

$\epsilon \rightarrow 0$ when the model satisfies the LAN condition. Linkov [35] obtained the LAN property of such models as observation time $T \rightarrow \infty$. Dietz [6] studied asymptotic properties of MLE as the intensity of noise $\epsilon \rightarrow 0$ or the observation time $T \rightarrow \infty$ when the model satisfies the LAN condition. Dietz [7] 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 $T \rightarrow \infty$. For a linear stochastic differential equation with time delay, the MLE shows eleven different behaviors for eleven parts of the parameter space [16].

However, in addition to asymptotic theory which certainly play a predominant role in statistical theory sequential asymptotics has certain advantages. Here we study the asymptotic properties of a sequential estimate of the parameter appearing in the drift coefficient of nonlinear nonhomogeneous diffusion type process, a non-Markovian process, when the solution process is observed continuously up to a random stopping time. In particular, we study the properties of sequential maximum likelihood estimate (SMLE), i.e., the maximum likelihood estimate based on observation of the process on the random time interval $[0, \tau]$ where τ is the stopping time which is the first passage time until the observed (stochastic) Fisher information of the process reaches a preassigned level of precision H . Under some regularity conditions, the SMLE is shown to be strongly consistent, asymptotically normal and locally asymptotically minimax in the Hajek-LeCam sense as $H \rightarrow \infty$.

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 [42] (see also [32]) proved that the SMLE of θ 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 [47] 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 [48] generalised this to the case $b(\theta, t, x) = \theta b_1(t, x) + b_0(t, x)$. Sørensen [45] gave a review of sequential maximum likelihood estimation in linearly parametrized diffusion type processes. Sørensen [46] (see also [24]) studied similar properties of SMLE for exponential families of stochastic processes. Musiela [40, 41] studied sequential ML estimation in a linear diffusion model. Le Breton and Musiela [31] studied similar properties of SMLE in linear homogeneous multidimensional SDE. Rozanskii [44] extended the work of Novikov [42] to a linear homogeneous diffusion field. Brown and Hewitt [5] studied the properties of sequential maximum likelihood estimator for diffusion branching process using a different type of stopping rule.

2. MODEL, ASSUMPTIONS AND THE SEQUENTIAL PLAN

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 non-Markovian diffusion type process $\{X_t, t \geq 0\}$ satisfying the functional stochastic differential equation (SDE)

$$\begin{aligned} dX_t &= f(\theta, t, X)dt + dW_t, \quad t \geq 0, \\ X_0 &= \xi \end{aligned} \tag{2.1}$$

where $\{W_t, t \geq 0\}$ is a standard Wiener process, ξ 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 θ) 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 σ -algebra \mathcal{B} and $\mathcal{B}_t = \sigma(X_s, s \geq t)$ are σ -algebras in the measurable space (C, \mathcal{B}) .

Prime denotes derivative with respect to θ throughout the paper. We assume the following conditions.

$$(A1) \int_0^T f^2(\theta, t, x)dt < \infty, \quad T < \infty, \quad x \in C, \quad \theta \in \mathbb{R},$$

$$(A2) P\left(\int_0^\infty f'^2(\theta, t, X)dt = \infty\right) = 1, \quad \theta \in \mathbb{R},$$

$$(A3) |f(\theta, t, x) - f(\theta, t, y)| \leq M_1 \int_0^t |x_s - y_s|ds + M_2|x_t - y_t|,$$

$$f^2(\theta, t, x) \leq M_1 \int_0^t (1 + |x_s|)ds + M_2(1 + |x_t|),$$

where $x_s, y_s \in C, \theta \in \mathbb{R}, M_1$ and M_2 are constants.

Under the condition (A3), it is well known that equation (2.1) has a unique solution (see [24]).

We estimate the unknown parameter θ from the observation of the process $\{X_t, t \geq 0\}$ 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 [1] (see also [11, 12]). This type of stopping rule was used for autoregressive parameter estimation in [8, 9, 29, 38, 43] among others.

Let $\tau = \tau(X)$ be a Markov stopping time with respect to the system $\{\mathcal{B}_t\}_{t \geq 0}$ and $\delta(t, X)$ be a progressively measurable real process defined on (C, \mathcal{B}) . The pair of functions $D = D(\tau, \delta)$, where $\tau(X)$ is a Markov time and $\delta = \delta(\tau(X), X)$ is an estimate of θ resulting from the observation of the trajectory of X on the time interval $[0, \tau(X)]$ yields a sequential sampling plan.

Let P_θ^T be the measure generated by the process $\{X_t, 0 \leq t \leq T\}$ on the space (C_T, \mathcal{B}_T) of continuous functions from $[0, T] \rightarrow \mathbb{R}$ with the associated Borel σ -algebra

\mathcal{B}_T under the supremum norm and P_W^T be the measure generated by $\xi + W_t$ on the space (C_T, \mathcal{B}_T) . Let E_θ^T denote the expectation w.r.t. measure P_θ^T . Under the assumption (A1), $P_\theta^T \ll P_W^T$ and the Radon-Nikodym derivative (likelihood function based on sample path $X_0^T := (X_s, 0 \leq s \leq T)$) is given by

$$L_T(\theta) = \frac{dP_\theta^T}{dP_W^T}(X_0^T) = \exp \left\{ \int_0^T f(\theta, t, X) dX_t - \frac{1}{2} \int_0^T f^2(\theta, t, X) dt \right\}. \quad (2.2)$$

Maximization of the likelihood $L_T(\theta)$ provides the maximum likelihood estimate (MLE) θ_T . Dietz [6] proved the consistency and asymptotic mixed normality of θ_T as $T \rightarrow \infty$. Kutoyants [25, 27] also studied the asymptotic properties of MLE as the diffusion coefficient $\epsilon \rightarrow 0$ (for fixed T).

Let $l_T(\theta) := \log L_T(\theta)$. We assume that differentiation under the stochastic and ordinary integral sign is valid (see [23]).

Note that θ_t is the MLE based on the sample path $X_0^t := (X_s, 0 \leq s \leq t)$. The MLE process satisfies the following evolution equation (see [34]) when $X = X_t$ in the drift coefficient in equation (2.1):

$$\begin{aligned} d\theta_t &= -l_t''^{-1}(\theta_t) \{ f'(\theta_t, t, X_t) [dX_t - f(\theta_t, t, X_t) dt] \\ &\quad + [\frac{1}{2} l_t'''(\theta_t) l_t''^{-2}(\theta_t) f'^2(\theta_t, t, X_t) - l_t''^{-1}(\theta_t) f'(\theta_t, t, X_t) f''(\theta_t, t, X_t)] dt \}, \end{aligned} \quad (2.3)$$

with initial conditions: $\theta_{t_0} < \infty$, $l_t'(\theta_{t_0}) = 0$, $l_t''(\theta_{t_0}) < 0$ where the initial time $t_0 > 0$ is fixed. This shows that in this case the MLE process itself is a continuous semimartingale.

We have

$$\begin{aligned} l_T'(\theta) &= \int_0^T f'(\theta, t, X) dX_t - \int_0^T f(\theta, t, X) f'(\theta, t, X) dt \\ &= \int_0^T f'(\theta, t, X) [dX_t - f(\theta, t, X) dt] \\ &= \int_0^T f'(\theta, t, X) dW_t \end{aligned} \quad (2.4)$$

when θ is the true parameter.

We shall use the Fisher information to define a stopping rule. The a priori Fisher information is

$$I(\theta) = \text{var}[l_T'(\theta)] = E_\theta^T [l_T'(\theta)]^2 = E_\theta^T \int_0^T f'^2(\theta, t, X) dt. \quad (2.5)$$

But $I(\theta)$ can not be observed since it depends on the unknown parameter. The a posteriori observed Fisher information or stochastic Fisher information is given by

$$I_T = \int_0^T f'^2(\theta_t, t, X) dt \quad (2.6)$$

where θ_t is the MLE based on the observation $\{X_s, 0 \leq s \leq t\}$.

Note that in the linear case the a priori Fisher information does not contain the unknown parameter. We are now ready to define the sequential sampling plan (τ, θ_τ) as follows :

The stopping time τ is defined as

$$\tau \equiv \tau_H := \inf\{t \geq 0 : \int_0^t f'^2(\theta_s, s, X)ds \geq H\} \tag{2.7}$$

where $H > 0$ specifies the desired precision which is predetermined and θ_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

$$L_\tau(\theta) = \frac{dP_\theta^\tau}{dP_W^\tau}(X_0^\tau) = \exp\left\{\int_0^\tau f(\theta, t, X)dX_t - \frac{1}{2}\int_0^\tau f^2(\theta, t, X)dt\right\}. \tag{2.8}$$

(see [32]). The sequential maximum likelihood estimate (SMLE) is defined as

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

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}_τ measurable SMLM since $L_\tau(\theta)$ is continuous in θ and Θ is compact. Hereafter we assume the existence of such a measurable SMLM.

3. CONSISTENCY AND ASYMPTOTIC NORMALITY

We assume the following conditions to prove the main results of this section.

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

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

(B3) $\lim_{H \rightarrow \infty} \frac{1}{H} \int_0^\tau f''(\theta, t, X)dW_t = 0$ in P_θ -probability.

Remark 3.1. Condition (B1) is the identifiability condition. Due to the consistency of the MLE θ_t and by the definition of θ_τ this condition can be verified. Condition (B2) is provided for regularity. Condition (B3) is related to the law of large numbers for stochastic integrals.

Lemma 3.1. *Let $A_t^\theta := f(\theta \pm \delta, t, X) - f(\theta, t, X)$ for some $\delta > 0$. Then under (A1)–(A3), we have*

$$\frac{\int_0^\tau (A_t^\theta)^2 dt}{\int_0^\tau (A_t^{\theta_t})^2 dt} \rightarrow 1 \quad \text{a.s. } [P_\theta^\tau] \text{ as } H \rightarrow \infty.$$

Proof. We have

$$\begin{aligned} A_t^\theta &= f(\theta \pm \delta, t, X) - f(\theta, t, X), \\ A_t^{\theta_t} &= f(\theta_t \pm \delta, t, X) - f(\theta_t, t, X). \end{aligned}$$

Thus

$$(A_t^\theta)^2 - (A_t^{\theta_t})^2 = [f(\theta \pm \delta, t, X) - f(\theta, t, X)]^2 - [f(\theta_t \pm \delta, t, X) - f(\theta_t, t, X)]^2.$$

Since θ_t is a strongly consistent estimator of θ , $\theta_t \rightarrow \theta$ a.s. as $t \rightarrow \infty$. Since f is continuous, $(A_t^\theta)^2 - (A_t^{\theta_t})^2 \rightarrow 0$ a.s. as $t \rightarrow \infty$.

Further

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

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

$$\int_0^\tau [(A_t^\theta)^2 - (A_t^{\theta_t})^2] dt \rightarrow 0 \text{ a.s. as } H \rightarrow \infty.$$

Thus

$$\frac{\int_0^\tau [(A_t^\theta)^2 - (A_t^{\theta_t})^2] dt}{\int_0^\tau (A_t^{\theta_t})^2 dt} \rightarrow 0 \text{ a.s. } [P_\theta^\tau] \text{ as } H \rightarrow \infty.$$

This completes the proof of the lemma. \square

The definition of first order efficiency (in the sense of C.R. Rao) of estimators for stochastic processes (cf. [17]) may be generalized to sequential sampling case as follows:

Definition 3.1. An estimator V_{τ_H} of θ is said to be *first order efficient* (FOE) if

$$\sqrt{H} \left[V_{\tau_H} - \theta - \gamma(\theta) H^{-1} \frac{\partial l_\tau(\theta)}{\partial \theta} \right] \xrightarrow{P_\theta} 0 \text{ as } H \rightarrow \infty$$

where $\gamma(\theta)$ is a nonrandom function.

We establish following properties of the SMLE.

Theorem 3.1. (a) Under the assumption (A1) – (A3), the sampling plan is closed, i.e., $\tau_H < \infty$ P_θ -a.s.

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., $\lim_{H \rightarrow \infty} \theta_\tau = \theta$ a.s. $[P_\theta]$

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

(d) θ_τ is first order efficient as $H \rightarrow \infty$.

Proof. (a) From assumption (A2) it is obvious that $\tau_H < \infty$ a.s. $[P_\theta]$.

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

$$\begin{aligned}
 & l_\tau(\theta \pm \delta) - l_\tau(\theta) \\
 = & \log \frac{dP_{\theta \pm \delta}^\tau}{dP_\theta^\tau} \\
 = & \int_0^\tau [f(\theta \pm \delta, t, X) - f(\theta, t, X)]dX_t - \frac{1}{2} \int_0^\tau [f^2(\theta \pm \delta, t, X) - f^2(\theta, t, X)]dt \\
 = & \int_0^\tau [f(\theta \pm \delta, t, X) - f(\theta, t, X)]dW_t - \frac{1}{2} \int_0^\tau [f(\theta \pm \delta, t, X) - f(\theta, t, X)]^2 dt \\
 = & \int_0^\tau A_t^\theta dW_t - \frac{1}{2} \int_0^\tau (A_t^\theta)^2 dt.
 \end{aligned}$$

Let

$$K_\tau := \int_0^\tau (A_t^\theta)^2 dt.$$

Then

$$\begin{aligned}
 \frac{l_\tau(\theta \pm \delta) - l_\tau(\theta)}{K_\tau} &= \frac{\int_0^\tau A_t^\theta dW_t}{\int_0^\tau (A_t^\theta)^2 dt} - \frac{1}{2} \\
 &= \frac{W^*(\int_0^\tau (A_t^\theta)^2 dt)}{\int_0^\tau (A_t^\theta)^2 dt} - \frac{1}{2} \\
 &= \frac{W^*(K_\tau)}{K_\tau} - \frac{1}{2}
 \end{aligned} \tag{3.1}$$

by the Skorohod embedding of the martingale $\int_0^\tau A_t^\theta dW_t$ where W^* is some other Brownian motion with independent of K_τ .

Using the assumption (A2) and Lemma 3.1, and the strong law of large numbers for Brownian motion (see [32]) the first term on the r.h.s. of (3.1) converges to zero a.s. as $H \rightarrow \infty$. Hence,

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

Furthermore, $K_\tau > 0$ a.s. $[P_\theta]$ by (B1). Therefore, for almost every $w \in \Omega, \delta$ and θ , there exist some H_0 such that for $H \geq H_0$

$$l_\tau(\theta \pm \delta) < l_\tau(\theta). \tag{3.2}$$

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 θ_τ in $[\theta - \delta, \theta + \delta]$. In view of (3.2), $\theta_\tau \in (\theta - \delta, \theta + \delta)$ for $H > H_0$. Since $l_\tau(\theta)$ is differentiable with respect to θ , it follows that $l'_\tau(\theta_\tau) = 0$ for $H \geq H_0$ and $\theta_\tau \rightarrow \theta$ a.s. as $H \rightarrow \infty$.

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

$$0 = l'_\tau(\theta_\tau) = l'_\tau(\theta) + (\theta_\tau - \theta)l''_\tau(\theta + \beta_\tau(\theta_\tau - \theta)) \tag{3.3}$$

where $|\beta_\tau| \leq 1$ a.s. for sufficiently large H .

Since $\theta_\tau \rightarrow \theta$ a.s. as $H \rightarrow \infty$ by (b) and since $l''_\tau(\theta)$ is continuous by (B3), it follows that $l''_\tau(\theta + \beta_\tau(\theta_\tau - \theta)) - l''_\tau(\theta) \rightarrow 0$ in P_θ^τ -probability as $H \rightarrow \infty$.

Hence as $H \rightarrow \infty$

$$l'_\tau(\theta) + (\theta_\tau - \theta)l''_\tau(\theta) \rightarrow 0 \text{ in } P_\theta^\tau\text{-probability.} \quad (3.4)$$

We have $l'_\tau(\theta) = \int_0^\tau f'(\theta, t, X)dW_t$. Hence using the C.L.T. for stochastic integral (see [2]) and Lemma 3.1, we obtain

$$\frac{l'_\tau(\theta)}{\sqrt{H}} \xrightarrow{\mathcal{D}[P_\theta]} \mathcal{N}(0, 1) \text{ as } H \rightarrow \infty \quad (3.5)$$

Note that when θ is the true parameter

$$\begin{aligned} l''_\tau(\theta) &= \int_0^\tau f''(\theta, t, X)dX_t - \int_0^\tau [f(\theta, t, X)f''(\theta, t, X) + f'^2(\theta, t, X)]dt \\ &= \int_0^\tau f''(\theta, t, X)dW_t - \int_0^\tau f'^2(\theta, t, X)dt. \end{aligned}$$

By Lemma 3.1 and (B3) it follows that

$$\frac{l''_\tau(\theta)}{H} \rightarrow -1 \text{ in } P_\theta\text{-probability as } H \rightarrow \infty. \quad (3.6)$$

From (3.5) and (3.6) it follows that

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

(d) From (3.5) and (3.6) and the definition of first order efficiency with $\gamma(\theta) = 1$, SMLE θ_τ is first order efficient as $H \rightarrow \infty$. \square

4. LOCAL ASYMPTOTIC MINIMAXITY

In this section we prove the local asymptotic minimaxity of the SMLE θ_τ 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 [21, 33]), 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 [21]). 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 θX_t . This process exhibits qualitatively different behavior for different values of the parameter θ . 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) for all θ . Similar situations occur in its discrete time counterpart : the Gaussian autoregressive process of first order and other processes like the Galton-Watson branching processes, pure birth processes etc. (see [15]). Greenwood and Shirayayev [13] 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. Greenwood and Wefelmeyer [14] showed that local asymptotic minimax bound is attained by asymptotically centering estimators even at critical points, which requires sequential sampling. In the nonlinear diffusion type model with sequential sampling, the more general functionally LAQ condition is satisfied which allows us to obtain locally asymptotic minimax bound for the normalized error of the sequential MLE. We now recall the definition of LAQ and functionally LAQ conditions.

Definition 4.1 ([22, 33]). Let $\mathcal{E}_n = (\Omega_n, \mathcal{F}_n, (P_\theta^n, \theta \in \Theta)), n \geq 1$ be a sequence of statistical experiments, where Θ is an open subset of \mathbb{R} . We denote by

$$\Lambda_{\eta, \theta}^n = \log \left(\frac{dP_\eta^n}{dP_\theta^n} \right)$$

the log-likelihood between η and θ at stage n .

We say that the sequence \mathcal{E}_n satisfies the *local asymptotically quadratic* (LAQ) condition at a point $\theta \in \Theta$ if there are random variables Δ_n and Γ_n defined on $(\Omega_n, \mathcal{F}_n), \Gamma_n > 0$ a.s. $[P_\theta^n]$ and a positive numerical sequence $\phi_n \downarrow 0$ such that for each bounded sequence of numbers u_n ,

$$\Lambda_{\theta + \phi_n u_n, \theta}^n - \left(u_n \Delta_n - \frac{1}{2} u_n^2 \Gamma_n \right) \xrightarrow{P_\theta^n} 0$$

and

$$(\Delta_n, \Gamma_n) \rightarrow (\Delta, \Gamma) \text{ in } P_\theta^n\text{-distribution}$$

where Δ and Γ are random variables on a measurable space (Ω, \mathcal{F}, P) with $\Gamma > 0$ a.s. (P) and

$$E_P \exp \left(u \Delta - \frac{1}{2} u^2 \Gamma \right) = 1.$$

The sequence of experiments is called *locally asymptotically Brownian functional* (LABF) if $\Delta = \int_0^1 F_s dW_s$ and $\Gamma = \int_0^1 F_s^2 ds$ with W a standard Brownian motion and F a predictable process with respect to some filtration in \mathcal{F} . It is called *locally asymptotically mixed normal* (LAMN) if $\Delta = \Gamma^{1/2}W_1$ with W_1 standard normal variable independent of Γ and *locally asymptotically normal* (LAN) if, in addition, Γ is nonrandom.

Definition 4.2 ([15]). Let the sequence (\mathcal{E}_n) satisfy LAQ condition at a point $\theta \in \Theta$. A sequence of estimators $\hat{\theta}_n$ is called *asymptotically centering* (AC) sequence at θ if

$$\phi_n^{-1}(\hat{\theta}_n - \theta) - \Gamma_n^{-1} \Delta_n \xrightarrow{P_\theta^n} 0.$$

Asymptotically centering estimators are also called *asymptotic maximum likelihood estimators*.

Let $(\Omega^n, \mathcal{F}^n, \{\mathcal{F}_t^n\}_{t \geq 0}, \{P_\theta^n, \theta \in \Theta\})$ where Θ is an open subset of \mathbb{R} , be a sequence of filtered statistical experiments. Let $P_{\theta,t}^n$ be the restriction of P_θ^n to \mathcal{F}_t^n .

The log-likelihood process $\Lambda_{\eta,\theta}^n$ at stage n between η and θ at time t is defined by

$$\Lambda_{\eta,\theta,t}^n = \log \left(\frac{dP_{\eta,t}^n}{dP_{\theta,t}^n} \right), \quad t \geq 0.$$

Definition 4.3 ([14]). We call the sequence of filtered experiments *functionally LAQ* at θ if there exist processes Δ_n and $\Gamma_n > 0$ a.s. and a positive numerical sequence $\phi_n \downarrow 0$ such that for each bounded sequence of numbers u_n and all $t > 0$,

$$\sup_{s \leq t} \left| \Lambda_{\theta + \phi_n u_n, \theta, s}^n - \left(u_n \Delta_s^n - \frac{1}{2} u_n^2 \Gamma_s^n \right) \right| \xrightarrow{P_\theta^n} 0$$

and

$$(\Delta_t^n, \Gamma_t^n) \rightarrow (\Delta_t, \Gamma_t) \text{ in } P_\theta^n\text{-distribution}$$

where Δ_t and Γ_t are processes on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ with $\Gamma_t > 0$ a.s. and

$$E_P \exp \left(u \Delta_t - \frac{1}{2} u^2 \Gamma_t \right) = 1 \text{ for all } t \geq 0.$$

We call the sequence of filtered models *functionally LABF* if $\Delta_t = \int_0^t F_s dW_s$ and $\Gamma_t = \int_0^t F_s^2 ds$ where W is a standard Brownian motion and F is a predictable process.

Theorem 4.1 ([14]). *Suppose the sequence of filtered models is functionally LAQ at θ , with Δ a continuous local martingale, Γ equal to the quadratic variation $\langle \Delta \rangle$ of Δ , and Γ strictly increasing to infinity a.s. Let ω be a bounded, symmetric, bowl shaped loss function on the real line, and let $\hat{\theta}^n$ be a sequence of estimator processes. For each $n \in \mathbb{N}$ and $t \geq 0$, let*

$$\mathcal{T}_t^n = \inf \{s : \Gamma_s^n > t\}.$$

Then

$$\lim_{b \rightarrow \infty} \liminf_{n \rightarrow \infty} \sup_{|u| \leq b} E_{\theta + \phi_n u} \omega \left(\phi_n^{-1}(\hat{\theta}_{\mathcal{T}_t^n} - (\theta + \phi_n u)) \right) \geq E\omega(W_{1/t}).$$

If a filtered model satisfies the functionally LAN condition, then the filtered model time changed by T^n satisfies the functionally LAQ condition, which is evident from the following result in [14].

Lemma 4.1 *Under the assumptions of Theorem 4.1, for each bounded sequence of numbers u_n and all $t > 0$,*

$$\sup_{s \leq t} \left| \Lambda_{\theta + \phi_n u_n, \theta, \mathcal{T}_s^n} - \left(u_n \Delta_{\mathcal{T}_s^n} - \frac{1}{2} u_n^2 s \right) \right| \xrightarrow{P_\theta^n} 0$$

and

$$\Delta_{\mathcal{T}_t^n} \rightarrow W_1 \text{ in } P_\theta^n\text{-distribution.}$$

If the assertions of Lemma 4.1 hold then the sequence of models is said to satisfy the *functionally LAN* condition.

Definition 4.4 ([14]). *If the sequence of filtered models is functionally LAQ at θ , we call a sequence of estimator process $\hat{\theta}^n$ functionally AC at θ , if for all $t > 0$,*

$$\sup_{s \leq t} \left| \phi_n^{-1}(\hat{\theta}_s^n - \theta) - \Delta_s^n (\Gamma_s^n)^{-1} \right| \xrightarrow{P_\theta^n} 0.$$

The filtered model observed up to \mathcal{T}_t^n is LAN, hence by the classical theory, the local asymptotic minimax theorem applies and $\hat{\theta}_{\mathcal{T}_t^n}^n$ are locally asymptotically minimax if and only if they are AC.

Let us return to our problem. Let us introduce the time changed estimator processes. For each $n \in \mathbb{N}$, we define the retimed processes $X_t^n = n^{-1/2} X_{nt}$ 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 X^n , $\mathcal{F}_t^n = \{X_s^n, s \leq t\} = \sigma\{X_s, s \leq nt\}$. The log-likelihood process at stage n and time t between θ_1 and θ is given by

$$\Gamma_{\theta_1, \theta, t}^n = \int_0^t [f(\theta_1, s, X_s^n) - f(\theta, s, X_s^n)] dW_s^n - \frac{1}{2} \int_0^t [f(\theta_1, s, X_s^n) - f(\theta, s, X_s^n)] ds.$$

We assume the following conditions in this section. There exist $r_n \uparrow \infty$ as $n \uparrow \infty$ and $\nu > 0$ such that:

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

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

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

Theorem 4.2. *Under the assumptions (C1) – (C3), the sequence of filtered models $(\Omega^n, \mathcal{F}^n, \{\mathcal{F}_t^n\}_{t \geq 0}, \{P_\theta^n, \theta \in \Theta\})$ generated by X_t^n satisfy the functionally LAQ condition at θ with*

$$\Delta_t = \int_0^t f'(\theta, s, X_s) dW_s \text{ and } \Gamma_t = \int_0^t f'^2(\theta, s, X_s) ds.$$

Proof. For $\nu \leq t$, we have

$$\begin{aligned} \Lambda_{\theta+r_n^{-1}u_n, \theta, \nu}^n &= \int_0^\nu [f(\theta + r_n^{-1}u_n, s, X_s^n) - f(\theta, s, X_s^n)] dW_s^n \\ &\quad - \frac{1}{2} \int_0^\nu [f(\theta + r_n^{-1}u_n, s, X_s^n) - f(\theta, s, X_s^n)]^2 ds. \end{aligned}$$

By Taylor's formula

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

where

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

Hence

$$\begin{aligned} \Lambda_{\theta+r_n^{-1}u_n, \theta, \nu}^n &= r_n^{-1} u_n \int_0^\nu f'(\theta, s, X_s^n) dW_s^n + \frac{1}{2} r_n^{-2} u_n^2 \int_0^\nu f''(\bar{\theta}, s, X_s^n) dW_s^n \\ &\quad - \frac{1}{2} r_n^{-2} u_n^2 \int_0^\nu f'^2(\theta, s, X_s^n) ds - \frac{1}{8} r_n^{-4} u_n^4 \int_0^\nu f''^2(\bar{\theta}, s, X_s^n) ds \\ &\quad - \frac{1}{8} r_n^{-3} u_n^3 \int_0^\nu f'(\theta, s, X_s^n) f''(\bar{\theta}, s, X_s^n) ds \\ &= r_n^{-1} u_n \int_0^\nu f'(\theta, s, X_s^n) dW_s^n - \frac{1}{2} r_n^{-2} u_n^2 \int_0^\nu f'^2(\theta, s, X_s^n) ds + o_{P_\theta^n}(1) \\ &\quad \text{(by assumption (C1) – (C3))} \\ &= u_n \Delta_\nu^n - \frac{1}{2} u_n^2 \Gamma_\nu^n + o_{P_\theta^n}(1), \end{aligned}$$

where

$$\Delta_\nu^n = r_n^{-2} \int_0^\nu f'(\theta, s, X_s^n) dW_s^n \text{ and } \Gamma_\nu^n = r_n^{-2} \int_0^\nu f'^2(\theta, s, X_s^n) ds.$$

Let

$$\Delta_\nu = \int_0^\nu f'(\theta, s, X_s) dW_s \text{ and } \Gamma_\nu = \int_0^\nu f'^2(\theta, s, X_s) ds.$$

By the functional CLT for martingales and stability of weak convergence (see [20]), we obtain

$$(\Delta_\nu^n, \Gamma_\nu^n) \xrightarrow{\mathcal{D}[P_\theta^n]} (\Delta_\nu, \Gamma_\nu) \text{ as } n \rightarrow \infty.$$

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

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

Thus the sequence of filtered models are functionally LAQ at θ . □

We now obtain the local asymptotic minimax theorem for SMLE processes.

Theorem 4.3. *Let ω be a bounded, symmetric, bowl shaped loss function on the real line. Let θ_t^n be a sequence of MLE processes w.r.t. 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

$$\tau_H^n = \inf \{ t \geq 0 : r_n^{-2} \int_0^t f'^2(\theta_s^n, s, X_s^n) ds \geq H \}$$

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

$$\lim_{b \rightarrow \infty} \liminf_{n \rightarrow \infty} \sup_{|u| \leq b} E_{\theta_{\tau_H^n}^{-1} u} \omega \left(\theta_{\tau_H^n}^n - (\theta + r_n^{-1} u) \right) \geq E \omega(W_{1/H}).$$

Proof. Since by Theorem 4.2 the sequence of filtered models generated by X_n^t is functionally LAQ at θ with Δ_t a continuous local martingale, Γ_t equal to the quadratic variation $\langle \Delta \rangle_t$ of Δ_t , and $\Gamma_t \uparrow \infty$ a.s. hence the theorem follows from Theorem 4.1.

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

$$\sup_{G \leq H} \left| \Lambda_{\theta_{\tau_H^n}^{-1} u_n, \theta, \tau_G^n}^n - \left(u_n \Delta_{\tau_G^n}^n - \frac{1}{2} u^2 G \right) \right| = o_{P_{\theta^n}}(1)$$

and

$$\Delta_{\tau_H^n}^n \overset{\mathcal{D}[P_{\theta^n}]}{\underset{\mathcal{P}}{\rightsquigarrow}} W_1$$

by Lemma 4.1. Hence by Theorem 4.1 we obtain the local asymptotic minimax theorem. □

Theorem 4.4. *Under the conditions (B1) - (B3) and (C1) – (C3) the sequence of estimators $\theta_{\tau_H^n}^n$ are locally asymptotically minimax (LAM), i.e., they attain the lower bound in Theorem 4.3.*

Proof. Here we have only to show that $\theta_{\tau_H^n}^n$ are asymptotically centering (AC), that is

$$r_n(\theta_{\tau_H^n}^n - \theta) - H^{-1} \Delta_{\tau_H^n}^n = o_{P_{\theta^n}}(1) \text{ as } n \rightarrow \infty. \tag{4.1}$$

Note that

$$\begin{aligned} r_n(\theta_{\tau_H^n}^n - \theta) &= \frac{-r_n^{-1}l'_{\tau_H^n}(\theta)}{\gamma_n^{-2}l''_{\tau_H^n}(\theta + \beta_{\tau_H^n}(\theta_{\tau_H^n}^n - \theta))} \\ &\sim \frac{-r_n^{-1} \int_0^{\tau_H^n} f'(\theta, s, X_s^n) dW_s^n}{r_n^{-2} \int_0^{\tau_H^n} f''(\theta, s, X_s^n) dW_s^n - r_n^{-2} \int_0^{\tau_H^n} f'^2(\theta, s, X_s^n) ds} \end{aligned} \quad (4.2)$$

by the arguments similar to the proof of Theorem 3.1 (c). On the other hand,

$$H^{-1}\Delta_{\tau_H^n}^n = \frac{\gamma_n^{-1} \int_0^{\tau_H^n} f'(\theta, s, X_s^n) dW_s^n}{\gamma_n^{-2} \int_0^{\tau_H^n} f'^2(\theta, s, X_s^n) ds}. \quad (4.3)$$

From (4.2) and (4.3) using (C3) and Lemma 3.1, we see that (4.1) holds. \square

5. ORNSTEIN-UHLENBECK PROCESS WITH MEMORY

We illustrate the results of the previous sections through the exponential memory Ornstein-Uhlenbeck process. Consider the Ornstein-Uhlenbeck type process which is a solution of the stochastic integro-differential equation (SIDE)

$$dX_t = \theta \left(\alpha \int_0^t e^{-\alpha(t-s)} X_s ds \right) dt + dW_t, \quad X_0 = 0$$

where $\alpha > 0$ is known memory parameter and $\theta \in \mathbb{R}_+$ is the parameter of interest. For each θ the solution X is a diffusion type process where the nonanticipative drift coefficient depends on the whole past of the process. This is a nonhomogeneous Volterra integro-differential equation with a white noise input. Despite losing the Markov property, it describes a linear stochastic system. Here the range of the memory is governed by the parameter α . When $\alpha \rightarrow \infty$, the process losses memory and converges almost surely to the classical Ornstein-Uhlenbeck process. Dietz [7] showed that for $\theta < 0$, the model is LAN and for $\theta > 0$ the model is LAMN. Here the stopping time is given by

$$\tau_H = \inf\{v > 0 : \int_0^v \left(\int_0^t e^{-\alpha(t-s)} X_s ds \right)^2 dt \geq H\}$$

and the SMLE is given by

$$\theta_{\tau_H} = \frac{\int_0^{\tau_H} \left(\int_0^t e^{-\alpha(t-s)} X_s ds \right) dX_t}{\int_0^{\tau_H} \left(\int_0^t e^{-\alpha(t-s)} X_s ds \right)^2 dt}.$$

Here the model satisfies the functional LAN condition and the SMLE is unbiased, strongly consistent and uniformly normally distributed and locally asymptotically minimax.

6. CONCLUDING REMARKS

(1) It is worth investigating to generalize the results in this paper to a nonlinear semimartingale model which includes both the continuous and the discrete time models, e.g., diffusion processes, autoregressive processes and branching processes. For the linear semimartingale model, properties of sequential maximum likelihood estimator were studied in [39] extending the work in [42] for diffusion processes.

(2) It would be interesting to study the properties of sequential Bayes estimators for nonlinear diffusion models. Sequential Bayes estimation for exponential type processes was studied [10].

(3) Obtaining the limiting distribution of the stopping time would be an interesting problem to investigate.

(4) Likelihood ratio processes for interacting particle systems was studied in [36, 37]. It is worth investigating results of this paper for such models.

(5) Replacing the standard Brownian motion W by fractional Brownian motion, produces non-Markovian non-semimartingale solution process [3]. In our case, with still noise driven by standard Brownian motion, using memory in the drift coefficient, we have non-Markovian solution, but still in a semimartingale framework.

REFERENCES

- [1] F.J. Anscombe, Large sample theory of sequential estimation, *Proc. Camb. Phil. Soc.*, 48: 600-607, 1952.
- [2] I.V. Basawa and B.L.S. Prakasa Rao, *Statistical Inference for Stochastic Processes*, Academic Press, New York-London, 1980.
- [3] J.P.N. Bishwal, *Parameter Estimation in Stochastic Differential Equations*, Lecture Notes in Mathematics, Vol. 1923, Springer-Verlag 2008.
- [4] V. Borkar and A. Bagchi, Parameter estimation in continuous time stochastic processes, *Stochastics*, 8:193-212, 1982.
- [5] B.M. Brown and J.T. Hewitt, Inference for the diffusion branching process, *J. Appl. Prob.*, 12:588-594, 1975.
- [6] H.M. Dietz, Asymptotic properties of maximum likelihood estimators in diffusion type models, Preprint No. 228, Humboldt-Universitat Zu Berlin, Sektion Mathematik, 1989.
- [7] H.M. Dietz, A non-Markovian relative to the Ornstein-Uhlenbeck process and some of its local statistical properties, *Scand. J. Statist.*, 19:363-379, 1992.

- [8] A.A. Dmitrienko and V.V. Konev, On guaranteed estimation of autoregressive parameter for an unknown variance of noise, *Automation and Remote Control*, 55:218-228, 1994.
- [9] A.A. Dmitrienko, V. Konev and A. Pergamenshchikov Sequential generalised least squares estimator for an autoregressive parameter, *Sequential Analysis*, 16:25-46, 1997.
- [10] J. Franz and R. Magiera, Admissible estimators in sequential plans for exponential type processes, *Scand. J. Statist.*, 17:275-285, 1990.
- [11] M. Ghosh, N. Mukhopadhyaya, P.K. Sen, *Sequential Estimation*, John Wiley, New York, 1997.
- [12] P. Grambsch, Sequential sampling based on the observed Fisher information to guarantee the accuracy of the maximum likelihood estimator, *Ann. Statist.*, 11:68-77, 1983.
- [13] P.E. Greenwood and A.N. Shiriyayev, Asymptotic minimaxity of a sequential estimator for a first order autoregressive model, *Stoch. Stoch. Rep.*, 38:49-65, 1992.
- [14] P.E. Greenwood and W. Welfelmeyer, Asymptotic minimax results for stochastic process families with critical points, *Stoch. Proc. Appl.*, 44:107-116, 1993.
- [15] A.A. Gushchin, On the asymptotic optimality of estimators of parameters under the LAQ condition, *Theory. Probab. Applications*, 40:261-272, 1995.
- [16] A.A. Gushchin and U. Küchler, Asymptotic inference for a linear stochastic differential equation with time delay, To appear in *Bernoulli*, 5:1059-1098, 1999.
- [17] P. Hall and C.C. Heyde, *Martingale Limit Theory and its Applications*, Academic Press, New York, 1980.
- [18] Höpfner, R. On statistics of Markov processes: representation of log-likelihood ratio processes in filtered local models, *Probab. Theory Rel. Fields*, 94:375-398, 1993.
- [19] R. Höpfner, Asymptotic inference for Markov step processes: observation up to a random time, *Stoch. Process. Appl.*, 48:295-310, 1993.
- [20] J. Jacod and A.N. Shiriyayev, *Limit Theorems for Stochastic Processes*, Springer, Berlin, 1987.
- [21] P. Jeganathan, On the asymptotic theory of statistical inference when the limit of the log-likelihood ratios is mixed normal, *Sankhyā Ser. A*, 44:173-212, 1982.
- [22] P. Jeganathan, Some aspects of asymptotic theory with applications to time series models, *Econometric Theory*, 11:818-887, 1995.

- [23] R.L. Karandikar, Interchanging the order of stochastic integration and ordinary differentiation, *Sankhyā Ser. A*, 45:120-124, 1983.
- [24] U. Küchler, and M. Sørensen, *Exponential Family of Stochastic Processes*, Springer, NewYork, 1997.
- [25] Yu. A. Kutoyants, Parameter Estimation for diffusion type processes of observations, *Math. Operationsforsch. U. Statist., Ser. Statist.*, 15:541-551, 1984.
- [26] Yu. A. Kutoyants, *Parameter Estimation for Stochastic Processes*, Heldermann-Verlag, Berlin, 1984.
- [27] Yu. A. Kutoyants, *Identification of Dynamical Systems with Small Noise*, Kluwer, Dordrecht, 1994.
- [28] Yu. A. Kutoyants, *Statistical Inference for Ergodic Diffusion Processes*, Springer-Verlag, New York-Berlin, 2003.
- [29] T.L. Lai and D. Siegmund, Fixed accuracy estimation of an autoregressive parameter, *Ann. Statist.*, 11:478-485, 1985.
- [30] A. Le Breton and M. Musiela, On sequential estimation of parameters of continuous Gaussian Markov processes, *Proab. Math. Statist.*, 2:37-53, 1981.
- [31] A. Le Breton and M. Musiela, Some parameter estimation problems for hypoelliptic homogeneous Gaussian diffusions, *Sequential Methods in Statistics*, Banach Centre Publications, Vol. 16, 337-356, Warsaw, 1985.
- [32] R.S. Liptser and A.N. Shiriyayev, *Statistics of Random Processes I, II*, Springer, Berlin, 1977, 1978.
- [33] L. Le Cam and G.L. Yang, *Asymptotics in Statistics*, Springer, New York, 1990.
- [34] D. Levanony, A. Shwartz, and O. Zeitouni, Recursive identification in continuous-time stochastic processes, *Stoch. Proc. Appl.*, 49:245-275, 1994.
- [35] Yu. N. Linkov, Local asymptotic mixed normality for distributions of processes of diffusion type, *Theory Probab. Math. Statist.*, 40:67-73, 1990.
- [36] E. Löcherbach, Likelihood ratio processes for Markovian particle systems with killing and jumps, *Statistical Inference for Stochastic Processes*, 5:153-177, 2002.
- [37] E. Löcherbach, LAN and LAMN for systems of interacting diffusions with branching and immigration, *Ann. Inst. Henri Poincare*, 38:59-2002, 2002.
- [38] V.K. Malinovskii, Asymptotic expansions in sequential estimation of an autoregressive parameter, *Math. Methods Statist.*, 2:206-227, 1993.
- [39] A.V. Melnikov and A.A. Novikov, Sequential inferences with prescribed accuracy for semimartingales, *Theory Prob. Appl.*, 33:446-459, 1988.

- [40] M. Musiela, Sequential estimation of parameters of a stochastic differential equation, *Math. Operationsforsch. U. Statistik, Ser Statistics*, 8:483-498, 1977.
- [41] M. Musiela, Admissible and minimax estimation for two classes of stochastic processes, *Bulletin de L' Academic Polonaise des Sciences, Série des Sci. Math.*, 27:511-514, 1979.
- [42] A.A. Novikov, Sequential estimation of the parameters of diffusion type processes, *Mathematical Notes*, 12:812-818, 1972.
- [43] A. Pergamenshchikov Asymptotic properties of sequential design for estimating the parameter of a first order autoregression, *Theory Probab. Appl.*, 36:36-49, 1991.
- [44] R. Rozanskii, Markov stopping sets and stochastic integrals - Application in sequential estimation for a random diffusion field, *Stoch. Proc. Appl.*, 32:237-252, 1989.
- [45] M. Sørensen, On maximum likelihood estimation in randomly stopped diffusion-type processes, *Int. Stat. Review*, 51:93-110, 1983.
- [46] M. Sørensen, On sequential maximum likelihood estimation for exponential families of stochastic processes, *Int. Stat. Review*, 54:191-210, 1986.
- [47] M.S. Tikhov, On optimal sequential estimation plans for nonquadratic loss functions, *Theory Prob. Appl.*, 23: 132-138, 1978.
- [48] M.S. Tikhov, On the sequential estimation of the drift coefficient of a diffusion process with quadratic and nonquadratic losses, *Theory Prob. Appl.*, 25:607-614, 1980.