Institut für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

A minimum–distance estimator for diffusion processes with ergodic properties

H.M. Dietz[1] , Y. Kutoyants[2]

submitted: 21st October 1992

[1] Institut für Angewandte Analysis
und Stochastik
Hausvogteiplatz 5-7
D – O 1086 Berlin
Germany

[2] Université du Maine
Département de Mathématiques
B.P. 535
72017 Le Mans, Cedex
France

Preprint No. 17
Berlin 1992

1991 Mathematics Subject Classification. 62 M 05, 60 J 60.
Key words and phrases. minimum distance estimator, diffusion process, ergodic property, consistency, asymptotic normality.

Herausgegeben vom
Institut für Angewandte Analysis und Stochastik
Hausvogteiplatz 5-7
D – O 1086 Berlin

Fax: + 49 30 2004975
e-Mail (X.400): c=de;a=dbp;p=iaas-berlin;s=preprint
e-Mail (Internet): preprint@iaas-berlin.dbp.de

A MINIMUM - DISTANCE ESTIMATOR FOR DIFFUSION PROCESSES WITH ERGODIC PROPERTIES

HANS M. DIETZ AND YURIJ KUTOYANTS

ABSTRACT. Suppose one observes one path of a stochastic process $X = (X_t)_{t \geq 0}$ which is known to solve an equation of the form

$$dX_t^\theta = S(\theta,\ X_t^\theta)dt + dW_t, \quad t \geq 0, \quad \theta \in \Theta \subset \mathbb{R}^d \tag{0.1}$$

with a given coefficient functional S and given initial condition X_0, where Θ is a non-void bounded open subset of $\mathbb{R}^d$. In order to estimate the true but unknown parameter θ_0 the paper proposes the minimum distance estimator (MDE) $\hat{\theta}_T$ given by

$$\hat{\theta}_T \in \arg\inf_{\theta \in \Theta} \int_0^T (X_t - X(\theta)_t)^2 dt, \quad T > 0, \tag{0.2}$$

where

$$X(\theta)_t := X_0 + \int_0^t S(\theta, X_u)\, du, \quad t \geq 0 \tag{0.3}$$

and studies its asymptotic behaviour as $T \to \infty$. Under the main assumption that the observed process has an ergodic property and some further (less restrictive) conditions it is shown that $\hat{\theta}_T$ is strongly consistent and - in case $d = 1$ - asymptotically normal. In particular, the results apply to models where $S(\theta, x) = S(\theta - x)$. Several examples and a comparison with likelihood estimation are added.

1. INTRODUCTION

Suppose one observes one path of a stochastic process $X = (X_t)_{t \in [0,T]}$ which is known to solve an equation of the form

$$dX_t^\theta = S(\theta,\ X_t^\theta)dt + dW_t, \quad t \geq 0, \quad \theta \in \Theta \subset \mathbb{R}^d \tag{1.1}$$

with a given initial condition X_0, where S is a given coefficient functional and Θ a non-void bounded open parameter subset of $\mathbb{R}^d$. It is assumed that, for every $\theta \in \Theta$, (1.1) has a unique strong solution and that for the observation it holds $X = X^{\theta_0}$ for some θ_0.

The purpose of the present paper is to study the performance of minimum distance estimators (MDE) $\hat{\theta}_T$ for θ_0, given by

$$\hat{\theta}_T \in \arg\inf_{\theta \in \Theta} \int_0^T (X_t - X(\theta)_t)^2 dt, \quad T > 0, \tag{1.2}$$

where

$$X(\theta)_t := X_0 + \int_0^t S(\theta, X_u)\, du, \quad t \geq 0 \tag{1.3}$$

as $T \to \infty$. Under the basic assumption that the underlying process $X = X^{\theta_0}$ has an ergodic property and some further (less restrictive) conditions it shall be shown that MDE according to (1.2) exist a.s. for sufficiently large T, are strongly consistent and are - in case $d = 1$- asymptotically normal.

The minimum distance estimators of the above type were introduced by KUTOYANTS [3, 5, 4] in connection with stochastic differential equations similar to (1.1), with the main difference that in [5, 4] the observation "time" T was kept fixed, whereas consistency and asymptotic normality of the MDE were established under the assumption of asymptotically vanishing noise.

It should be underlined that (1.2) provides another example of what is likewise called "minimum contrast estimator" (MCE) or "minimum dispersion estimator" in the literature. General properties of such estimators have been investigated, e.g., by PFANZAGL [11], MILLAR [9, 10], DACUNHA-CASTELLE and DUFLO [2] and BASAWA and KOUL [1]. The main contributions to the study of such estimators in connection with diffusion (type) processes are due to LÁNSKA [6] and KUTOYANTS et al. [3, 5, 4]. As far as the asymptotics for increasing 'time' of observation are concerned, it has to be mentioned that in [6] strong consistency and asymptotic normality were proved for a large (abstract) class of minimum contrast estimators, containing the maximum likelihood estimator (MLE) - but, however, not our estimator (1.2). Let us note that the MLE is universal in the sense that is provides a map from the set of all (suitable) coefficient functionals S into the set of "contrast processes", and an interesting question should be whether there are further universal MCE, with possibly interesting properties. In [3], KUTOYANTS proposed another universal MDE for diffusion processes with ergodic properties, which is based on distances between empirical distributions of the observed process and the theoretical stationary distributions that correspond to different parameter values. However, in some applications these stationary distributions may be difficult to deal with explicitly. Thus, the main motivation for considering (1.2) may lie in the fact that it proposes another universal MDE which is quite natural and rather easily manageable.

The techniques for establishing the forementioned properties of the MDE (1.2) consist in an analysis of the underlying distance process rather than in referring to general asymptotic results as given, e. g., by DACUNHA- CASTELLE & DUFLO [2] or MILLAR [9, 10]. The reason for this is, in part, that the (consistency) result presented here is somewhat stronger than these in the mentioned references, and in part, that the assumptions required e.g. in [9] are not fulfilled in the present case.

The paper is organised as follows : In Section 2, the assumptions concerning the model are made precise. Section 3 is devoted to the existence and consistency result, while asymptotic normality is proved in Section 4. Section 5 is devoted to a particular class of models which shall be called "shifted coefficient models". Examples are given in Section 6, and Section 7 provides a comparison with likelihood estimators.

2. THE MODEL

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a basic probability space , $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ a filtration in $\mathcal{F}$ subject to the usual hypotheses, and $W = (W_t)_{t\geq 0}$ a standard Wiener process adapted to $\mathbb{F}$. Further let $\emptyset \neq \Theta \subset \mathbb{R}^d$ be open and bounded and $S : \Theta \times \mathbb{R} \longrightarrow \mathbb{R}$ be a mapping with the property that for each $\theta \in \Theta$ the function $x \longmapsto S(\theta, x)$ is Borel. Consider the stochastic differential equations (SDE) (1.1)

$$dX_t^\theta = S(\theta, X_t^\theta)dt + dW_t, \quad t \geq 0,$$

with a given initial condition X_0. Throughout the present paper it shall be assumed that S satisfies the following Lipschitz and growth conditions.

($\exists$!) For every $\theta \in \Theta$ there is a constant L_θ for which

$$\begin{aligned} |S(\theta,x) - S(\theta,y)| &\leq L_\theta |x-y|, \\ |S(\theta,x)| &\leq L_\theta(1+|x|), \end{aligned}$$

for every $x, y, \in \mathbb{R}$.

It is well-known that under these conditions for each $\theta \in \Theta$ the SDE (1.1) has a unique strong solution (cf. [7]).

It shall be assumed further that the observed process X solves (1.1) for a certain $\theta_0 \in \Theta$, i.e. $X = X^{\theta_0}$, and that it obeys the following ergodic property:

(ERG (θ_0)) There is a probability measure $\mu = \mu_{\theta_0}$ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that

(i) for each $f \in L_1(\mu_{\theta_0})$ it holds $\mathbb{P}$ - a.s.

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(X_t)dt = \int_{\mathbb{R}} f(x)\mu_{\theta_0}(dx) =: \mu_{\theta_0}(f) \ ,$$

(ii) $S(\theta_0, \cdot) \in L_1(\mu_{\theta_0})$.

The measure μ_{θ_0} could be referred to as the stationary distribution of X. Note that in case of existence of such a measure it is necessarily unique.

Remark 1. A constructive sufficient condition for (ERG) to hold is the following one: Let

$$B_{\theta_0}(x) := 2\int_0^x S(\theta_0, y)dy, \quad x \in \mathbb{R}, \tag{2.1}$$

and define a nonnegative measure m_{θ_0} on $(\mathbb{R}, \mathcal{B}, (\mathbb{R}))$ by

$$m_{\theta_0}(dx) = 2\exp B_{\theta_0}(x)dx \ .$$

If $\int_{\mathbb{R}} |x| m_{\theta_0}(dx) < \infty$ and ($\exists$!) is fulfilled then (ERG(θ_0)) holds with

$$\mu_{\theta_0} = m_{\theta_0}(\mathbb{R})^{-1} m_{\theta_0}$$

(see, e. g. [8]).

CONVENTIONS

In order to simplify notations let us omit unnecessary indices, variables, and so on, wherever any misunderstanding is excluded from the context. In particular, let us write λ for the Lebesque measure on $\mathbb{R}$ and $\int_0^T f d\lambda$ instead of $\int_0^T f(t)dt$ etc. Dependencies of random variables on $\omega \in \Omega$ shall be suppressed as far as possible. We shall briefly write $S(\theta)$ both for the function $x \mapsto S(\theta, x)$ and for the stochastic process $(S(\theta, X_t^{\theta_0}))_{t\geq 0}$, as the context shall always suggest the right interpretation.

3. EXISTENCE AND STRONG CONSISTENCY OF MDE

Let $D = (D_T(\theta))_{T\geq 0}^{\theta\in\Theta}$ be a given $\mathbb{R}$-valued random field with the property that for each $\theta \in \Theta$ the random process $D(\theta) = (D_T(\theta))_{T\geq 0}$ is $\mathbb{F}^X$ - adapted, where $\mathbb{F}^X = (\mathcal{F}_t^X)_{t\geq 0}$ and $\mathcal{F}_t^X :=^{\sigma}\langle X_s, s \leq t\rangle$. Let us agree to call D a *distance* if there is a family $(E_T)_{T\geq 0}$ of sets $E_T \in \mathcal{F}_T^X$ such that

$$\mathbb{P}(E_T) \longrightarrow 1 \quad \text{as} \quad T \to \infty \tag{3.1}$$

and for sufficiently large T

$$[\arg\inf_\theta D_T(\theta) = \emptyset] \subset \Omega \setminus E_T \,. \tag{3.2}$$

A $\mathbb{R}^d$-valued random process $(\hat\theta_T)_{T\geq 0}$ shall be called a MDE (w.r.t. the distance D) if for every $T \geq 0$ there is a $\mathbb{P}$-null set F_T such that

$$[\hat\Theta_T \neq \emptyset] \setminus F_T \subset [\hat\theta_T \in \hat\Theta_T]$$

where

$$\hat\Theta_T(\omega) := \arg\inf_{\theta\in\Theta} D_T(\theta)(\omega), \quad \omega \in \Omega \,. \tag{3.3}$$

In the following attention shall be directed to the random field D defined by

$$D_T(\theta) := \frac{1}{T^3}\int_0^T (X - X(\theta))^2 d\lambda \,, \tag{3.4}$$

where

$$X(\theta) := X_0 + \int_0^{\cdot} S(\theta, X_t)dt \,. \tag{3.5}$$

It is clear from the assumptions made in the preceding section that D is well-defined. The task of the present section is to show that D is a distance and that there exists a MDE w.r.t. D and, moreover, that any such MDE is strongly consistent.

Under $(\exists!)$ and $(ERG(\theta_0))$, introduce the following two assumptions :

$(HC(\theta_0))$ ("Hölder continuity of $S(\cdot, x)$")
There are a constant $\alpha > 0$ and a function $g \in L_1(\mu_{\theta_0})$ such that for every $\theta, \theta' \in \Theta$

$$|S(\theta, x) - S(\theta', x)| \leq |\theta - \theta'|^\alpha g(x), \; x \in \mathbb{R} \,. \tag{3.6}$$

If $(HC(\theta_0))$ is fulfilled, then obviously it holds that

$$S(\theta, \cdot) \in L_1(\mu_{\theta_0})\cdot$$

for every $\theta \in \Theta$. Thus, one can define a function $k_{\theta_0} : \Theta \longrightarrow \mathbb{R}$ as follows :

$$k_{\theta_0}(\theta) := \mu_{\theta_0}(S(\theta_0) - S(\theta)) \,, \; \theta \in \Theta \,. \tag{3.7}$$

As this function shall turn out to be basic for the contrast associated with D, one has to impose further the following condition :

$(ID(\theta_0))$ (Identifiability condition) For the function $\rho_{\theta_0} : [0, \infty] \to [0, \infty]$ defined by

$$\rho_{\theta_0}(c) := \inf\{|k_{\theta_0}(\theta)|^2 \; : \; \theta \in \Theta, |\theta - \theta_0| > c\} \tag{3.8}$$

it holds

$$c > 0 \implies \rho_{\theta_0}(c) > 0 \ . \tag{3.9}$$

(We use the convention $\inf \emptyset = \infty$.)

Remark 2. In case $d = 1$ $ID(\theta_0)$ is implied by the following condition :

$(ID_1(\theta_0))$ There is a neighbourhood $U(\theta_0) \subset \Theta$ of θ_0 such that
(i) the function $\theta \mapsto k_{\theta_0}(\theta)$ is strictly monotonous on $U(\theta_0)$
(ii) there is a constant $\kappa > 0$ with

$$\inf_{\theta \in \Theta \setminus U(\theta_0)} |k_{\theta_0}(\theta_0)| \geq \kappa > 0. \tag{3.10}$$

The main result of this section is

Theorem 3.1. *Assume* $(\exists!), (ERG(\theta_0), (HC(\theta_0)),$ *and* $(ID(\theta_0))$. *Then* D *is a distance, and*

(i) There exists a minimum distance estimator $(\hat{\theta}_T)_{T \geq 0}$ *w.r.t.* D.
(ii) Every such MDE is strongly consistent at θ_0 , *i.e.*

$$\mathbb{P}(\lim_{T \to \infty} \hat{\theta}_T = \theta_0) = 1 \ . \tag{3.11}$$

Before proving this theorem we shall provide a series of auxiliary results. As the reference parameter θ_0 remains fixed, it shall be suppressed in the notation.

Lemma 3.1. *Under* (ERG) *and* (HC) *the function* $\theta \mapsto k(\theta)$ *is Hölder continuous of order* α.

Proof. Indeed,

$$|k(\theta) - k(\theta')| = |\mu(S(\theta') - S(\theta))| \leq |\theta' - \theta|^\alpha |\mu(g)|.$$

□

Remark 3. If (ID) holds, then Lemma 1 implies that

$$\arg\inf_{\theta \in \Theta} k^2(\theta) = \{\theta_0\} \ . \tag{3.12}$$

Lemma 3.2. *Assume* $(\exists!), (ERG),$ *and* (HC).

(i) For each $T \geq 0$ *and each* $\omega \in \Omega$, *the function* $\theta \mapsto D_T(\theta)(\omega)$ *is continuous.*

(ii) There is a $\mathbb{P}$ *-null set* $N \subset \Omega$ *such that*

$$\forall \omega \in \Omega \setminus N \quad \exists M \geq 0 : \forall T \geq 1 \quad \forall \theta, \tilde{\theta} \in \Theta$$

$$|D_T(\theta)(\omega) - D_T(\tilde{\theta})(\omega)| \leq |\theta - \tilde{\theta}|^\alpha M \tag{3.13}$$

Remark 4. Another way of stating (3.13) is to say that for $\mathbb{P}$ -almost all ω the set of continuous functions $\{D_T(\cdot)(\omega), T \geq 1\} =: \mathcal{D}(\omega)$ is uniformly equicontinuous.

Proof of Lemma 2. First of all, note that the (unique strong) solution of (1.1) is uniquely defined for <u>all</u> $\omega \in \Omega$, as, in fact, the integral associated with (1.1) is a Lebesque integral and, thus, (1.1) can be solved path-wise uniquely.

(i) : Let $\theta, \tilde{\theta} \in \Theta$ be arbitrary. By definition of D and as $X = X^{\theta_0}$ solves

(1.1) for the parameter θ_0 it holds (for every (suppressed) $\omega \in \Omega$)

$$\begin{aligned} D_T(\theta) - D_T(\tilde{\theta}) &= \frac{1}{T^3} \int_0^T \left\{ \left(\int_0^\cdot [S(\theta_0) - S(\theta)] d\lambda \right)^2 - \left(\int_0^\cdot [S(\theta_0) - S(\tilde{\theta})] d\lambda \right)^2 + \right. \\ & \left. + 2W \int_0^\cdot [S(\tilde{\theta}) - S(\theta)] d\lambda \right\} d\lambda \end{aligned}$$

where, for brevity, $S(\theta) := (S(\theta, X_t))_{t \geq 0}, \theta \in \Theta$. Hence,

$$|D_T(\theta) - D_T(\tilde{\theta})| \leq H_T^1(\theta, \tilde{\theta}) + H_T^2(\theta, \tilde{\theta}) \tag{3.14}$$

where

$$\begin{aligned} H_T^1(\theta, \tilde{\theta}) &= \frac{1}{T^3} \int_0^T \left| \left(\int_0^\cdot [S(\theta_0) - S(\tilde{\theta})] d\lambda \right)^2 - \left(\int_0^\cdot [S(\theta_0) - S(\theta)] d\lambda \right)^2 \right| d\lambda, \\ H_T^2(\theta, \tilde{\theta}) &= \frac{2}{T^3} \int_0^T |W| \int_0^\cdot |S(\tilde{\theta}) - S(\theta)| d\lambda d\lambda. \end{aligned}$$

It is apparent that

$$\begin{aligned} H_T^1(\theta, \tilde{\theta}) &\leq \frac{1}{T^3} \int_0^T \left(\int_0^\cdot (|S(\theta_0) - S(\theta)| + |S(\theta_0) - S(\tilde{\theta})|) d\lambda \right) \times \\ & \times \int_0^\cdot |S(\tilde{\theta}) - S(\theta)| d\lambda d\lambda. \end{aligned} \tag{3.15}$$

An application of (2.6) to (2.15) gives

$$H_T^1(\theta, \tilde{\theta}) \leq |\theta - \tilde{\theta}|^\alpha \left(|\theta_0 - \theta|^\alpha + |\theta_0 - \tilde{\theta}|^\alpha \right) H_T^3 \tag{3.16}$$

with

$$H_T^3 = \frac{1}{T^3} \int_0^T \left(\int_0^\cdot g \circ X d\lambda \right)^2 d\lambda \tag{3.17}$$

(note that this term is free of θ and $\tilde{\theta}$). One can proceed in a similar manner to obtain the following estimate for $H_T^2(\theta, \tilde{\theta})$:

$$H_T^2(\theta, \tilde{\theta}) \leq |\theta - \tilde{\theta}|^\alpha H_T^4 \tag{3.18}$$

with

$$H_T^4 = \frac{2}{T^3} \int_0^T |W| \int_0^\cdot g \circ X d\lambda d\lambda, \tag{3.19}$$

(again this term is free of θ and $\tilde{\theta}$). But (3.14), (3.16) and (3.18) together prove (i).
($D_T(\cdot)(\omega)$ is even Hölder continuous of order α, with random time-dependend constant $H_T^3(\omega) + H_T^4(\omega)$.)

(ii) Fix a null set $N \subset \Omega$ such that for every $\omega \in \Omega \setminus N$ it holds

$$\begin{aligned} \lim_{t\to\infty} \frac{1}{t} W_t(\omega) &= 0 \\ \lim_{t\to\infty} \frac{1}{t} \int_0^t g \circ X(\omega) d\lambda &= \mu(g) , \end{aligned}$$

where g is the function from the (HC) condition. Now, by L'Hospital's rule, for every $\omega \in \Omega \setminus N$

$$\lim_{T\to\infty} H_T^3(\omega) = \lim_{T\to\infty} \frac{1}{3} \left(\frac{1}{T} \int_0^T g \circ X(\omega) d\lambda \right)^2 = \frac{1}{3} \mu(g)^2 \tag{3.20}$$

and

$$\lim_{T\to\infty} H_T^4(\omega) = \lim_{T\to\infty} \frac{2}{3} \cdot \frac{|W_T(\omega)|}{T} \cdot \frac{1}{T} \int_0^T g \circ X(\omega) d\lambda = 0. \tag{3.21}$$

As the function $T \mapsto H_T^3(\omega) + H_T^4(\omega)$ is continuous, (3.20) and (3.21) show that there is an $M = M(\omega)$ such that

$$0 \leq H_T^3(\omega) + H_T^4(\omega) \leq M(\omega) \quad for \quad T \geq 1,$$

hence, from (3.14), (3.16) and (3.18), (3.13) follows. □

Lemma 3.3. *Under* $(\exists!)$, *(ERG), and(HC) there is a* $\mathbb{P}$*-null set* $\tilde{N} \subset \Omega$ *such that for every* $\omega \in \Omega \setminus \tilde{N}$ *it holds*

$$D_T(\theta)(\omega) \longrightarrow k(\theta)^2 \quad as \quad T \to \infty , \tag{3.22}$$

uniformly $\omega.r.t.$ $\omega \in \Theta$.

Proof. Take the null set N from Lemma 3.2, then, for any $\omega \in \Omega \setminus N$, the set of functions $\{k^2(\cdot)\} \cup \mathcal{D}(\omega) = \{k^2(\cdot)\} \cup \{D_T(\cdot)(\omega),\, T \geq 1\}$, is uniformly Hölder continuous of order α . Further, for every $\theta \in \Theta$ one can find a null set N_θ such that for every $\omega \in \Omega \setminus N_\theta$

$$\frac{1}{T} W_T(\omega) \longrightarrow 0 \quad \text{as} \quad T \to \infty \quad \text{and}$$

$$\frac{1}{T} \int_0^T S(\tilde{\theta}, X_t(\omega)) dt \longrightarrow \mu(S(\tilde{\theta})) \quad \text{for } \tilde{\theta} \in \{\theta, \theta_0\} \quad \text{as} \quad T \to \infty,$$

hence,for $\omega \in \Omega \setminus N_\theta$, by L'Hospital's rule

$$\lim_{T\to\infty} D_T(\theta)(\omega) = \lim_{T\to\infty} \frac{1}{3} \left(\frac{1}{T} \int_0^T (S(\theta_0) - S(\theta)) d\lambda + \frac{W_T}{T} \right)^2 = k(\theta)^2 .$$

Now let $\tilde{N} := N \bigcup_{\theta \in \Theta \cap \mathbb{Q}^d} N_\theta$ and choose any $\omega \in \tilde{N}$. One can find a constant M_ω such that all functions $h \in \{k^2\} \cup \mathcal{D}(\omega)$ satisfy the same Hölder condition: $|h(\theta) - h(\tilde{\theta})| \leq |\theta - \tilde{\theta}|^\alpha M_\omega$. Given any $\varepsilon > 0$, let $\{\theta_1^\varepsilon, \dots, \theta_{n_\varepsilon}^\varepsilon\} =: \Theta_\varepsilon \subset \Theta \cap \mathbb{Q}^d$ be a finite δ-net for Θ , where δ is chosen such that $\delta_\alpha M_\omega < \varepsilon/3$, and choose T_0 so large that $\sup_{\tilde{\theta} \in \Theta_\varepsilon} |D_T(\tilde{\theta}) - k(\tilde{\theta})^2| < \varepsilon/3$ for $T \geq T_0$. Hence, for $T \geq T_0$, and any $\theta \in \Theta$

$$\begin{aligned} |D_T(\theta) - k(\theta)^2| &\leq \min_{\tilde{\theta} \in \Theta_\varepsilon} \{ |D_T(\theta) - D_T(\tilde{\theta})| + |D_T(\tilde{\theta}) - k(\tilde{\theta})^2| + |k(\tilde{\theta})^2 - k(\theta)^2| \} \\ &\leq \delta^\alpha M_\omega + \varepsilon/I + \delta^\alpha M_\omega < \varepsilon/I . \end{aligned}$$

This proves the assertion with the null set $\tilde{N}$. □
In the following lemma we use the notation

$$\begin{aligned} A - z &:= \{x - z : x \in A\} \quad \text{and} \\ \operatorname{diam} A &:= \sup\{|x-y| : x, y \in A\} \end{aligned}$$

for any set $A \subset \mathbb{R}^d$. The result is well-known :

Lemma 3.4. *Let (f_n) be a sequence of continuous functions $f_n : \Theta \to \mathbb{R}$ that converges uniformly to a continuous function $f : \Theta \to \mathbb{R}$. Suppose f is bounded from below and* $\arg\inf_{\theta \in \Theta} f(\theta) = \{\theta_0\} \subset \Theta$.
Then:

(i) for sufficiently large n,

$$\emptyset \neq \arg\inf_{\theta\in\Theta} f_n(\theta) \subset \Theta$$

(ii) $\operatorname{diam}\,(\arg\inf_{\theta\in\Theta} f_n(\theta) - \theta_0) \to 0 \quad \text{as} \quad n \to \infty.$ □

We can now proceed with the *Proof of Theorem* 3.1.
(i) By Remark 3, the function k^2 can take the role of the function f in Lemma 3.4. Lemma 3.3 together with Lemma 3.4 implies for every $\omega \in \Omega \setminus N$ that $\hat{\Theta}_T(\omega)$ is non-void for sufficiently large T . We show that the mapping $\omega \mapsto 1_{\{\hat{\Theta}_T(\omega)\neq\emptyset\}}$ is measurable for every $T > 0$. Indeed, it is clear from the proof of Lemma 3.2 (i) that for each $T \geq 0$ the assignment $\omega \mapsto D_T(\cdot)(\omega)$ defines a mapping $D_T : \Omega \to C_b(\Theta)$ where $C_b(\Theta)$ denotes the space of continuous bounded real-valued functions on Θ. If $C_b(\Theta)$ is endowed with the topology of uniform convergence and with the corresponding Borel σ-field $\mathcal{C}$ then it is easy to see that the mapping D_T is $\mathcal{F}_T^X - \mathcal{C}$ measurable. Hence, for proving that $\omega \mapsto 1_{\{\hat{\Theta}_T(\omega)\neq\emptyset\}}$ is measurable it is sufficient to show that

$$A^* := \{f \in C_b(\Theta) : \quad \arg\inf f = \emptyset\} \in \mathcal{C}\,.$$

Note that for each non-void subset K of Θ the mapping

$$I_K : C_b(\Theta) \to \mathbb{R} : f \mapsto \inf_{\theta\in K} f(\theta) =: I_K(f)$$

is continuous, thus, measurable. For every $\varepsilon > 0$ let

$$U_\varepsilon(\Theta^c) := \{\eta \in \mathbb{R}^d : \inf\{|\eta - \zeta|, \zeta \in \mathbb{R}^d \setminus \Theta\} < \varepsilon\}$$

and define

$$\Theta[\varepsilon] := \overline{\Theta \setminus U_\varepsilon(\Theta^c)}.$$

Now it is clear that

$$A^* = \bigcap_{n\in\mathbb{N}} \{f \in C_b(\theta) : I_\Theta(f) < I_{\Theta[1\setminus n]}\};$$

but the sets under the intersection sign are in $\mathcal{C}$, hence, so is A^*, proving the desired measurability property.
It is clear now that $\mathbb{P}(\hat{\Theta}_T \neq \emptyset) \to 1$ as $T \to \infty$, i.e. D is a distance.
The proof of the existence of a MDE reduces now to showing the existence of a measurable selector, which can be accomplished as in [11] (details can be omitted here).
(ii) Suppose now $(\hat{\Theta}_T)_{T\geq 0}$ to be any MDE w.r.t. D. Then its strong consistency is an obvious consequence of Lemma 3.4 (ii). □

Note: Theorem 3.1 does not tell anything about the uniqueness of MDE. If card $\hat{\Theta}_T(\omega) > 2$ then the problem arises which point has to be chosen out of this set.

However, part (ii) of Theorem 3.1 states that for any choice of this point the resulting estimator is strongly consistent.

Remark 5. Theorem 3.1 remains valid if Θ is assumed to be nonvoid and bounded only. (Indeed, all statements needed for consistency (e.g.3.4) remain valid even for non-open Θ. Morever, proving that $\hat{\Theta}_T$ is non-empty for sufficiently large T becomes even easier, as

$$\arg\inf_{\theta\in\Theta} f(\theta) \supset \arg\inf_{\theta\in int\Theta} f(\theta)$$

for every $f \in C_b(\Theta)$.)

4. ASYMPTOTIC NORMALITY OF THE MDE

In the present section attention shall be directed to one-dimensional open parameter sets, i.e., it shall be assumed that Θ is of the form $\Theta = (\alpha, \beta)$ for some $\alpha, \beta \in \mathbb{R}$ with $\alpha < \beta$. Again it is assumed that $\theta_0 \in \Theta$ is the true parameter.
The conditions required for the consistency result of Theorem 3.1 are 5 assumed throughout here. Furthermore, the continuity condition $(HC(\theta_0))$ shall be replaced by the following stronger condition.

$(DC(\theta_0))$ ("Differentiability and Hölder continuity"):
There is a neighbourhood $U(\theta_0) \subset \Theta$ of θ_0 with the following properties.

(i) $\forall x \in \mathbb{R}, \quad \theta \mapsto S(\theta, x)$ is twice continuously differentiable on $U(\theta_0)$ (the derivatives w.r.t. θ shall be denoted by $\dot{S}, \ddot{S}$ or by $S^{(i)}$, $i \leq 2$, likewise);

(ii) $\dot{S}(\theta_0, \cdot) \in L_1(\mu_{\theta_0}) \ni \ddot{S}(\theta_0, \cdot)$;

(iii) there are functions $w_0, w_1, w_2 \in L_1(\mu_{\theta_0})$ such that, for some $\kappa > 0$, the following "integrable" local Lipschitz and Hölder conditions hold : for every $x \in \mathbb{R}$ and $\theta \in U(\theta_0)$

$$|S(\theta, x) - S(\theta_0, x)| \leq |\theta - \theta_0| \omega_0(x) \tag{4.1}$$
$$|\dot{S}(\theta, x) - \dot{S}(\theta_0, x)| \leq |\theta - \theta_0| \omega_1(x) \tag{4.2}$$
$$|\ddot{S}(\theta, x) - \ddot{S}(\theta_0, x)| \leq |\theta - \theta_0|^{\kappa} \omega_2(x) \tag{4.3}$$

(iv) $\mu_{\theta_0}(\dot{S}(\theta_0)) \neq 0$.

Theorem 4.1. *Assume* $(\forall!), (ERG), (DC)$, *and* (ID). *Let* $(\hat{\theta}_T)_{T\geq 0}$ *be a* MDE *w.r.t.* D. *Then it holds*

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \Longrightarrow N(0, \frac{6}{5}\mu(\dot{S})^{-2}) \tag{4.4}$$

as $T \to \infty$.

Proof: The (ERG) and (DC) assumptions together with Theorem 3.1 imply that there is a $\mathbb{P}$-null set N with the property that for each $\omega \in \Omega \setminus N$ one has simultaneously, as $T \to \infty$,

$$\hat{\theta}_T(\omega) \to \theta_0$$

and

$$\frac{1}{T}\int_0^T f(X_t(\omega))dt \to \mu(f)$$

for every $f \in \{\dot{S}(\theta_0, \cdot), |\dot{S}(\theta_0, \cdot)|, |\ddot{S}(\theta_0, \cdot)|, w_0, w_1, w_2\}$. Let us fix such an $\omega \in \Omega \setminus N$ (and drop it in the following notation, as it remains fixed).

For sufficiently large T - say $T \geq T_0$ - it holds $\hat{\theta}_T \in int\ U(\theta_0)$ (where $U(\theta_0)$ is the neighbourhood of θ_0 from the (DC) condition). Thus, $\hat{\theta}_T$ must be a local minimum point of the function $D_T : \theta \mapsto D_T(\theta)$. As by (DC) it is twice continuously differentiable, one concludes

$$\dot{D}_T(\hat{\theta}_T) = 0 \ . \tag{4.5}$$

On the other hand, by Taylor's formula,

$$\dot{D}_T(\hat{\theta}_T) = \dot{D}_T(\theta_0) + \ddot{D}_T(\breve{\theta}_T)(\hat{\theta}_T - \theta_0) \tag{4.6}$$

with some $\breve{\theta}_T$ that satisfies $|\breve{\theta}_T - \theta_0| < |\hat{\theta}_T - \theta_0|$. ($\breve{\theta}_T$ can depend on ω; note however that there is no measurability problem in what follows, as by (4.6) the r.v. $\ddot{D}_T(\breve{\theta}_T)1_{\{\hat{\theta}_T \neq \theta_0\}}1_N$ is measurable.)
If one can show that for sufficiently large T - say $T \geq T_1 \geq T_0$ - it holds

$$\ddot{D}_T(\breve{\theta}_T) \neq 0 \tag{4.7}$$

then (4.5) and (4.6) imply

$$\sqrt{T}(\hat{\theta}_T - \theta_0) = -\sqrt{T}\dot{D}_T(\theta_0)\ddot{D}_T(\breve{\theta}_T)^{-1} \ . \tag{4.8}$$

In order to justify (4.7) we shall prove that

$$\ddot{D}_T(\breve{\theta}_T) \longrightarrow \frac{2}{3}\mu(\dot{S})^2 \tag{4.9}$$

(which is positive by $(DCiv)$).
Furthermore we shall show that

$$\sqrt{T}\dot{D}_T(\theta_0) \Longrightarrow N(0, \frac{8}{15}\mu(\dot{S})^2) \tag{4.10}$$

as $T \to \infty$; and the assertion follows from (4.8) - (4.10) at once .

Proof of (4.9) :
Let us start with calculating the derivatives of $D_T(\cdot)$. It is clear from (2.4) that

$$\begin{aligned} \dot{D}_T(\theta) &= -\frac{2}{T^3}\int_0^T \left(X - \int_0^\cdot S(\theta)d\lambda\right)\left(\int_0^\cdot \dot{S}(\theta)d\lambda\right)d\lambda \qquad (4.11)\\ \ddot{D}_T(\theta) &= \frac{2}{T^3}\int_0^T \left\{\left(\int_0^\cdot \dot{S}(\theta)d\lambda\right)^2 - \right.\\ &\quad \left. - \left(\int_0^\cdot (S(\theta_0) - S(\theta))d\lambda + W\right)\left(\int_0^\cdot \ddot{S}(\theta)d\lambda\right)\right\} d\lambda. \qquad (4.12) \end{aligned}$$

As $\hat{\theta}_T \to \theta_0$, so $\breve{\theta}_T \to \theta_0$; hence, (4.9) can be proved by showing first that the functions $\theta \mapsto \ddot{D}_T(\theta)$ are uniformly continuous at θ_0 for sufficiently large T and second, that

$$\ddot{D}_T(\theta_0) \longrightarrow \frac{2}{3}\mu(\dot{S})^2 \text{ as } T \to \infty.$$

For $\theta \in U(\theta_0)$ one has, by (4.12)

$$|\ddot{D}_T(\theta) - \ddot{D}_T(\theta_0)| \leq \Delta_1(T) + \Delta_2(T) + \Delta_3(T)$$

with

$$\Delta_1(T) = \frac{2}{T^3}\int_0^T \left|(\int_0^{\cdot} \dot S(\theta)d\lambda)^2 - (\int_0^{\cdot} \dot S(\theta_0)d\lambda)^2\right| d\lambda \tag{4.13}$$

$$\Delta_2(T) = \frac{2}{T^3}\int_0^T \left|\int_0^{\cdot} (S(\theta_0) - S(\theta))d\lambda\right|\left|\int_0^{\cdot} \ddot S(\theta)d\lambda\right| d\lambda \tag{4.14}$$

$$\Delta_3(T) = \frac{2}{T^3}\int_0^T |W|\left|\int_0^{\cdot} (\ddot S(\theta) - \ddot S(\theta_0))d\lambda\right| d\lambda\,. \tag{4.15}$$

Now we use the inequalities (4.1) - (4.3) to see that

$$\begin{aligned}\Delta_1(T) &\le \frac{2}{T^3}\int_0^T \left(\int_0^{\cdot} |\dot S(\theta) + \dot S(\theta_0)|d\lambda\right)\left(\int_0^{\cdot} |\dot S(\theta) - \dot S(\theta_0)|d\lambda\right) d\lambda \\ &\le \frac{2}{T^3}\int_0^T \left(\int_0^{\cdot} (2|\dot S(\theta_0)| + w_1(X)|\theta-\theta_0|)d\lambda\right)\left(\int_0^{\cdot} w_1(X)d\lambda\right) d\lambda|\theta-\theta_0|\,,\end{aligned}$$

the right hand expression converging to

$$|\theta-\theta_0|\,\frac{2}{3}\left(2\mu(|\dot S|) + |\theta-\theta_0|\mu(w_1)\right)\mu(w_1) \tag{4.16}$$

as $T \to \infty$;

$$\Delta_2(T) \le \frac{2}{T^3}\int_0^T \left(\int_0^{\cdot} |\theta-\theta_0|w_0(X)d\lambda\right)\left(\int_0^{\cdot} (|\ddot S(\theta_0)| + |\theta-\theta_0|^{\kappa}\, w_2(X))d\lambda\right) d\lambda$$

with the right hand limit

$$|\theta-\theta_0|\,\frac{2}{3}\mu(w_0)(\mu(|\ddot S|) + |\theta-\theta_0|^2\mu_0(w_2)) \tag{4.17}$$

and that

$$\Delta_3(T) \le |\theta-\theta_0|^{\kappa}\frac{2}{T^3}\int_0^T |W|\int_0^{\cdot} w_3(X)d\lambda d\lambda\,, \tag{4.18}$$

where the integral vanishes as $T \to \infty$. (4.16) - (4.18) together show that, for some $T_1 \ge T_0, M > 0$, and $|\theta-\theta_0| < 1$, $\theta \in U(\theta_0)$,

$$|\ddot D_T(\theta) - \ddot D_T(\theta_0)| \le |\theta-\theta_0|^{1\wedge\kappa} M\,,$$

which proves the desired uniform continuity of $\ddot D_T(\cdot)$ at θ_0 . (Note that this also implies $\ddot D_T(\breve\theta_T) \ne 0$ for sufficiently large T .) Now, the same reasoning as in the proof of (3.20) and (3.21) shows that

$$\ddot D_T(\theta_0) \longrightarrow \frac{2}{3}\mu(\dot S)^2 \text{ as } T \to \infty.$$

Proof of (4.10) :
(Let us underline again that $\dot D_T, \dot S$, etc. refer to $\dot D_T(\theta_0), \dot S(\theta_0, X)$, etc.)

Let us note that, following (4.11),

$$\sqrt{T}\dot D_T = -\frac{2}{T^{5/2}}\int_0^T W\int_0^{\cdot} \dot S d\lambda\, d\lambda = -\frac{2}{T^{5/2}}\mu(\dot S)N_T + H_T$$

where

$$\begin{aligned} N_T &= \int_0^T (1+t)W_t dt \\ H_T &= -\frac{2}{T^{5/2}} \int_0^T (1+t)W_t \left(\frac{1}{1+t}\int_0^t \dot{S}d\lambda - \mu(\dot{S})\right) dt. \end{aligned} \tag{4.19}$$

It turns out that

$$H_T \xrightarrow{\mathbb{P}} 0 \quad \text{as } T \to \infty . \tag{4.20}$$

Indeed, for every $\delta > 0,\ \tau > 0$, with $W_T^* := \sup_{t \le T} |W_t|$,

$$\begin{aligned} \mathbb{P}(|H_T| \ge \delta) &\le \mathbb{P}\left(\frac{2}{\sqrt{T}} W_T^* \cdot \frac{1}{T^2}\int_0^T (1+t)\left|\frac{1}{1+t}\int_0^t \dot{S}d\lambda - \mu(\dot{S})\right| dt \ge \delta\right) \\ &\le \mathbb{P}\left(\frac{2}{\sqrt{T}} W_T^* \ge \tau\right) + \mathbb{P}\left(\frac{1}{T^2}\int_0^T (1+t)\left|\frac{1}{1+t}\int_0^t \dot{S}d\lambda - \mu(\dot{S})\right| dt \ge \frac{\delta}{\tau}\right). \end{aligned}$$

As the right-hand probability vanishes asymptotically and as

$$\mathbb{P}\left(\frac{2}{\sqrt{T}} W_T^* \ge \tau\right) \le \frac{1}{2}\exp\left\{-\frac{1}{2T}\left(\frac{\tau\sqrt{T}}{2}\right)^2\right\}$$

it follows that, for every $\tau > 0$,

$$\overline{\lim_{T\to\infty}} \mathbb{P} \quad (|H_T| \ge \delta) \le \frac{1}{2}\exp\left(-\frac{\tau^2}{8}\right),$$

proving (4.20). In view of (4.20) it is enough to calculate the asymptotic variance of the Gaussian r.v. $-\frac{2}{T^{5/2}}\mu(\dot{S})N_T$ as $T \to \infty$.

Note that N_T can be rewritten as

$$N_T = \int_0^T (1+t)W_t dt = \int_0^T W_t d_t \frac{(1+t)^2}{2} = \frac{(1+T)^2}{2} W_T - \int_0^T \frac{(1+t)^2}{2} dW_t$$

by Itô's formula, i.e.

$$N_T = \int_0^T \left(\frac{(1+T)^2}{2} - \frac{(1+t)^2}{2}\right) dW_t$$

and hence $N_T \in N(0, \sigma_T^2)$, where

$$\begin{aligned} \sigma_T^2 &= \int_0^T \left(\frac{(1+T)^2}{2} - \frac{(1+t)^2}{2}\right)^2 dt \\ &= \frac{(1+T)^4}{4}T - \frac{(1+T)^2[(1+T)^3-1]}{6} + \frac{(1+T)^5-1}{20}. \end{aligned}$$

It follows that

$$-\frac{2}{T^{5/2}} N_T \mu(\dot{S}) \to N(0, \sigma_\infty^2) \quad \text{with} \quad \sigma_\infty^2 = \frac{8}{15}\mu(\dot{S})^2,$$

hence (4.10). □

Remark 6. By exactly the same arguments one can prove that in case that Θ is no longer one-dimensional but a non-empty bounded open subset of $\mathbb{R}^d$, under the further assumptions of Theorem 4.1, it holds as $T \to \infty$

$$\sqrt{T}\mu(\dot{S})\mu(\dot{S})^T(\hat{\theta}_T - \theta_0) \Longrightarrow N(0, \frac{6}{5}\mu(\dot{S})\mu(\dot{S})^T)$$

(where the T–sign means transpose) . This indicates that only the one-dimensional functional $\mu_{\theta_0}(\dot{S}(\theta_0,\cdot))^T\theta_0$ of θ_0 can be estimated $\sqrt{T}$ - consistently (although, by Theorem 3.1, the full estimate $\hat{\theta}_T$ for θ_0 itself is strongly consistent).

5. "THE SHIFTED COEFFICIENT MODEL"

The purpose of the present section is to present a particular model of the type (1.1) which has the advanteage that verifying the assumptions of Theorems 3.1 and 4.1 becomes rather easy. The model is defined by the SDE

$$dX_t^\theta = c(X_t^\theta - \theta)dt + dW_t, \quad t \geq 0,\ X_0 \quad \text{given} \tag{5.1}$$

with $\theta \in \Theta \subset \mathbb{R}^1$ (Θ bounded, non-void), where $c : \mathbb{R} \to \mathbb{R}$ is a measurable function. Clearly this model is of type (1.1) with $S(\theta, x) = c(x - \theta), i.e.$ the parameter θ determines a shift of the coefficient function which explaines the above terminology.

Consider the following assumptions:

$$(\exists!_S) \qquad \exists L > 0 : \forall x, y \quad \begin{aligned} |\, c(x) - c(y)\, | &\leq L|x-y| \\ |c(x)| &\leq L(1+|x|) \end{aligned}$$

$$(ERG_S) \qquad A \mapsto m(A) := \int_A \exp\{2\int_0^x c(y)dy\}dx \quad \text{defines a}$$
$$\text{finite measure on} \quad (\mathbb{R}, \mathcal{B}(\mathbb{R}))\,.$$

(Note that these conditions imply $(\exists!), (ERG_{\theta_0})$ and (HC_{θ_0}) to hold for any $\theta_0 \in \Theta$.)
Letting $\mu := m(\mathbb{R})^{-1}m$, the following identifiability condition is imposed.

(ID_S) The function h , defined by

$$h(\Delta) := \int_{\mathbb{R}} (c(x) - c(x-\Delta))\mu_0(dx),\ \Delta \in \mathbb{R},$$

is strictly monotonous in a neighbourhood U of 0 and satisfies

$$|k(\Delta)| \geq \tilde{c} > 0$$

for some constant $\tilde{c}$ and every $\Delta \notin U$.

Corollary 5.1. Under $(\exists!_S), (ERG_S)$, and (ID_S) :

(i) There is a MDE $(\hat{\theta}_T)$ w.r.t. D for the model (5.1).

(ii) Every such MDE is strongly consistent for every $\theta_0 \in \Theta$.

Proof. It is straightforward to see that the above assumptions imply those of Theorem 3.1.
□

In order to obtain the asymptotic distribution of the estimators one needs the following somewhat stronger condition.

(DC_S) : (i) The function c is twice continuously differentiable (with derivatives c', c'') .

(ii) There are constants $M > 0$ and $\kappa > 0$ such that for every $x, y \in \mathbb{R}$

$$|c'(x) - c'(y)| \leq M|x-y|$$

and

$$|c''(x) - c''(y)| \leq M|x-y|^{\kappa}.$$

(iii) It holds

$$\mu_0(c') \neq 0.$$

Again it is only one straightforward step to check that (DC_S) implies that (DC_{θ_0}) is fulfilled for every θ_0. Thus, one obtains

Corollary 5.2. Under $(\exists!_S), (ERG_S), (ID_S)$, and (DG_S), let $(\hat{\theta}_T)$ be a MDE for the model (5.1). Then it holds for every $\theta_0 \in \Theta$

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \Longrightarrow N(0, \frac{6}{5}\mu_0(c')^{-2})$$

as $T \to \infty$. □

6. EXAMPLES

6.1. "Shifted ORNSTEIN-UHLENBECK model". Consider the SDE

$$dX_t^{\theta} = \rho(X_t - \theta), \quad t \geq 0, \quad X_0 \text{ given} \tag{6.1}$$

where ρ is a fixed constant and $\theta \in \Theta$ an unknown location parameter. It is well-known - and easy to see as well - that in case $\rho < 0$ the assumptions $(\exists!_S)$, (ERG_S) and (DC_S) are fulfilled with $\mu_0 = N(0, -\frac{1}{2\rho})$ (hence the stationary distribution of X^{θ_0} is $N(\theta_0, -\frac{1}{2\rho})$); moreover, (ID_S) is fulfilled as well as $k(\theta) = \theta$. We find from Corollaries 5.1 and 5.2 that there is a MDE $(\hat{\theta}_T)$ for θ_0 and that every MDE is strongly consistent and satisfies

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \Longrightarrow N(0, \frac{6}{5}\rho^{-2}) \tag{6.2}$$

as $T \to \infty$.

Howewer, in the present example the random field D is just given by

$$D_T(\theta) = \frac{1}{T^3}\int_0^T \left(X_t - X_0 - \int_0^t \rho(X_n - \theta)du\right)^2 dt\,, \quad T > 0. \tag{6.3}$$

Hence, if Θ is open then for sufficiently large T there is a *unique* MDE which is given by

$$\hat{\theta}_T = \frac{3}{\rho T^3}\int_0^T t\left(\rho\int_0^t X_n du + X_0 - X_t\right)dt\,, \quad T > 0. \tag{6.4}$$

It might be of some interest to compare the behaviour of this estimator with that of the maximum likelihood estimator (MLE) (θ_T^*) , determined by

$$\theta_T^* \in \arg\max\left\{\int_0^T \rho(X_t - \theta)dX_t - \frac{1}{2}\int_0^T \rho^2(X_t - \theta)^2 dt\right\} \tag{6.5}$$

(cf. [7]), which - under the above assumptions on θ - yields

$$\theta_T^* = \frac{1}{T}\left(\int_0^T X_t dt - \frac{1}{\rho}X_T\right). \tag{6.6}$$

One obtains (see also (7.4)) for every θ_0

$$\sqrt{T}(\theta_T^* - \theta_0) \Longrightarrow N(0, \rho^{-2}). \tag{6.7}$$

6.2. "Counter - example". Now let us consider the model (6.1) again, but with respect to $\rho < 0$ as unknown parameter, i.e. $S(\rho, x) = \rho(x - \theta)$ with θ kept fixed. This is *not* a shifted coefficient model, howewer, it is straightforward that $(\exists!), (ERG_{\rho_0})$, and (DC_{ρ_0}) are fulfilled with $\mu_{\rho_0} = N(\theta, -\frac{1}{2\rho_0})$.

Note that the function k_{ρ_0} of (3.7) takes the form

$$k_{\rho_0}(\rho) = (\rho_0 - \rho) \int_{\mathbb{R}} x \mu_{\rho_0}(dx) = (\rho_0 - \rho)\theta. \tag{6.8}$$

This implies that (ID_{ρ_0}) is fulfilled if $\theta \neq 0$; the model is *not* identifiable if $\theta = 0$!

By writing down $D_T(\rho)$ and assuming the parameter domain to be open one finds the *MDE*

$$\hat{\rho}_T = \frac{\int_0^T (X_t - X_0) \int_0^t (X_u - \theta) du dt}{\int_0^T \left(\int_0^t (X_u - \theta) du \right)^2 dt}, \tag{6.9}$$

which is strongly consistent and satisfies

$$\sqrt{T}(\hat{\rho}_T - \rho_0) \Longrightarrow N(0, \frac{6}{5}\theta^{-2}). \tag{6.10}$$

In order to derive a MLE one has to maximize the right- hand expression in (6.5) w.r.t. ρ and obtains

$$\rho_T^* = \frac{\int_0^T (X_t - \theta) dX_t}{\int_0^T (X_t - \theta)^2 dt} \tag{6.11}$$

there (cf. again (7.4))

$$\sqrt{T}(\hat{\rho}_T - \rho_0) \Longrightarrow N(0, (\theta^2 - \frac{1}{2\rho_0})^{-1}). \tag{6.12}$$

6.3. A non- linear example. Let us consider also an example where the coefficient function depends on the parameter θ in a non- linear way, namely the SDE

$$dX_t^\theta = \frac{2(\theta - X_t^\theta)}{1 + (\theta - X_t^\theta)^2} dt dW_t, \quad t \geq 0, \quad \text{with } X_0 \text{ given} \tag{6.13}$$

where the parameter θ varies in some bounded interval $\Theta = (\alpha, \beta)$. Clearly this example is of the form (5.1) with

$$c(x) = \frac{-2x}{1 + x^2}, \quad x \in \mathbb{R}\,. \tag{6.14}$$

We claim that it is easy to check the conditions introduced in Section 5 .

Indeed, let us start from checking the (ERG_S) condition . As

$$\int_0^x c(y) dy = -ln(1 + x^2)$$

we see that

$$m(dx) = \frac{1}{(1+x^2)} dx \ ,$$

hence

$$m(\mathbb{R}) = \frac{\pi}{2}$$

and

$$\mu_0(dx) = \frac{\pi}{2} \frac{1}{(1+x^2)^2} \, dx \ .$$

Furthermore, the function c is thrice continuously differentiable with derivatives

$$c'(x) = 2\frac{x^2-1}{(1+x^2)^2} \ , \quad c''(x) = -4\frac{x^3-3x}{(1+x^2)^3} \ , \quad c'''(x) = 12\frac{x^4-6x^2+1}{(1+x^2)^4} \ ,$$

and each of these functions is absolutely bounded by the constant 2 . This proves the $(\exists!_S)$ and (DC_S) conditions at once,

as

$$\mu_0(c') = \frac{4}{\pi} \int_{\mathbb{R}} \frac{x^2-1}{(1+x^2)^4} dx = 1 \ .$$

It remains to check the (ID_S) condition by considering the function

$$\Delta \mapsto h(\Delta) \ := \frac{4}{\pi} \int_{\mathbb{R}} \frac{\Delta - u}{1+(\Delta-u)^2} \frac{1}{(1+u^2)^2} du.$$

One can write

$$\begin{aligned} h(\Delta) &= h_1(\Delta) + h_2(\Delta) \quad \text{with} \\ h_1(\Delta) &= -\frac{4\Delta}{\pi} \int_{\mathbb{R}} \frac{1}{1+(\Delta-u)^2} \frac{1}{(1+u^2)^2} du \\ h_2(\Delta) &= -\frac{2}{\pi} \int_{\mathbb{R}} \frac{1}{1+(\Delta-u)^2} \frac{-2u}{(1+u^2)^2} du \ . \end{aligned}$$

Integration by parts and the substitution $u = \Delta - v$ give

$$\begin{aligned} h_2(\Delta) &= -\frac{2}{\pi} \int \frac{2(\Delta-u)}{(1+(\Delta-u)^2)^2} \frac{1}{1+u^2} du \\ &= -\frac{4\Delta}{\pi} \int \frac{1}{(1+v^2)^2} \frac{1}{1+(\Delta-v)^2} dv - h(\Delta); \end{aligned}$$

i.e.

$$h(\Delta) = -\frac{4\Delta}{\pi} J(\Delta) \quad \text{with} \quad J(\Delta) = \int_{\mathbb{R}} \frac{1}{1+(\Delta-u)^2} \frac{1}{(1+u^2)^2} du.$$

Observe that the function J is continuously differentiable, strictly positive on $\mathbb{R}$ and has a strict maximum at $\Delta = 0$. This implies that h is strictly monotonous in a vicinity $U(0)$ of 0 and bounded away from 0 on $(\Theta - \theta_0) \setminus U(0)$. Hence, (ID_S) is fulfilled, too.

Altogether, Corollaries 5.1 and 5.2 are proved to hold for every $\theta_0 \in \Theta$, i.e. the MDE $(\hat{\theta}_T)$ exists and is strongly consistent, and it holds

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \Longrightarrow N(0, \frac{6}{5})$$

as $T \to \infty$.

7. CONCLUDING REMARKS

A few remarks are in order to relate the proposed MDE $(\hat{\theta}_T)$ according to (1.2) to estimators that can be obtained from a likelihood function, which is available under the $(\exists!)$ condition, e.g. in the form (cf. [7])

$$\Lambda_T(\theta) = \exp\left\{\int_0^T S(\theta, X_t)\, dX_t - \int_0^T S(\theta, X_t)^2\, dt\right\} . \tag{7.1}$$

LÁNSKA [6] has shown that for compact $\Theta \subset \mathbb{R}$ under conditions close to that of Theorem 3.1 there is a strongly consistent MLE (θ_T^*) , i.e. an estimator satisfying

$$\theta_T^* \in \arg\max_{\theta\in\Theta} \Lambda_T(\theta) \tag{7.2}$$

Moreover, under additional assumptions in [6] it is shown that there exists a strongly consistent estimator (θ_T^{**}) that satisfies

$$\frac{\partial}{\partial\theta} \log \Lambda_T(\theta)\Bigg|_{\theta=\theta_T^{**}} = 0 \tag{7.3}$$

and

$$\sqrt{T}(\theta_T^{**} - \theta_0) \Longrightarrow N\left(0, \mu_{\theta_0}(\dot{S}(\theta_0, \cdot)^2)^{-1}\right) . \tag{7.4}$$

(Θ is assumed open here.) (θ_T^{**} can be referred to as a likelihood equation estimator (LEE)).

The MDE $(\hat{\theta}_T)$ proposed here has two obvious disadvanteages in comparison with the MLE (θ_T^*) or the LEE (θ_T^{**}) : First, the identifiability condition (ID) can be violated in cases where MLE / LEE still work, as Example 6.2 shows. And second, a comparison of (7.4) and (4.4) shows that the asymptotic variance of $(\hat{\theta}_T)$ exceeds that of (θ_T^{**}) at least at 20 per cent.

Howewer, let us add three notes reflecting certain (possible) advanteages of the MDE $(\hat{\theta}_T)$.

First, it may be unconvenient from the practical point of view to deal with the stochastic integral in (7.1) . If the function $S(\theta, \cdot)$ is differentiable then the stochastic integral can be replaced by means of a Lebesque integral - if not, the MDE might be a convenient alternative.

Second, the assumptions needed in order to prove (7.4) are considerably stronger than the assumptions imposed here.

And third, one may hope that the MDE provides a procedure that is robust w.r.t. small deviations of the model (opposite to other estimators) . Unfortunately, there is no proof of this property, yet.

REFERENCES

1. I.V. BASAWA and H.L. KOUL, *Large sample statistics based on quadratic dispersion*, Internat. Statist. Review **56** (1988), no. 3, 199 – 219.
2. D. DACUNHA CASTELLE and M. DUFLO, *Probability and statistics*, vol. I, Springer, Berlin, 1986.
3. Y. KUTOYANTS, *Minimum distance parameter estimation for diffusion type observations*, C.R. Acad. Sci. Paris, Serie I **312** (1991), 637 – 642.
4. Y. KUTOYANTS and P. PILIBOSSIAN, *On minimum l_1-norm estimate of the parameter of Ornstein-Uhlenbeck process*, Tech. Report 167, Université Paris VI, LSTA Mathématiques, 1992.

5. ______, *On minimum uniform metric estimate of parameters of diffusion type processes*, Tech. report, Univ. Paris VI, 1992, preliminary version.
6. V. LÁNSKA, *Minimum contrast estimation in diffusion processes*, J. Appl. Prob. **16** (1979), 65 – 75.
7. R. S. LIPTSER and A. N. SHIRYAEV, *Statistics of random processes*, Nauka, Moscow, 1974. (russian)
8. P. MANDL, *Analytical treatment of one-dimensional markov processes*, Springer, Berlin, 1968.
9. P.W. MILLAR, *The minimax principle in asymptotic statistical theory*, Ecole d'Eté de Probabilités de Saint Flour XI - 1981 (P.L. Hennequin, ed.), Lecture Notes in Mathematics No. 976, Springer, Berlin, 1983.
10. P.W. MILLAR, *A general approach to the optimality of minimum distance estimators*, Trans. Am. Math. Soc. **286** (1984), 377 – 418.
11. I. PFANZAGL, *On the measurability and consistency of minimum contrast estimates*, Metrika **14** (1969), 249 – 272.

INSTITUT FÜR ANGEWANDTE ANALYSIS UND STOCHASTIK (IAAS), MOHRENSTRASSE 39, D – O – 1086 BERLIN

UNIVERSITÉ DU MAINE, DEPARTEMENT DE MATHÉMATIQUES, B.P. 535, 72017 LE MANS, FRANCE

Veröffentlichungen des Instituts für Angewandte Analysis und Stochastik

Preprints 1992

1. D.A. Dawson and J. Gärtner: Multilevel large deviations.

2. H. Gajewski: On uniqueness of solutions to the drift–diffusion–model of semiconductor devices.

3. J. Fuhrmann: On the convergence of algebraically defined multigrid methods.

4. A. Bovier and J.-M. Ghez: Spectral properties of one–dimensional Schrödinger operators with potentials generated by substitutions.

5. D.A. Dawson and K. Fleischmann: A super–Brownian motion with a single point catalyst.

6. A. Bovier, V. Gayrard: The thermodynamics of the Curie–Weiss model with random couplings.

7. W. Dahmen, S. Prößdorf, R. Schneider: Wavelet approximation methods for pseudodifferential equations I: stability and convergence.

8. A. Rathsfeld: Piecewise polynomial collocation for the double layer potential equation over polyhedral boundaries. Part I: The wedge, Part II: The cube.

9. G. Schmidt: Boundary element discretization of Poincaré–Steklov operators.

10. K. Fleischmann, F. I. Kaj: Large deviation probability for some rescaled superprocesses.

11. P. Mathé: Random approximation of finite sums.

12. C.J. van Duijn, P. Knabner: Flow and reactive transport in porous media induced by well injection: similarity solution.

13. G.B. Di Masi, E. Platen, W.J. Runggaldier: Hedging of options under discrete observation on assets with stochastic volatility.

14. J. Schmeling, R. Siegmund–Schultze: The singularity spectrum of self-affine fractals with a Bernoulli measure.

15. A. Koshelev: About some coercive inequalities for elementary elliptic and parabolic operators.

16. P.E. Kloeden, E. Platen, H. Schurz : Higher order approximate Markov chain filters.