$$\frac{\tau_h}{r_2^2(h)} \int (\log Y)^2 I_{\{|\log Y| > r_2(h)\varepsilon\}} dL_0 \to 0 \tag{4.19}$$
in $P_0$-probability. The assertion now follows from (4.15) and CLT for martingales. □

Define the stopping time
$$\tau_h = \inf\{s \geq 0 : \int_0^s X_r^2 dr > h\}.$$

The stopping time is based on the observed Fisher information process exceeding a prescribed level and the asymptotic is for level $\to \infty : \tau(sh) = \tau_h(s)$.

*Asymptotic Dubins-Schwarz-Dambis (DDS) Theorem*: $\sigma_h$ is a deterministic time change such that $h \mapsto \sigma_h(s)$ increases to infinity for every $s$ and let $t \mapsto 1/\delta_h^2$ be continuous and strictly increasing to infinity.

**Case 1** (Ergodic/Subcritical case) : If $\theta_1 > 0$, let $\delta_{1,h} = h^{-1/2}$. If $\theta_1 > 0$, choose $\delta_{2,h} = h^{-1/2}$. Assume $\int x^2 dL(\theta_2) < \infty$ for every $\theta_2$. Then
$$h^{-1} \int_0^{sh} X_r^2 dr \to a_H(\theta)s \quad P_\theta - \text{a.s.} \tag{4.20}$$
for every $s \geq 0$ where
$$a_H(\theta) := \frac{(E_\theta V_1(\theta_1))^2}{\theta_1^{4H}} + \frac{Var_\theta(V_1(\theta_1))H\Gamma(2H)}{\theta_1^{2H}}. \tag{4.21}$$
This implies
$$h^{-1}\tau_h(s) \to \frac{s}{a_H(\theta)} \quad P_\theta - \text{a.s.}$$

for every $s \geq 0$. By the functional limit theorems for martingales, the model is functionally LAN with $\langle M \rangle_s = a_H(\theta)^{-1} \int f'(\theta_2)^2 dL(\theta_2) s$.

**Case 2** (Singular/Critical Case) : If $\theta_1 = 0$, let $\delta_{2,h} = h^{-1/6}$. From the SLLN for Levy processes

$$h^{-1} \int_0^{sh^{1/3}} X_r^2 dr \to \frac{1}{3}(E_\theta V_1(\theta_1))^2 s^3 \quad P_\theta - \text{a.s.} \tag{4.22}$$

for every $s \geq 0$.

Hence

$$h^{-1/3}\tau_h(s) \to (3(E_\theta V_1(\theta_1))^{-2} s)^{1/3} \quad P_\theta - \text{a.s.} \tag{4.23}$$

for every $s \geq 0$. The model is functionally LAN with $\langle M \rangle_s = (3(E_\theta V_1(\theta_1))^{-2})^{1/3} \int f'(\theta_2)^2 dL(\theta_2) s^{1/3}$.

**Case 3** (Nonergodic/Supercritical case) : If $\theta_1 < 0$, choose $\delta_{2,h} = (\log(1+h))^{-1/2}$.

$$e^{2\theta_1 h} \int_0^h X_r^2 dr \to U \quad P_\theta - \text{a.s.} \tag{4.24}$$

where $U$ is a random variable with $U > 0$ a.s., one obtains

$$h^{-1} \int_0^{\varsigma_h(s)} X_r^2 dr \to Us \quad P_\theta - \text{a.s.} \tag{4.25}$$

for every $s \geq 0$ where $\varsigma_h(s) = \varsigma(sh)$ and $\varsigma$ denotes the inverse function of $h \mapsto e^{-2\theta_1 h} - 1$. Hence $\varsigma_h^{-1}(\tau_h(s)) \to s/U$ $P_\theta$-a.s. This implies $-2\theta_1\tau_h(s) - \log(1+h) \to \log(s/U)$ $P_\theta$-a.s. Hence

$$(\log(1+h))^{-1}\tau_h(s) \to -\frac{1}{2\theta_1} \quad P_\theta - \text{a.s.} \tag{4.26}$$

The model is functionally LAN with $\langle M \rangle_s = -\frac{1}{\theta_1} \int f'(\theta_2)^2 dL(\theta_2)$. Further $E_\theta(\tau_h) = O(h)$.

For every $\theta_1 \in \mathbb{R}$ ($\theta_1 > 0$, $\theta_1 = 0$, $\theta_1 < 0$), the model is functionally LAN. The SMLE is asymptotically normal for every $\theta_1 \in \mathbb{R}$.

**Remark** One should obtain the Berry-Esseen bound for the sequential maximum likelihood estimator for the discretely sampled OU process using Shiryayev and Spokoiny (2000). The rate is $O(h^{-1/13})$ where $h$ is the precision.

## 5. Sequential Estimation in Nonlinear Fractional Diffusions

State space transform (SST) of regular diffusions was first studied by Karlin and Taylor (1981, Theorem 2.1, page 172). A standard tool in the theory of diffusions is the concept of scale function (see Itô and McKean (1974) and Revuz and Yor (1991)) which turns one dimensional diffusions into local martingales.

A function $f$ is called a state space transform (SST) if it is continuous and strictly increasing: the open interval $I = f(\mathbb{R})$ is called the state space.

Let $X_t$ be a stationary O-U process satisfying

$$dX_t = -\gamma X_t dt + dW_t, \; t > 0, \; X_0 = \int_{-\infty}^0 e^{\gamma s} dW_s \tag{5.1}$$

with $\gamma > 0$. Let $Y_t := f(X_t)$ where $f$ is a monotone transform which satisfies

$$dY_t = \mu(Y_t)dt + \sigma(Y_t)dW_t. \tag{5.2}$$

Buchmann and Kluppelberg (2006) answered what functions of $\mu$ and $\sigma$ allow for the solution $Y$ to be stationary. Buchmann and Kluppelberg (2005) studied maxima of diffusions using the state space transform.

Let $(I, \mu, \sigma)$ be proper and the friction coefficient (FC) $\gamma > 0$. The distribution of $Y_t$ has a Lebesgue density $p(\cdot)$ where

$$p(\cdot) = \frac{\sqrt{\gamma}}{\sqrt{\pi\sigma(\cdot)^2}} \exp\left[-\frac{1}{\gamma}\left(\frac{\mu(\cdot)}{\sigma(\cdot)}\right)^2\right]. \tag{5.3}$$

Using the SST $f(x) = \sigma x - \alpha/\beta, x \in \mathbb{R}$, for $\beta < 0, \sigma > 0$ and $\alpha \in \mathbb{R}$ and setting $\gamma = -\beta$, the process $f(X_t)$ is a Vasicek process satisfying

$$dV_t = \mu(V_t)dt + \sigma dW_t \tag{5.4}$$

where $\mu(x) = \alpha + \beta x$.

First consider the fCIR model

$$dX_t = a(b - X_t)dt + \sigma\sqrt{X_t}dW_t^H. \tag{5.5}$$

Then by Proposition 5.7 of Buchmann and Kluppelberg (2006), we have

$$X_t = f(Y_t)$$

where

$$dY_t = a(b - Y_t)dt + dW_t^H, \quad Y_0 = f^{-1}(X_0), \quad t \in [0, T] \tag{5.6}$$

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

We use Proposition 5.7 of Buchmann and Kluppelberg (2006) to represent the following nonlinear fractional diffusion as a SST of fOU process. Then we use fundamental semimartingale representation of the fOU model to study sequential estimation.

From Theorem 3.4 in Buchman and Kluppelberg (2006), the fractional diffusion (5.5) above can be represented as a monotone and differentiable functional of the fO-U process using the state space transform (SST) representation. Hence $\widetilde{Z}$ can be represented as a SST of semimartingale in terms of $Z$.

Now we focus on the fundamental semimartingale behind the f-O-U model (5.6). Define

$$\begin{aligned}
\kappa_H &:= 2H\Gamma(3/2 - H)\Gamma(H + 1/2), \quad k_H(t, s) := \kappa_H^{-1}(s(t - s))^{\frac{1}{2} - H}, \\
\eta_H &:= \frac{2H\Gamma(3 - 2H)\Gamma(H + \frac{1}{2})}{\Gamma(3/2 - H)}, \quad v_t \equiv v_t^H := \eta_H^{-1}t^{2-2H}, \quad \mathcal{M}_t^H := \int_0^t k_H(t, s)dM_s^H.
\end{aligned} \tag{5.7}$$

For using Girsanov theorem for Brownian motion, since a Radon-Nikodym derivative process is always a martingale, a central problem is how to construct an appropriate martingale which generates the same filtration, up to sets of measure zero, as the non-semimartingale called the *fundamental martingale*.

Extending Norros *et al.* (1999) it can be shown that $\mathcal{M}_t^H$ is a martingale, called the fundamental martingale whose quadratic variation $\langle \mathcal{M}^H \rangle_t$ is $v_t^H$. Moreover, the natural filtration of the martingale $\mathcal{M}^H$ coincides with the natural filtration of the FLP $M^H$ since

$$M_t^H := \int_0^t K(t,s) d\mathcal{M}_s^H \tag{5.8}$$

holds for $H \in (1/2, 1)$ where

$$K_H(t,s) := H(2H-1) \int_s^t r^{H-\frac{1}{2}} (r-s)^{H-\frac{3}{2}} dr, \quad 0 \le s \le t \tag{5.9}$$

and for $H = 1/2$, the convention $K_{1/2} \equiv 1$ is used.

Define

$$Q_t \equiv Q(t) := \frac{d}{dv_t} \int_0^t k_H(t,s) Y_s ds. \tag{5.10}$$

It is easy to see that

$$Q(t) = \frac{\eta_H}{2(2-2H)} \left\{ t^{2H-1} Z(t) + \int_0^t r^{2H-1} dZ(s) \right\}. \tag{5.11}$$

Define the process $Z = (Z(t), t \in [0,T])$ by

$$Z_t \equiv Z(t) := \int_0^t k_H(t,s) dY(s). \tag{5.12}$$

Extending Kleptsyna and Le Breton (2002), we have:
(i) $Z$ is the fundamental semimartingale associated with the process $Y$.
(ii) $Z$ is a $(\mathcal{F}_t)$ -semimartingale with the decomposition

$$Z(t) = (a - b\int_0^t Q(s) dv_s) + \mathcal{M}_t^H. \tag{5.13}$$

(iii) $X$ admits the representation

$$Y(t) = \int_0^t K_H(t,s) dZ(s). \tag{5.14}$$

(iv) The natural filtration $(\mathcal{Z}(t))$ of $Z$ and $(\mathcal{Y}(t))$ of $Y$ coincide.

We focus on our observations now. Note that for equally spaced data (homoscedastic case)

$$v_{t_k} - v_{t_{k-1}} = \eta_H^{-1} \left(\frac{T}{n}\right)^{2-2H} [k^{2-2H} - (k-1)^{2-2H}], \quad k = 1, 2, \cdots, n. \tag{5.15}$$

For $H = 0.5$,

$$v_{t_k} - v_{t_{k-1}} = \eta_H^{-1} \left(\frac{T}{n}\right)^{2-2H} [k^{2-2H} - (k-1)^{2-2H}] = \frac{T}{n}, \quad k = 1, 2, \ldots, n.$$

We have

$$\begin{aligned} Q(t) &= \frac{d}{dv_t}\int_0^t k_H(t,s)Y(s)ds = \kappa_H^{-1}\frac{d}{dv_t}\int_0^t s^{1/2-H}(t-s)^{1/2-H}Y(s)ds \\ &= \kappa_H^{-1}\eta_H t^{2H-1}\frac{d}{dt}\int_0^t s^{1/2-H}(t-s)^{1/2-H}Y(s)ds \\ &= \kappa_H^{-1}\eta_H t^{2H-1}\int_0^t \frac{d}{dt}s^{1/2-H}(t-s)^{1/2-H}Y(s)ds \\ &= \kappa_H^{-1}\eta_H t^{2H-1}\int_0^t s^{1/2-H}(t-s)^{-1/2-H}Y(s)ds. \end{aligned} \tag{5.16}$$

The process $Q$ depends continuously on $Y$ and therefore, the discrete observations of $Y$ does not allow one to obtain the discrete observations of $Q$. The process $Q_i$ can be approximated by

$$\widetilde{Q}(n) = \kappa_H^{-1}\eta_H n^{2H-1}\sum_{j=0}^{n-1} j^{1/2-H}(n-j)^{-1/2-H}Y(j). \tag{5.17}$$

It is easy to show that $\widetilde{Q}_i(n) \to Q_i(t)$ almost surely as $n \to \infty$, see Tudor and Viens (2007).

Define a new partition $0 \leq r_1 < r_2 < r_3 < \cdots < r_{m_k} = t_k, \quad k = 1, 2, \cdots, n$. Define

$$\widetilde{Q}(t_k) = \kappa_H^{-1}\eta_H t_k^{2H-1}\sum_{j=1}^{m_k} r_j^{1/2-H}(r_{m_k} - r_j)^{-1/2-H}u_i(r_j)(r_j - r_{j-1}), \quad k = 1, 2, \cdots, n. \tag{5.18}$$

It is easy to show that $\widetilde{Q}(t_k) \to Q(t)$ almost surely as $m_k \to \infty$ for each $k = 1, 2, \cdots, n$.

We use this approximate observation in the calculation of our estimators. Thus with $a = 0$ and $b = -\theta$ our observations are

$$Y(t) \approx \int_0^t K_H(t,s)d\widetilde{Z}(s) \quad \text{where} \quad \widetilde{Z}(t) = \theta\int_0^t \widetilde{Q}(s)dv_s + \mathcal{M}_t^H. \tag{5.19}$$

Now we consider the general case. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a *stochastic basis* satisfying the usual hypotheses on which we have a real valued fractional non-Markovian diffusion type process $\{X_t, t \geq 0\}$ satisfying the fractional stochastic differential equation (fSDE)

$$\begin{aligned} dX_t &= f(\theta, t, X)dt + dW_t^H, \quad t \geq 0, \\ X_0 &= \xi \end{aligned} \tag{5.20}$$

where $\{W_t, t \geq 0\}$ is a fractional Wiener process, $\xi$ is a $\mathcal{F}_0$-measurable random variable with $P\{|\xi| < \infty\} = 1$, $\theta \in \Theta$ a compact subset of $\mathbb{R}$ is the unknown parameter to be estimated on the basis of observation of the process $\{X_t, t \geq 0\}$. The measurable function $f(\theta, t, x), t \geq 0, \theta \in \Theta$ and $x \in C$ are assumed to be (for each fixed $\theta$) nonanticipative, that is $\mathcal{B}_t$-measurable for each $t \geq 0$. Here $(C, \mathcal{B})$ is the space of continuous functions $\{x_t, t \geq 0\}$ with $x_0 = \xi$ with the associated Borel $\sigma$-algebra $\mathcal{B}$ and $\mathcal{B}_t = \sigma(X_s, s \geq t)$ are $\sigma$-algebras in the measurable space $(C, \mathcal{B})$.

We are now ready to define the sequential sampling plan $(\tau, \theta_\tau)$ as follows : The stopping time $\tau$ is defined as

$$\tau \equiv \tau_h := \inf\left\{t \geq 0 : \int_0^t f'^2(\theta_s, s, Q)dv_s \geq h\right\} \tag{5.21}$$

where $h > 0$ specifies the desired precision which is predetermined and $\theta_s$ is the MLE based on the observation of $X$ in the time interval $[0, s]$. Note that by Sudakov's lemma, the likelihood based on $[0, \tau]$ is given by

$$L_\tau(\theta) = \frac{dP_\theta^\tau}{dP_W^\tau}(Z_0^\tau) = \exp\left\{\int_0^\tau f(\theta, t, Q)dZ_t - \frac{1}{2}\int_0^\tau f^2(\theta, t, Q)dv_t\right\}. \tag{5.22}$$

(see Liptser and Shiryayev (1977, 1978) for the non-fractional case). The sequential maximum likelihood estimate (SMLE) is defined as

$$\theta_\tau := \arg\max_{\theta\in\Theta} L_\tau(\theta). \tag{5.23}$$

We call the procedure here a two stage estimation procedure since we use an estimated stopping time to define the sequential estimate. One can show that there exists a $\mathcal{F}_\tau$ measurable SMLE since $L_\tau(\theta)$ is continuous in $\theta$ and $\Theta$ is compact. Hereafter we assume the existence of such a measurable SMLE.

Let us introduce the time changed estimator processes: For each $n \in \mathbb{N}$, we define the retimed processes $Z_t^n := n^{-1/2}Z_{nt}$, $Q_t^n := n^{-1/2}Q_{nt}$, and $W_t^n := n^{-1/2}W_{nt}, t \geq 0$ and the filtration $\{\mathcal{F}_t^n\}_{t\geq 0}$ be generated by $Z^n$, $\mathcal{F}_n^t = \{Z_s^n, s \leq t\} = \sigma\{Z_s, s \leq nt\}$. The log-likelihood process at stage $n$ and time $t$ between $\theta_1$ and $\theta$ is given by

$$\Gamma^n_{\theta_1,\theta,t} := \int_0^t [f(\theta_1, s, Q_s^n) - f(\theta, s, Q_s^n)]dZ_s^n - \frac{1}{2}\int_0^t [f(\theta_1, s, Q_s^n) - f(\theta, s, Q_s^n)]dv_s. \tag{5.24}$$

Prime denotes derivative with respect to $\theta$ throughout the paper. We assume the following conditions in this section:

(A1) $\int_0^T f^2(\theta, t, x)dv_t < \infty,\ T < \infty,\ x \in C[0, T],\ \theta \in \mathbb{R}.$

(A2) $P_\theta\left(\int_0^\infty f'^2(\theta, t, Q)dv_t = \infty\right) = 1,\ \theta \in \mathbb{R}.$

(A3)

$$|f(\theta, t, x) - f(\theta, t, y)| \leq R_1\int_0^t |x_s - y_s|ds + R_2|x_t - y_t|,$$

$$f^2(\theta, t, x) \leq R_1\int_0^t (1 + |x_s|)ds + R_2(1 + |x_t|)$$

where $x_s, y_s \in C[0, T],\ \theta \in \mathbb{R},\ R_1$ and $R_2$ are constants.

(B1) (Identifiability condition) $P_{\theta_1}^\tau \neq P_{\theta_2}^\tau$ for $\theta_1 \neq \theta_2$ in $\Theta$.

(B2) $l_\tau(\theta)$ is twice continuously differentiable in a neighborhood $U_\theta$ of $\theta$ for every $\theta \in \Theta$.

(B3) $\lim_{h\to\infty}\frac{1}{h}\int_0^\tau f''(\theta, t, Q)d\mathcal{M}_t = 0$ in $P_\theta -$ probability.

There exist $r_n \uparrow \infty$ as $n \uparrow \infty$ and $\nu > 0$ such that

(C1) $\frac{1}{r_n^2}\int_0^\nu f''^2(\theta, s, Q_s^n)dv_s \xrightarrow{P_\theta^n} \zeta_\nu(\theta)$ as $n \to \infty$ where $P_\theta^n[\zeta_\nu(\theta) > 0] > 0.$

(C2) $\frac{1}{r_n^2}\int_0^\nu f'(\theta,s,Q_s^n)dv_s \xrightarrow{P_\theta^n} \xi_\nu(\theta)$ as $n\to\infty$ where $P_\theta^n[\xi_\nu(\theta)>0]>0$.

(C3) $\frac{1}{r_n^2}\int_0^\nu f''(\theta,s,Q_s^n)dZ_s^n \xrightarrow{P_\theta^n} 0$ as $n\to\infty$.

**Lemma 5.1**(a) Let $A_t^\theta := f(\theta\pm\delta,t,Q)-f(\theta,t,Q)$ for some $\delta>0$. Then under (A1) – (A3), we have

$$\frac{\int_0^\tau (A_t^\theta)^2 dv_t}{\int_0^\tau (A_t^{\theta_t})^2 dv_t}\to 1 \text{ a.s. } [P_\theta^\tau] \text{ as } h\to\infty.$$

Under the assumptions (A1) – (A3) and (B1) – (B3), we have

(b) There exists a root of the likelihood equation which is strongly consistent, i.e.,

$$\lim_{h\to\infty}\theta_\tau=\theta \text{ a.s. } [P_\theta].$$

(c) $\sqrt{h}(\theta_\tau-\theta)\xrightarrow{\mathcal{D}[P_\theta]}\mathcal{N}(0,1)$ as $h\to\infty$ uniformly in $\theta\in\Theta$.

**Proof.** (a) We have $A_t^\theta = f(\theta\pm\delta,t,Q)-f(\theta,t,Q)$, $A_t^{\theta_t}=f(\theta_t\pm\delta,t,Q)-f(\theta_t,t,Q)$. Thus $(A_t^\theta)^2-(A_t^{\theta_t})^2=[f(\theta\pm\delta,t,Q)-f(\theta,t,Q)]^2-[f(\theta_t\pm\delta,t,Q)-f(\theta_t,t,Q)]^2$. Since $\theta_t$ is a strongly consistent estimator of $\theta$, $\theta_t\to\theta$ a.s. as $t\to\infty$. Since $f$ is continuous, $(A_t^\theta)^2-(A_t^{\theta_t})^2\to 0$ a.s. as $t\to\infty$. Further

$$\int_0^\tau (A_t^{\theta_t})^2 dt\to c(\theta) \text{ a.s. as } h\to\infty \text{ and } \delta\ \to 0 \tag{5.25}$$

where $c(\theta)$ is a positive constant and

$$\int_0^\tau [(A_t^\theta)^2-(A_t^{\theta_t})^2]dv_t\to 0 \text{ a.s. as } h\to\infty. \tag{5.26}$$

Thus

$$\frac{\int_0^\tau [(A_t^\theta)^2-(A_t^{\theta_t})^2]dv_t}{\int_0^\tau (A_t^{\theta_t})^2 dv_t}\to 0 \text{ a.s. } [P_\theta^\tau] \text{ as } h\to\infty. \tag{5.27}$$

(b) Observe that, for $\delta>0$

$$\begin{aligned}
&l_\tau(\theta\pm\delta)-l_\tau(\theta)=\log\frac{dP_{\theta\pm\delta}^\tau}{dP_\theta^\tau}\\
=&\int_0^\tau [f(\theta\pm\delta,t,Q)-f(\theta,t,Q)]dZ_t-\frac12\int_0^\tau [f^2(\theta\pm\delta,t,Q)-f^2(\theta,t,Q)]dv_t\\
=&\int_0^\tau [f(\theta\pm\delta,t,Q)-f(\theta,t,Q)]dM_t-\frac12\int_0^\tau [f(\theta\pm\delta,t,Q)-f(\theta,t,Q)]^2dv_t\\
=&\int_0^\tau A_t^\theta dM_t-\frac12\int_0^\tau (A_t^\theta)^2dv_t.
\end{aligned} \tag{5.28}$$

Let $K_\tau := \int_0^\tau (A_t^\theta)^2 dv_t$. Then

$$\frac{l_\tau(\theta\pm\delta)-l_\tau(\theta)}{K_\tau} = \frac{\int_0^\tau A_t^\theta d\mathcal{M}_t}{\int_0^\tau (A_t^\theta)^2dv_t}-\frac12=\frac{W^*(\int_0^\tau (A_t^\theta)^2dv_t)}{\int_0^\tau (A_t^\theta)^2dv_t}-\frac12=\frac{W^*(K_\tau)}{K_\tau}-\frac12 \tag{5.29}$$

by the Skorohod embedding of the martingale $\int_0^\tau A_t^\theta d\mathcal{M}_t$ where $W^*$ is some other Brownian motion which is independent of $K_\tau$.

Using the assumption (A2) and Lemma 5.1 (a), and the strong law of large numbers for Brownian motion (see Liptser and Shiryayev (1978)) $\frac{W^*(K_\tau)}{K_\tau}$ converges to zero a.s. as $h \to \infty$. Hence,

$$\frac{l_\tau(\theta \pm \delta) - l_\tau(\theta)}{K_\tau} \to -\frac{1}{2} \text{ a.s. } [P_\theta] \text{ as } h \to \infty. \tag{5.30}$$

Furthermore, $K_\tau > 0$ a.s. $[P_\theta]$ by (B1). Therefore, for almost every $w \in \Omega, \delta$ and $\theta$, there exist some $H_0$ such that for $h \geq h_0$, we have $l_\tau(\theta \pm \delta) < l_\tau(\theta)$. Since $l_\tau(\theta)$ is continuous on the compact set $[\theta - \delta, \theta + \delta]$, it has a local maximum and it is attained at a measurable $\theta_\tau$ in $[\theta - \delta, \theta + \delta]$. Since $l_\tau(\theta \pm \delta) < l_\tau(\theta)$, $\theta_\tau \in (\theta - \delta, \theta + \delta)$ for $h > h_0$. Since $l_\tau(\theta)$ is differentiable with respect to $\theta$, it follows that $l'_\tau(\theta_\tau) = 0$ for $h \geq h_0$ and $\theta_\tau \to \theta$ a.s. as $h \to \infty$.

(c) In view of the assumption (B3), we can apply Taylor's expansion, for $l'_\tau(\theta)$ around $\theta_\tau$ and write

$$0 = l'_\tau(\theta_\tau) = l'_\tau(\theta) + (\theta_\tau - \theta) l_\tau''(\theta + \beta_\tau(\theta_\tau - \theta)) \tag{5.31}$$

where $|\beta_\tau| \leq 1$ a.s. for sufficiently large $h$. Since $\theta_\tau \to \theta$ a.s. as $h \to \infty$ by (b) and since $l''_\tau(\theta)$ is continuous by (B3), it follows that $l''_\tau(\theta + \beta_\tau(\theta_\tau - \theta)) - l_\tau''(\theta) \to 0$ in $P^\tau_\theta$ - probability as $h \to \infty$. Hence $l'_\tau(\theta) + (\theta_\tau - \theta) l_\tau''(\theta) \to 0$ in $P^\tau_\theta$-probability as $h \to \infty$. We have $l'_\tau(\theta) = \int_0^\tau f'(\theta, t, Q) d\mathcal{M}_t$, see Karandikar (1983). Hence using the central limit theorem for stochastic integrals (see Basawa and Prakasa Rao (1980)) and Lemma 5.1(a), we obtain

$$\frac{l'_\tau(\theta)}{\sqrt{h}} \stackrel{\mathcal{D}[P_\theta]}{\to} \mathcal{N}(0,1) \text{ as } h \to \infty. \tag{5.32}$$

Note that when $\theta$ is the true parameter

$$\begin{aligned} l''_\tau(\theta) &= \int_0^\tau f''(\theta, t, Q) dZ_t - \int_0^\tau [f(\theta, t, Q) f\,''(\theta, t, Q) + f'^2(\theta, t, Q)] dv_t \\ &= \int_0^\tau f''(\theta, t, Q) d\mathcal{M}_t - \int_0^\tau f'^2(\theta, t, Q) dv_t. \end{aligned} \tag{5.33}$$

By Lemma 5.1(a) and (B3) it follows that

$$\frac{l''_\tau(\theta)}{h} \to -1 \text{ in } P_\theta\text{-probability as } h \to \infty. \tag{5.34}$$

Hence it follows that

$$\sqrt{h}(\theta_\tau - \theta) \stackrel{\mathcal{D}[P_\theta]}{\to} \mathcal{N}(0,1) \text{ as } h \to \infty.$$

This completes the proof of the lemma. □

**Theorem 5.2** Under the assumptions (C1) – (C3) the sequence of filtered models $(\Omega^n, \mathcal{F}^n, \{\mathcal{F}^n_t\}_{t \geq 0},$
$\{P^n_\theta, \theta \in \Theta\})$ generated by $Z^n_t$ satisfy the functionally LAQ condition at $\theta$ with

$$\Delta_t = \int_0^t f'(\theta, s, Q_s) dZ_s \text{ and } \Gamma_t = \int_0^t f'^2(\theta, s, Q_s) dv_s.$$

**Proof.** For $\nu \leq t$,

$$\Lambda^n_{\theta + r_n^{-1} u_n, \theta, \nu} = \int_0^\nu [f(\theta + r_n^{-1} u_n, s, Q^n_s) - f(\theta, s, Q^n_s)] dZ^n_s - \frac{1}{2} \int_0^\nu [f(\theta + r_n^{-1} u_n, s, Q^n_s) - f(\theta, s, Q^n_s)]^2 dv_s. \tag{5.35}$$

By Taylor's formula

$$f(\theta + r_n^{-1}u_n, s, Q_s^n) - f(\theta, s, Q_s^n) = r_n^{-1}u_n f'(\theta, s, Q_s^n) + \frac{1}{2}r_n^{-2}u_n^2 f''(\bar{\theta}, s, Q_s^n) \tag{5.36}$$

where

$$\bar{\theta} := \theta + q(s, Q_s^n)r_n^{-1}u_n, \quad |q(\cdot,\cdot)| < 1. \tag{5.37}$$

Hence

$$\begin{aligned}
\Lambda^n_{\theta+r_n^{-1}u_n,\theta,\nu} &= r_n^{-1}u_n \int_0^\nu f'(\theta, s, Q_s^n)dZ_s^n + \frac{1}{2}r_n^{-2}u_n^2 \int_0^\nu f''(\bar{\theta}, s, Q_s^n)dZ_s^n \\
&\quad -\frac{1}{2}r_n^{-2}u_n^2 \int_0^\nu f'^2(\theta, s, Q_s^n)dv_s - \frac{1}{8}r_n^{-4}u_n^4 \int_0^\nu f''^2(\bar{\theta}, s, Q_s^n)dv_s \\
&\quad -\frac{1}{8}r_n^{-3}u_n^3 \int_0^\nu f'(\theta, s, Q_s^n)f''(\bar{\theta}, s, Q_s^n)dv_s \\
&= r_n^{-1}u_n \int_0^\nu f'(\theta, s, Q_s^n)dZ_s^n - \frac{1}{2}r_n^{-2}u_n^2 \int_0^\nu f'^2(\theta, s, Q_s^n)dv_s + o_{P_\theta^n}(1) \\
&\qquad \text{(by assumption (C1) – (C3))} \\
&= u_n\Delta_\nu^n - \frac{1}{2}u_n^2\Gamma_\nu^n + o_{P_\theta^n}(1)
\end{aligned} \tag{5.38}$$

where

$$\Delta_\nu^n := r_n^{-2}\int_0^\nu f'(\theta, s, Q_s^n)dZ_s^n \text{ and } \Gamma_\nu^n := r_n^{-2}\int_0^\nu f'^2(\theta, s, Q_s^n)dv_s. \tag{5.39}$$

Let

$$\Delta_\nu := \int_0^\nu f'(\theta, s, Q_s)dZ_s \text{ and } \Gamma_\nu := \int_0^\nu f'^2(\theta, s, Q_s)dv_s. \tag{5.40}$$

By the functional CLT for martingales and stability of weak convergence (see Jacod and Shiryayev (1987)), we obtain

$$(\Delta_\nu^n, \Gamma_\nu^n) \overset{\mathcal{D}[P_\theta^n]}{\to} (\Delta_\nu, \Gamma_\nu) \text{ as } n \to \infty. \tag{5.41}$$

Here $\Delta_\nu$ and $\Gamma_\nu$ are processes on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ with $\Gamma_\nu > 0$ a.s. and

$$E_P \exp(u\Delta_\nu - \frac{1}{2}u^2\Gamma_\nu) = 1 \text{ for all } \nu \geq 0. \tag{5.42}$$

Thus the sequence of filtered models are functionally LAQ at $\theta$. □

We now state the Hajek's local asymptotic minimax (LAM) theorem for SMLE processes.

**Theorem 5.3** Let $\omega$ be a bounded, symmetric, bowl shaped loss function on the real line. Let $\theta_t^n$ be a sequence of MLE processes with respect to the filtered model $(\Omega^n, \mathcal{F}^n, \{\mathcal{F}_t^n\}_{t\geq 0}, \{P_\theta, \theta \in \Theta\})$. For each $n \in \mathbb{N}$ and $t \geq 0$, let

$$\tau_h^n := \inf\{t \geq 0 : r_n^{-2}\int_0^t f'^2(\theta_s^n, s, Q_s^n)dv_s \geq h\}. \tag{5.43}$$

Let conditions (C1) – (C3) be satisfied. Then

$$\lim_{v\to\infty} \liminf_{n\to\infty} \sup_{|u|\leq v} E_{\theta+r_n^{-1}u}\omega(\theta^n_{\tau_h^n} - (\theta + r_n^{-1}u)) \geq E\omega(W_{1/h}). \tag{5.44}$$

**Proof.** Since by Theorem 5.2 the sequence of filtered models generated by $Z_n^t$ is functionally LAQ at $\theta$ with $\Delta_t$ a continuous local martingale, $\Gamma_t$ equal to the quadratic variation $\langle\Delta\rangle_t$ of $\Delta_t$, and $\Gamma_t \uparrow \infty$ a.s. hence the theorem follows from Theorem 5.1.

Note that the filtered model time-changed by $\tau_h^n$ is functionally LAN i.e., for each bounded sequence of numbers $u_n$ and all $h > 0$

$$\sup_{g\le h}\left|\Lambda^n_{\theta+r_n^{-1}u_n,\theta,\tau_g^n} - \left(u_n\Delta^n_{\tau_g^n} - \frac{1}{2}u_n^2 g\right)\right| = o_{P_{\theta_0^n}}(1) \tag{5.45}$$

and

$$\Delta^n_{\tau^n} \stackrel{\mathcal{D}[P_\theta^n]}{\to} W_1 \tag{5.46}$$

by Lemma 4.1 in Bishwal (2018). Hence by Theorem 5.1 we obtain the local asymptotic minimax theorem. □

**Theorem 5.4** Under the conditions (B1) – (B3) and (C1) – (C3) the sequence of estimators $\theta^n_{\tau_h^n}$ are locally asymptotically minimax (LAM), i.e., they attain the lower bound in Theorem 5.3.

**Proof** : Here we have only to show that $\theta^n_{\tau_h^n}$ are asymptotically centering (AC), that is

$$r_n(\theta^n_{\tau_h^n} - \theta) - h^{-1}\Delta^n_{\tau_h^n} = o_{P_\theta^n}(1) \ \text{ as } n\to\infty. \tag{5.47}$$

Note that

$$\begin{aligned} r_n(\theta^n_{\tau_h^n} - \theta) &= \frac{-r_n^{-1}l'_{\tau_h^n}(\theta)}{\gamma_n^{-2}l''_{\tau_h^n}(\theta+\beta_{\tau_h^n}(\theta_{\tau_h^n}-\theta))} \\ &\simeq \frac{-r_n^{-1}\int_0^{\tau_h^n} f'(\theta,s,Q_s^n)dZ_s^n}{r_n^{-2}\int_0^{\tau_h^n} f''(\theta,s,Q_s^n)dZ_s^n - r_n^{-2}\int_0^{\tau_h^n} f'^2(\theta,s,Q_s^n)dv_s} \end{aligned} \tag{5.48}$$

by the arguments similar to the proof of Lemma 5.1(c). On the other hand,

$$h^{-1}\Delta^n_{\tau_h^n} = \frac{r_n^{-1}\int_0^{\tau_h^n} f'(\theta,s,Q_s^n)dZ_s^n}{r_n^{-2}\int_0 f'^2(\theta_s^n,s,Q_s^n)dv_s}. \tag{5.49}$$

From (5.48) and (5.49) using (C3) and Lemma 5.1(a), one observes that (5.47) holds. □

## Concluding Remarks

1. Alternatively one can observe $m$ independent discretely observed trajectories of $X$ and let $m\to\infty$.

2. Density of $\tau_h$: The density of the first passage time is obtained in Bibbona and Ditlevsen (2013).

3. It would be interesting to study the properties of sequential Bayes estimators for nonlinear fractional diffusion models. Sequential Bayes estimation for exponential type processes was studied by Franz and Magiera (1990).

## References


[1] L. Alili, P. Patie, J.L. Pedersen, Representations of first hitting time density of an Ornstein-Uhlenbeck process, Stoch. Models 21 (2005), 967–980.

[2] F.J. Anscombe, Large sample theory of sequential estimation, Proc. Camb. Philos. Soc. 48 (1952), 600–607.

[3] I.V. Basawa, B.L.S. Prakasa Rao, Statistical inference for stochastic processes, Academic Press, New York–London, 1980.

[4] E. Bibbona, S. Ditlevsen, Estimation in discretely observed diffusions killed at a threshold, Scand. J. Stat. 40 (2013), 274–293.

[5] J.P.N. Bishwal, Bayes and sequential estimation in Hilbert space valued stochastic differential equations, J. Korean Stat. Soc. 28 (1999), 93–106.

[6] J.P.N. Bishwal, Sequential maximum likelihood estimation in semimartingales, J. Stat. Appl. 1 (2006), 15–25.

[7] J.P.N. Bishwal, Parameter estimation in stochastic differential equations, Lecture Notes in Mathematics 1923, Springer-Verlag, 2008.

[8] J.P.N. Bishwal, Berry-Esseen inequalities for the discretely observed Ornstein-Uhlenbeck-Gamma process, Markov Process. Relat. Fields 17 (2011), 119–150.

[9] J.P.N. Bishwal, Valuation of real options under persistent shocks, J. Stat. Manag. Syst. 20 (2017), 801–815.

[10] J.P.N. Bishwal, Sequential maximum likelihood estimation in nonlinear non-Markov diffusion type processes, Dyn. Syst. Appl. 27 (2018), 107–124.

[11] J.P.N. Bishwal, Parameter estimation in stochastic volatility models, Springer Nature, Cham, 2022.

[12] V. Borkar, A. Bagchi, Parameter estimation in continuous time stochastic processes, Stochastics 8 (1982), 193–212.

[13] B.M. Brown, J.T. Hewitt, Inference for the diffusion branching process, J. Appl. Probab. 12 (1975), 588–594.

[14] B. Buchmann, C. Kluppelberg, Maxima of stochastic processes driven by fractional Brownian motion, Adv. Appl. Probab. 37 (2005), 743–764.

[15] B. Buchmann, C. Kluppelberg, Fractional integral equations and state space transform, Bernoulli 12 (2006), 431–456.

[16] J.C. Cox, Notes on option pricing I: Constant elasticity of diffusions, J. Portf. Manag. 23 (1996), 15–17.

[17] O. Dehtiar, Y. Mishura, K. Ralchenko, Two methods of estimation of the drift parameters of the Cox-Ingersoll-Ross process; Continuous observations, Commun. Stat. Theory Methods 50 (2021), 1–16.

[18] H.M. Dietz, Asymptotic properties of maximum likelihood estimators in diffusion type models, Preprint No. 228, Humboldt-Universität zu Berlin, Sektion Mathematik, 1989.

[19] H.M. Dietz, A non-Markovian relative to the Ornstein-Uhlenbeck process and some of its local statistical properties, Scand. J. Stat. 19 (1992), 363–379.

[20] S. Ditlevsen, A. Samson, Estimation in the partially observed stochastic Morris-Lecar neuronal model with particle filter and stochastic approximation method, Ann. Appl. Stat. 8 (2014), 674–702.

[21] A.A. Dmitrienko, V.V. Konev, On guaranteed estimation of autoregressive parameter for an unknown variance of noise, Autom. Remote Control 55 (1994), 218–228.

[22] A. Dmitrienko, V. Konev, V. Pergamenshchikov, Sequential generalised least squares estimator for an autoregressive parameter, Seq. Anal. 16 (1997), 25–46.

[23] J. Franz, R. Magiera, Admissible estimators in sequential plans for exponential type processes, Scand. J. Stat. 17 (1990), 275–285.

[24] M. Ghosh, N. Mukhopadhyaya, P.K. Sen, Sequential estimation, John Wiley, New York, 1997.

[25] A. Going-Jaeschke, M. Yor, A survey and some generalizations of Bessel processes, Bernoulli 9 (2003), 313–349.

[26] P. Grambsch, Sequential sampling based on the observed Fisher information to guarantee the accuracy of the maximum likelihood estimator, Ann. Stat. 11 (1983), 68–77.

[27] P.E. Greenwood, A.N. Shiryayev, Asymptotic minimaxity of a sequential estimator for a first order autoregressive model, Stoch. Stoch. Rep. 38 (1992), 49–65.

[28] P.E. Greenwood, W. Welfelmeyer, Asymptotic minimax results for stochastic process families with critical points, Stoch. Process. Appl. 44 (1993), 107–116.

[29] A.A. Gushchin, On the asymptotic optimality of estimators of parameters under the LAQ condition, Theory Probab. Appl. 40 (1995), 261–272.

[30] A.A. Gushchin, U. Küchler, Asymptotic inference for a linear stochastic differential equation with time delay, Bernoulli 5 (1998), 1059–1098.

[31] P. Hall, C.C. Heyde, Martingale limit theory and its applications, Academic Press, New York, 1980.

[32] R. Höpfner, On statistics of Markov processes: representation of log-likelihood ratio processes in filtered local models, Probab. Theory Relat. Fields 94 (1993), 375–398.

[33] R. Höpfner, Asymptotic inference for Markov step processes: observation up to a random time, Stoch. Process. Appl. 48 (1993), 295–310.

[34] K. Itô, H.P. McKean Jr., Diffusion processes and their sample paths, Springer-Verlag, Berlin, 1974.

[35] J. Jacod, A.N. Shiryayev, Limit theorems for stochastic processes, Springer, Berlin, 1987.

[36] P. Jeganathan, On the asymptotic theory of statistical inference when the limit of the log-likelihood ratios is mixed normal, Sankhyā Ser. A 44 (1982), 173–212.

[37] P. Jeganathan, Some aspects of asymptotic theory with applications to time series models, Econom. Theory 11 (1995), 818–887.

[38] R.L. Karandikar, Interchanging the order of stochastic integration and ordinary differentiation, Sankhyā Ser. A 45 (1983), 120–124.

[39] S. Karlin, H.M. Taylor, A second course in stochastic processes, Academic Press, New York, 1981.

[40] M.L. Kleptsyna, A. Le Breton, Statistical analysis of the fractional Ornstein-Uhlenbeck type process, Stat. Inference Stoch. Process. 5 (2002), 229–248.

[41] U. Küchler, M. Sørensen, Exponential family of stochastic processes, Springer, New York, 1997.

[42] Yu.A. Kutoyants, Parameter estimation for diffusion type processes of observations, Math. Operationsforsch. Stat. Ser. Stat. 15 (1984), 541–551.

[43] Yu.A. Kutoyants, Parameter estimation for stochastic processes, Heldermann-Verlag, Berlin, 1984.

[44] Yu.A. Kutoyants, Identification of dynamical systems with small noise, Kluwer, Dordrecht, 1994.

[45] Yu.A. Kutoyants, Statistical inference for ergodic diffusion processes, Springer-Verlag, New York–Berlin, 2003.

[46] T.L. Lai, D. Siegmund, Fixed accuracy estimation of an autoregressive parameter, Ann. Stat. 11 (1985), 478–485.

[47] A. Le Breton, M. Musiela, On sequential estimation of parameters of continuous Gaussian Markov processes, Probab. Math. Stat. 2 (1981), 37–53.

[48] A. Le Breton, M. Musiela, Some parameter estimation problems for hypoelliptic homogeneous Gaussian diffusions, Sequential Methods in Statistics, Banach Centre Publ. 16 (1985), 337–356.

[49] C. Lee, J.P.N. Bishwal, M.H. Lee, Sequential maximum likelihood estimation for reflected Ornstein-Uhlenbeck process, J. Stat. Plann. Inference 142 (2012), 1234–1242.

[50] A. Lenkasas, V. Mackevicius, A second order weak approximation of Heston model by discrete random variables, Lith. Math. J. 55 (2015), 555–572.

[51] G. Lileika, V. Mackevicius, Weak approximation of CKLS and CEV process by discrete random variables, Lith. Math. J. 60 (2020), 208–224.

[52] L. Li, V. Linetsky, Optimal stopping and early exercise: an eigenfunction expansion approach, Oper. Res. 61 (2013), 625–643.

[53] L. Li, V. Linetsky, Time-changed Ornstein-Uhlenbeck processes and their applications in commodity derivative models, Math. Finance 24 (2014), 289–330.

[54] L. Li, V. Linetsky, Discretely monitored first passage problems and barrier options: an eigenfunction expansion approach, Finance Stoch. 19 (2015), 941–977.

[55] V. Linetsky, Computing hitting time densities for CIR and OU diffusions: application to mean reverting models, J. Comput. Finance 7 (2004), 1–22.

[56] A. Lipton, V. Kaushansky, On the first hitting time density for a reducible diffusion process, Quant. Finance 20 (2020), 723–743.

[57] R.S. Liptser, A.N. Shiryayev, Statistics of random processes I, II, Springer, New York, 1977–1978.

[58] L. Le Cam, G.L. Yang, Asymptotics in statistics, Springer, New York, 1990.

[59] D. Levanony, A. Shwartz, O. Zeitouni, Recursive identification in continuous-time stochastic processes, Stoch. Process. Appl. 49 (1994), 245–275.

[60] E. Lőcherbach, Likelihood ratio processes for Markovian particle systems with killing and jumps, Stat. Inference Stoch. Process. 5 (2002), 153–177.

[61] E. Lőcherbach, LAN and LAMN for systems of interacting diffusions with branching and immigration, Ann. Inst. Henri Poincaré 38 (2002), 59–2002.

[62] V. Mackevicius, Weak approximation of CIR equation by discrete random variables, Lith. Math. J. 51 (2011), 385–401.

[63] V. Mackevicius, Verhulst versus CIR, Lith. Math. J. 55 (2015), 119–133.

[64] V.K. Malinovskii, Asymptotic expansions in sequential estimation of an autoregressive parameter, Math. Methods Stat. 2 (1993), 206–227.

[65] A.V. Melnikov, A.A. Novikov, Sequential inferences with prescribed accuracy for semimartingales, Theory Probab. Appl. 33 (1988), 446–459.

[66] Y. Mishura, A. Yurchenko-Tytarenko, Fractional Cox-Ingersoll-Ross model with non-zero mean, Mod. Stoch. Theory Methods 5 (2018), 99–111.

[67] M. Musiela, Sequential estimation of parameters of a stochastic differential equation, Math. Operationsforsch. Stat. Ser. Stat. 8 (1977), 483–498.

[68] M. Musiela, Admissible and minimax estimation for two classes of stochastic processes, Bull. Acad. Polon. Sci. Sér. Sci. Math. 27 (1979), 511–514.

[69] A.A. Novikov, Sequential estimation of the parameters of diffusion type processes, Math. Notes 12 (1972), 812–818.

[70] L. Overbeck, Estimation for continuous branching processes, Scand. J. Stat. 25 (1998), 111–126.

[71] L. Overbeck, T. Ryden, Estimation in the Cox-Ingersoll-Ross model, Econom. Theory 13 (1997), 430–461.

[72] S.M. Pergamenshchikov, Asymptotic properties of sequential design for estimating the parameter of a first order autoregression, Theory Probab. Appl. 36 (1991), 36–49.

[73] D. Revuz, M. Yor, Continuous martingales and Brownian motion, Springer-Verlag, 1991.

[74] L.M. Ricciardi, S. Sato, First-passage-time density and moments of the Ornstein-Uhlenbeck process, J. Appl. Probab. 25 (1988), 43–57.

[75] R. Rozanskii, Markov stopping sets and stochastic integrals – Application in sequential estimation for a random diffusion field, Stoch. Process. Appl. 32 (1989), 237–252.

[76] A.N. Shiryayev, V.G. Spokoiny, Statistical experiments and decision: Asymptotic theory, World Scientific, Singapore, 2000.

[77] M. Sørensen, On maximum likelihood estimation in randomly stopped diffusion-type processes, Int. Stat. Rev. 51 (1983), 93–110.

[78] M. Sørensen, On sequential maximum likelihood estimation for exponential families of stochastic processes, Int. Stat. Rev. 54 (1986), 191–210.

[79] M.S. Tikhov, On optimal sequential estimation plans for nonquadratic loss functions, Theory Probab. Appl. 23 (1978), 132–138.

[80] M.S. Tikhov, On the sequential estimation of the drift coefficient of a diffusion process with quadratic and nonquadratic losses, Theory Probab. Appl. 25 (1980), 607–614.

[81] C.A. Tudor, F. Viens, Statistical aspects of the fractional stochastic calculus, Ann. Stat. 35 (2007), 1183–1212.