Estimation of an Ergodic Diffusion from Discrete Observations

MATHIEU KESSLER
Université P. et M. Curie

ABSTRACT. We consider a one-dimensional diffusion process $X$, with ergodic property, with drift $b(x, \theta)$ and diffusion coefficient $a(x, \sigma)$ depending on unknown parameters $\theta$ and $\sigma$. We are interested in the joint estimation of $(\theta, \sigma)$. For that purpose, we dispose of a discretized trajectory, observed at $n$ equidistant times $t_{ni} = i h_n$, $1 \leq i \leq n$. We assume that $h_n \to 0$ and $nh_n \to \infty$. Under the condition $nh_p n \to 0$ for an arbitrary integer $p$, we exhibit a contrast dependent on $p$ which provides us with an asymptotically normal and efficient estimator of $(\theta, \sigma)$.

Key words: discrete time sampling, inference for diffusion processes, stochastic differential equation

1. Introduction

Assume one wants to estimate the unknown parameter $\alpha := (\theta, \sigma)$ in the stochastic differential equation

$$\begin{cases} dX_t = a(X_t, \sigma) \, dW_t + b(X_t, \theta) \, dt \\ X_0 = x_0; \end{cases}$$

(1.1)

$\alpha$ belongs to $H$, a compact of $\mathbb{R}^2$; $W$ is the standard Brownian motion; $a$ and $b$ are known $\mathbb{R}$-valued functions defined on $\mathbb{R}^2$, which are assumed to be smooth enough to ensure the uniqueness in law of the solution to (1.1). We denote this law by $P_\alpha$. The observation consists of the discretized trajectory $(X_{t_i})_{0 \leq i \leq n}$ with $t_{ni} = i h_n$; $h_n$ is called the discretization step. We consider here the case when $h_n \to 0$ and $nh_n \to \infty$, assuming that the process $X$ has the ergodic property.

This estimation problem has been studied by several authors, see Florens-Zmirou (1989), Prakasa-Rao (1983, 1988) and Yoshida (1992). The main difficulty one has to face is that the transition density of the process $X$ is unknown, hence the likelihood function is unknown as well, and the maximum likelihood estimator which has the usual good properties, see Dacunha-Castelle & Florens-Zmirou (1986), is not a solution in practice. A common way to overcome this difficulty is to base the inference on a discretization of the continuous likelihood, see Yoshida (1992) and Genon-Catalot (1990). Assume that $\alpha$ is known in (1.1) and denote by $Q_\alpha$ the law of the process solution of $dY_t = a(Y_t, \sigma) \, dW_t$, if we dispose of the observation of the whole trajectory up to time $nh_n$, the continuous log likelihood log $\{dP_\alpha/dQ_\alpha(X \in [0, nh_n])\}$ is

$$\int_0^{nh_n} \frac{b(X_t, \theta)}{a^2(X_t, \sigma)} \, dX_t - \frac{1}{2} \int_0^{nh_n} \frac{b^2(X_t, \theta)}{a^2(X_t, \sigma)} \, dt.$$  

(1.2)

A discretization of (1.2) yields the contrast

$$\sum_{i=1}^{n} \left\{ \frac{b(X_{t_{i+1}}, \theta)}{a^2(X_{t_{i+1}}, \sigma)} (X_{t_i} - X_{t_{i+1}}) - \frac{1}{2} \frac{b^2(X_{t_{i+1}}, \theta)}{a^2(X_{t_{i+1}}, \sigma)} h_n \right\}.$$  

(1.3)
Another way to construct a contrast was introduced by Florens-Zmirou (1989) by using the approximate discrete-time scheme known as Euler–Maruyama’s. In the case when \( a(\cdot, \sigma) = \sigma \), Florens-Zmirou considered the approximate model obtained by discretizing (1.1)

\[
\begin{align*}
Z_{t_n} - Z_{t_{n-1}} &= h_n b(Z_{t_{n-1}}, \theta) + \sigma(W_{t_n} - W_{t_{n-1}}) & \text{for } i \geq 1 \\
Z_0 &= x_0.
\end{align*}
\]  

(1.4)

The log likelihood function of \( (Z_{t_n})_{0 \leq i \leq n} \) is given by

\[
C_i \left\{ \sum_{i=1}^n \frac{(Z_{t_i} - Z_{t_{i-1}} - h_n b(Z_{t_{i-1}}, \theta))^2}{\sigma^2} + n \log \sigma^2 \right\},
\]

where \( C \) is a constant which does not depend on \( \theta, \sigma \). A contrast for the estimation of \( \theta \) is derived from (1.5) by substituting \( (Z_{t_i})_{0 \leq i \leq n} \) with \( (X_{t_i})_{0 \leq i \leq n} \). The resulting contrast is

\[
C_i \left\{ \sum_{i=1}^n \frac{(X_{t_i} - X_{t_{i-1}} - h_n b(X_{t_{i-1}}, \theta))^2}{\sigma^2} + n \log \sigma^2 \right\}.
\]

(1.6)

Let us note that, up to a multiplicative constant and an additive random variable not depending on \( \theta, \sigma \), (1.6) coincides, when \( a(\cdot, \sigma) = \sigma \), with the discretized continuous log likelihood (1.3). Moreover, the estimator minimizing (1.6) was also studied by Prakasa-Rao (1983) as the least square estimator for \( \theta \).

As far as the joint estimation of \((\theta, \sigma)\) is concerned, Yoshida (1992) (in a d-dimensional framework) as well as Florens-Zmirou (1989) made two important assumptions. First, \( a(\cdot, \sigma) \) is multiplicative i.e. \( a(x, \sigma) = \sigma a(x) \). This allowed them to propose as an estimator of \( \theta \) the value that minimizes (1.6) and as an estimator of \( \sigma \) the one based on the quadratic variation. Secondly, they had to impose a condition on the rate at which the discretization step decreases to zero: Florens-Zmirou assumed that \( nh_n^d \to 0 \) (this is the “rapidly increasing experimental design” assumption of Prakasa-Rao (1983)) and Yoshida, through somewhat complicated corrections of (1.6), changed this assumption into the less restrictive \( nh_n^3 \to 0 \). Under these assumptions, the estimators they studied proved to be asymptotically efficient.

In this work, we exhibit a single contrast for the joint estimation of \((\theta, \sigma)\). Under the general assumption \( nh_n^d \to 0 \) for an arbitrary integer \( p \), we prove that the derived estimator is asymptotically efficient. We do not assume that the diffusion coefficient is multiplicative. In order to construct our contrast we begin by noticing that (1.6) would be the log likelihood of \((X_{t_i})_{0 \leq i \leq n}\) if the transition probability was a \( \mathcal{N}(h_n b(x, \theta), h_n \sigma^2) \) law. This leads us to consider a Gaussian approximation to the transition density. The most natural one is achieved through choosing its mean and variance to be the mean and variance of the transition density. Namely, if we set

\[
\begin{align*}
m(a, X_{t_{i+1}}) &= \mathbb{E}_a[X_{t_i}|X_{t_{i+1}}] \\
m_2(a, X_{t_{i+1}}) &= \mathbb{E}_a[(X_{t_i} - m(a, X_{t_{i+1}}))^2|X_{t_{i+1}}],
\end{align*}
\]

(1.7)

we approximate the transition density by the density of a \( \mathcal{N}(m(a, x), m_2(a, x)) \) law. Thus, we consider the contrast

\[
\sum_{i=1}^n \left\{ \frac{(X_{t_i} - m(\theta, X_{t_{i+1}}))^2}{m_2(\theta, X_{t_{i+1}})} + \log m_2(\theta, X_{t_{i+1}}) \right\}.
\]

(1.8)

But, as the transition density is unknown, in general there is no closed expression for \( m(\theta, x) \) and \( m_2(\theta, x) \), hence (1.8) is not explicit. Our method will consist in substituting in (1.8), closed approximations of \( m \) and \( m_2 \). In fact, the contrast (1.6) is an example of such an
approximation since it is known that \( m(\theta, x) \simeq x + h_\theta b(\theta, x) \) and \( m_2(\theta, x) \simeq h_\theta a^2(x, \sigma) \) i.e. \( h_\theta a^2 \) in the case when \( a(\cdot, \sigma) \equiv \sigma \).

To simplify the exposition, we assume that the process \( X \) is one-dimensional; nevertheless, the method we propose easily extends to a higher dimensional framework.

This paper consists of five parts. In section 2 we prepare notation and assumptions used later on, in section 3 we state our main result. The proofs are given in section 4. Finally, an appendix contains some useful technical results.

2. Notation—assumptions

We introduce the notation used in the rest of the paper and the assumptions.

1. \( \alpha = (\theta, \sigma) \) and \( \theta_0, \sigma_0 \) and \( a_0 \) denote the true values of \( \theta, \sigma \) and \( a \). \( c(x, \sigma) := a^2(x, \sigma) \).
2. \( C \) is a positive generic constant independent of all others variables (\( n, h_n, \alpha, \) etc. . . ). In the case when it depends on a fixed quantity, for instance an integer \( k \), we may write \( C_k \).
3. \( f^{(k)} \) denotes the \( k \)th derivative of \( f \) with respect to \( x \).
4. \( \partial_\theta := \frac{\partial}{\partial \theta}, \partial_\sigma := \frac{\partial}{\partial \sigma}, \partial^2_\theta := \frac{\partial^2}{\partial \theta^2}, \partial^2_\sigma := \frac{\partial^2}{\partial \sigma^2} \).
5. If \( f : \mathbb{R} \times H \to \mathbb{R} \), we denote by \( f_{i-1}(\alpha) \) the value \( f(X^i_{\tau_{i-1}}, \alpha) \); example: \( \partial_\alpha b_{i-1}(\theta_0) = \frac{\partial b}{\partial \alpha}(X^i_{\tau_{i-1}}, \theta_0) \). (The value of \( n \) will be clear from the context.)
6. For \( 0 < i \leq n, t^n_i = i h_n \). \( \mathcal{F}_i := \sigma(W_s, s \leq t^n_i) \).
7. If \( u_n \) is a \( \mathbb{R} \)-valued sequence, \( R \) denotes a function \( H \times \mathbb{R}^2 \to \mathbb{R} \) for which there exists a constant \( C \) such that

\[ R(\alpha, u_n, x) \leq u_0 C(1 + |x|)^C \] for all \( \alpha, n, x \).

We make the following assumptions

A1
There exists a constant \( C \) such that

\[ |a(x, \sigma_0) - a(y, \sigma_0)| + |b(x, \theta_0) - b(y, \theta_0)| \leq C|x - y|, \]

A1 implies the existence and uniqueness of a solution to (1.1) for the value \( a_0 \) of the parameter.

A2
The process \( X \) is ergodic for \( \alpha = \alpha_0 \). Let \( \mu_0(dx) \) be its invariant probability, and we assume that \( \mu_0(dx) \) has all moments finite.

A3
\( \inf_{x, \sigma} c(x, \sigma) > 0. \)

A4
For all \( p \geq 0, \sup E[|X_t|^p] < \infty. \)

A5 \([k]\)
(i) The functions \( a \) and \( b \) are continuously differentiable with respect to \( x \) up to order \( k \), for all \( \theta \) and \( \sigma \); their derivatives up to order \( k \) are of polynomial growth in \( x \), uniformly in \( \alpha \).

(ii) \(a(x, \cdot)\) (resp. \(b(x, \cdot)\)) and all its \(x\)-derivatives up to order \(k\), are three times differentiable with respect to \(\sigma\) (resp. \(\theta\)) for all \(x \in \mathbb{R}\). Moreover, these derivatives up to the third order with respect to \(\sigma\) (resp. \(\theta\)) are of polynomial growth in \(x\) uniformly in \(\alpha\).

Note that \(A5[k](i)\) for \(k \geq 1\) implies the existence and uniqueness of a solution to (1.1), we denote by \(P_\alpha\) the law of \(X\) for the value \(\alpha\) of the parameter. We set

\[ L_\alpha \text{ denotes the generator of the diffusion (1.1)} \]

\[ \text{if } f \in C^2(\mathbb{R}), \quad L_\alpha f(x) := b(x, \theta) \frac{\partial f}{\partial x}(x) + \frac{c(x, \alpha)}{2} \frac{\partial^2 f}{\partial x^2}(x), \quad L_0 = L_{\alpha_0}. \]

In this work, we consider the domain \(D\) of \(L_\alpha\) to be \(C^2(\mathbb{R})\). Hence, under the assumption \(A5[2(k-1)](i)\), we can define the \(k\)th iterate of \(L_\alpha\). We denote it by \(L_\alpha^k\). Its domain is \(C^{2k}(\mathbb{R})\). We set \(L_\alpha^0 = Id\).

We now make an identifiability assumption.

A6

\[ b(x, \theta) = b(x, \theta_0) \quad \text{for } \mu_0 \text{ a.s. all } x \Rightarrow \theta = \theta_0, \]

\[ a(x, \sigma) = a(x, \sigma_0) \quad \text{for } \mu_0 \text{ a.s. all } x \Rightarrow \sigma = \sigma_0. \]

3. The main result

3.1. Definition of the contrast functions

In order to define the contrast function, we first study expansions of \(m(\alpha, x)\) and \(m_2(\alpha, x)\) (see (1.7)). For that purpose, the main tool is lem. 1 from Florens-Zmirou (1989). We define \(r_l\) and \(\Gamma_l\) as follows. For \(l \geq 0\), under assumption \(A5[2(l-1)](i)\), let \(f(y) = y\), and

\[ r_l(h, x, \alpha) := \sum_{i=0}^l \frac{h_i}{i!} L_\alpha^i f(x). \]

The following lemma will be proved in section 4.

**Lemma 1**

We have

\[ m(\alpha, X_{r_{l+1}}) = r(h_n, X_{r_{l+1}}, \alpha) + R(\alpha, h_n^{l+1}, X_{r_{l+1}}), \]

where \(R\) was defined in section 2.

Now let \(g_{h_n,x,\alpha,l}(y) = (y - r(h_n, x, \alpha))^2\). For fixed \(x, y\) and \(\alpha\), \(g_{h_n,x,\alpha,l}\) is a polynomial in \(h_n\) of degree \(2l\). Let us denote by \(\overline{g}_{h_n,x,\alpha,l}\) the sum of its first terms up to degree \(l\). We have

\[ \overline{g}_{h_n,x,\alpha,l}(y) = \sum_{j=0}^l \frac{h_j}{j!} \overline{g}_{h_n,x,\alpha,l}(y) \]

\[ \begin{cases} \overline{g}_{h_n,x,\alpha,l}(y) = (y-x)^2 \\ \text{for } 1 \leq j \leq l, \quad \overline{g}_{h_n,x,\alpha,l}(y) = -2(y-x) \frac{L_j f(x)}{j!} + \sum_{r,s \geq 1, r + s = j} \frac{L_r f(x) L_s f(x)}{r! s!} \end{cases}. \]

From (3.2), we deduce that, under the assumption \(A5[2(l-1)](i)\), \(L_0^l \overline{g}_{h_n,x,\alpha,l}\) is well defined for \(r + j = l\), so we set

\[\text{© Board of the Foundation of the Scandinavian Journal of Statistics 1997.}\]
\[ \Gamma_l(h_n, x, \alpha) := \sum_{j=0}^{l} \frac{h_n}{l!} \frac{d^l}{dr^l} f(x, \alpha)(x). \]

\( \Gamma_l \) can be written as \( \Gamma_l(h_n, x, \alpha) := \sum_{i=0}^{l} h_n^i \gamma_i(x, \alpha) \). It is straightforward to see that \( \gamma_0(x, \alpha) = 0 \), \( \gamma_1(x, \alpha) = c(x, \sigma) \).

We set

\[ \tilde{\Gamma}_l(h_n, x, \alpha) := \frac{\sum_{i=2}^{l} h_n^i \gamma_i(x, \alpha)}{h_n c(x, \sigma)} \]

so that

\[ \Gamma_l(h_n, x, \alpha) := h_n c(x, \sigma) \{1 + \tilde{\Gamma}_l(h_n, x, \alpha)\}. \]  

We can now state the lemma that gives us the desired approximation of \( m_2 \).

**Lemma 2**

We have

\[ E_n[(X_{r} - r(h_n, X_{r-1}^+, \alpha))^2 | X_{r-1}^+] = h_n c_{r-1}(\sigma) (1 + \tilde{\Gamma}_l(h_n, X_{r-1}^+, \alpha)) + R(\alpha, h_n^{l+1}, X_{r-1}^+). \]  

(3.4)

It is worth noticing that we have a closed expression for \( r_l \) and \( \tilde{\Gamma}_l \) since it involves only the coefficients of the diffusion (1.1) and its derivatives with respect to \( x \).

We are now able to define the contrast. The natural way to do it would be to substitute \( m \) and \( m_2 \) in (1.8) with the expansions (3.1) and (3.4). However, in order to avoid technical difficulties (control of the denominator and logarithm), we will consider the expansions in \( h_n \) of \( (\Gamma_l)^{-1} \) and \( \log \Gamma_l \).

For the condition \( nh_n^p \to 0 \), we set \( k_0 = [p/2] \). Under the assumption A5[2k_0](i), we can define \( d_j \), resp. \( e_j \), as the coefficient of \( h_n^j \) in the Taylor expansion of \((1 + \tilde{\Gamma}_l(h_n, X_{r-1}^+, \alpha))^{-1}\), resp. \( \log(1 + \tilde{\Gamma}_l(h_n, X_{r-1}^+, \alpha)) \). We define the contrast

\[ l_{p,n}(\alpha) := \sum_{i=1}^{n} \left( \frac{(X_{r} - r(h_n, X_{r-1}^+, \alpha))^2}{h_n c_{r-1}(\sigma)} \left( \sum_{j=1}^{k_1} h_n^j d_j(X_{r-1}^+, \alpha) \right) \right) + \sum_{i=1}^{n} \left( \log c_{r-1}(\sigma) + \sum_{j=1}^{k_0} h_n^j e_j(X_{r-1}^+, \alpha) \right). \]  

(3.5)

Once again, let us stress the fact that \( l_{p,n} \) is explicit.

**Remark 1.** By assumptions A3 and A5 and by construction, for all \( j \leq k_0, d_j \) and \( e_j \) are three times differentiable with respect to \( \alpha \). Moreover, all their derivatives with respect to \( \alpha \) are of polynomial growth in \( x \) uniformly in \( \alpha \). Furthermore it is easy to check that

\[ d_1(x, \alpha) = -e_1(x, \alpha) = -\frac{\gamma_2(x, \alpha)}{c(x, \sigma)}. \]  

(3.6)

### 3.2 The main result

We denote by \( \hat{a}_{p,n} = (\hat{a}_{p,n}, \hat{a}_{p,n}) \) any of the solutions to \( l_{p,n}(\hat{a}_{p,n}) = \inf_{\alpha} l_{p,n}(\alpha) \). Our main result is theorem 1 below.
**Theorem 1**

Let $p$ be an integer and $k_0 = \lceil p/2 \rceil$. Under assumptions A1 to A4, A5 (2$k_0$) and A6, if $h_n \to 0$ and $nh_n \to \infty$, then

$$\hat{\alpha}_{p,n} \overset{p}{\to}_{n \to \infty} \alpha_0.$$  

If, in addition, $nh_n^p \to 0$ and $\alpha_0$ is in the interior of $H$

$$(\sqrt{n}h_n(\hat{\theta}_{p,n} - \theta_0), \sqrt{n}(\hat{\sigma}_{p,n} - \sigma_0)) \overset{p}{\to}_{n \to \infty} \mathcal{N}(0, K_0),$$

where

$$K_0 := \begin{bmatrix} \left( \int \frac{(\partial_0 b)^2}{c}(x, \alpha_0) \mu_0(dx) \right)^{-1} & 0 \\ 0 & 2 \left( \int \frac{(\partial_0 c)}{c}(x, \sigma_0) \mu_0(dx) \right)^{-1} \end{bmatrix}.$$  

**Remark 2.** The asymptotic efficiency for $\hat{\theta}_{p,n}$ is obtained since

$$\left( \int \frac{(\partial_0 b)^2}{c}(x, \alpha_0) \mu_0(dx) \right)$$

is the asymptotic Fisher information of the continuous time diffusion. (see e.g. Dacunha-Castelle & Florens-Zmirou, 1986). As for $\hat{\sigma}_{p,n}$, if we denote by $\Pi_{h_n}$ the transition density of the process $X$, an expansion of $\partial_0 \Pi_{h_n}/\Pi_{h_n}$ can be achieved through the representation of $\Pi_{h_n}$ given in Dacunha-Castelle & Florens-Zmirou (1986). From this expansion, the following inequality can be deduced (see Genon-Catalot & Jacod (1993)). Setting

$$I_{h_n}^\alpha(x) = \left( \int \frac{(\partial_0 c)}{c}(x, \sigma) \mu_0(dx) \right)^2 \Pi_{h_n}(\alpha, x, y) dy,$$

we have

$$\left| I_{h_n}^\alpha(x) - \frac{1}{2} \left( \frac{\partial_0 c}{c} \right)^2 (x, \sigma) \right| \leq R(\alpha, \sqrt{h_n}, x).$$

Hence

$$\left| \frac{1}{n} \sum_{i=1}^n I_{h_n}^\alpha(X_{t_{i-1}^n}) - \frac{1}{2} \frac{\partial_0 c}{c} (X_{t_{i-1}^n}, \sigma) \right| \leq \frac{1}{n} \sum_{i=1}^n R(\alpha, \sqrt{h_n}, X_{t_{i-1}^n}).$$

We conclude, thanks to lemma 8 below, that

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^n I_{h_n}^\alpha(X_{t_{i-1}^n}) = \frac{1}{2} \left( \frac{\partial_0 c}{c} \right)^2 (x, \sigma_0) \mu_0(dx).$$

**Remark 3.** For the joint estimation of $(\theta, \sigma)$, it is possible to adopt the same point of view as Yoshida (1992), and consider different contrasts for the estimation of $\theta$ and the estimation of $\sigma$. An adaptive procedure must then be used. Asymptotically efficient estimators under the condition $nh_n^p \to 0$ can be obtained. We refer to Kessler (1995) for the statement of these results. The proofs are not basically different from the ones made in this paper.
4. Proofs

Proof of Lemma 1. Lem. 1 from Florens-Zmirou (1989) provides us with the following expansion for every function $\phi$ in $C^{2l+1}$:

$$
\mathbb{E}_{t_0} [\phi(X_{t_1}^n) | \mathcal{F}_{t_1}^n] = \sum_{j=0}^{l} \frac{h_j^l}{j!} L_0^j \phi(X_{t_1}^n) + \int_0^{h_0} \cdots \int_0^{h_l} \mathbb{E}_{t_0} \left[ L_0^{l+1} \phi(X_{t_1+u_1}^n) | \mathcal{F}_{t_1}^n \right] du_1 \cdots du_{l+1}.
$$

To prove (3.1), it is enough to notice that, by assumption A5(i),

$$
|L_0^{l+1} f(X_{t_1+u_1}^n)| \leq C(1 + |X_{t_1+u_1}^n|)^3.
$$

Hence, (A.2) of lemma 6 (see appendix) applies and

$$
\int_0^{h_0} \cdots \int_0^{h_l} \mathbb{E}_{t_0} \left[ L_0^{l+1} f(X_{t_1+u_1}^n) | \mathcal{F}_{t_1}^n \right] du_1 \cdots du_{l+1} = R(\alpha, h_n^{l+1}, X_{t_1}^n).
$$

Proof of lemma 2. Is identical to the previous one.

Proof of theorem 1. Let us begin with the consistency of $\hat{\sigma}_{p,n}$. An application of lemmas 8 and 9 readily yields

$$
\frac{1}{n} l_{p,n}(\alpha) \xrightarrow{p=0} U(\alpha, \sigma_0) := \int \left[ \frac{c(x, \alpha_0)}{c(x, \alpha)} + \log c(x, \alpha) \right] \mu_0(dx) \tag{4.1}
$$

uniformly in $\alpha$. In order to prove that (4.1) implies the consistency of $\hat{\sigma}_{p,n}$, since the convergence in probability is equivalent to the existence, for any subsequence of a subsequence converging almost surely, we will consider that the convergence in (4.1) is almost sure and prove that it implies $\hat{\sigma}_{p,n} \rightarrow \sigma_0$ a.s. For a fixed $\omega$, thanks to the compactness of $H$, there exists a subsequence $n_k$ such that $(\theta_{p,n_k}, \hat{\sigma}_{p,n_k})$ tends to a limit $\alpha_\infty := (\theta_\infty, \sigma_\infty)$. Hence, (4.1) together with the continuity of $\sigma \rightarrow U(\alpha, \sigma_0)$, implies

$$
\frac{1}{n_k} l_{p,n_k}(\hat{\theta}_{p,n_k}, \hat{\sigma}_{p,n_k}) (\omega) \rightarrow U(\sigma_\infty, \sigma_0).
$$

But, by definition of $\hat{\sigma}_{p,n}$,

$$
\frac{1}{n_k} l_{p,n_k}(\hat{\theta}_{p,n_k}, \hat{\sigma}_{p,n_k}) \leq \frac{1}{n_k} l_{p,n_k}(\hat{\theta}_{p,n_k}, \sigma_0).
$$

So, by (4.1) again, we get $U(\sigma_\infty, \sigma_0) \leq U(\sigma_0, \sigma_0)$. On the other hand, since for all $y > 0$, $y_0 > 0$, $(y_0/y) + \log y \geq 1 + \log y_0$, we deduce from this inequality and assumption A6 that $\sigma_\infty = \sigma_0$. We have proved that any convergent subsequence of $\hat{\sigma}_{p,n}$ tends to $\sigma_0$, hence $\hat{\sigma}_{p,n} \xrightarrow{p=0} \sigma_0$ and we are done.

For the proof of $\hat{\theta}_{p,n} \xrightarrow{p=0} \theta_0$, we state the following lemma.

Lemma 3

$$
\frac{1}{n_h} l_{p,n}(\theta, \sigma) = \frac{1}{n_h} l_{p,n}(\theta_0, \sigma) \xrightarrow{p=0} \int \left[ \frac{(b(x, \theta_0) - b(x, \theta))^2}{c(x, \sigma)} \mu_0(dx) \right]
$$

$$
+ \int \left[ \frac{(c(x, \sigma_0)}{c(x, \sigma)} - 1 \right] \left\{ d_1(x, \theta, \sigma) - d_1(x, \theta_0, \sigma) \right\} \mu_0(dx)
$$

uniformly in $\alpha$.

In order to deduce the consistency of \( \hat{\theta}_{p,n} \) from lemma 3, the method is similar to the previous one. We know now that \( (\hat{\theta}_{p,n}, \hat{\alpha}_{p,n}) \) tends to \( (\theta_\infty, \alpha_0) \), hence

\[
\frac{1}{n_k} l_{p,n_1}(\hat{\theta}_{p,n_1}, \hat{\alpha}_{p,n_1}) = \frac{1}{n_k} l_{p,n_1}(\theta_0, \hat{\alpha}_{p,n_1}) \xrightarrow{n \to \infty} \mathbb{P} \left( \frac{b(x, \theta_0) - b(x, \theta_\infty)^2}{c(x, \alpha_0)} \mu_0(dx) \right).
\]

But

\[
\frac{1}{n_k} l_{p,n_1}(\hat{\theta}_{p,n_1}, \hat{\alpha}_{p,n_1}) - \frac{1}{n_k} l_{p,n_1}(\theta_0, \hat{\alpha}_{p,n_1}) \leq 0
\]

and we conclude by A6. Thus if we prove lemma 3, we are done.

**Proof of lemma 3.** Elementary computations yield

\[
I_{p,n}(\alpha) = \beta_{1,n}(\alpha) + \beta_{2,n}(\alpha) + \beta_{3,n}(\alpha)
\]

with

\[
\beta_{1,n}(\alpha) := \sum_{i=1}^{n} \left( \frac{(X_i^2 - r_{k_1}(h_n, X_i^{c_{i-1}}), \alpha_0))^2}{h_n c_{i-1}(\alpha)} \right) \left( 1 + \sum_{j=1}^{k_n} h_n^j d_j(X_i^{c_{i-1}}, \alpha) \right)
\]

\[
\beta_{2,n}(\alpha) := \sum_{i=1}^{n} \left( \frac{2(X_i^2 - r_{k_1}(h_n, X_i^{c_{i-1}}), \alpha_0)(r_{k_1}(h_n, X_i^{c_{i-1}}), \alpha_0) - r_{k_1}(h_n, X_i^{c_{i-1}}), \alpha))^2}{h_n c_{i-1}(\alpha)} \right) \left( 1 + \sum_{j=1}^{k_n} h_n^j d_j(X_i^{c_{i-1}}, \alpha) \right)
\]

\[
\beta_{3,n}(\alpha) := \sum_{i=1}^{n} \left( \frac{(r_{k_1}(h_n, X_i^{c_{i-1}}), \alpha_0) - r_{k_1}(h_n, X_i^{c_{i-1}}), \alpha))^2}{h_n c_{i-1}(\alpha)} \right) \left( 1 + \sum_{j=1}^{k_n} h_n^j d_j(X_i^{c_{i-1}}, \alpha) \right)
\]

+ \sum_{i=1}^{n} \left( \frac{\log c_{i-1}(\alpha) + \sum_{j=1}^{k_n} h_n^j e_j(X_i^{c_{i-1}}, \alpha)}{h_n c_{i-1}(\alpha)} \right). \tag{4.3}
\]

Lemma 9 applies and we deduce

\[
\frac{1}{n h_n} \beta_{1,n}(\theta, \sigma) - \frac{1}{n h_n} \beta_{1,n}(\theta_0, \sigma) \xrightarrow{n \to \infty} \mathbb{P} \left( \frac{\log c(x, \alpha_0)}{c(x, \alpha)} (d(x, \theta, \sigma) - d(x, \theta_0, \sigma)) \mu_0(dx) \right)
\]

uniformly in \( \sigma, \theta \). In addition, from (A.3) in lemma 7 below,

\[
r_j(h_n, X_i^{c_{i-1}}, \alpha_0) - r_j(h_n, X_i^{c_{i-1}}, \alpha) = h_n(b_{i-1}(\theta_0) - b_{i-1}(\theta)) + R(\alpha, h_n^2, X_i^{c_{i-1}}).
\]

Hence, lemma 10 yields \( (1/n h_n)^{\beta_{2,n}(\alpha)} \xrightarrow{n \to \infty} 0 \) uniformly in \( \alpha \), and, from lemma 8,

\[
\frac{1}{n h_n} \beta_{3,n}(\theta, \sigma) - \frac{1}{n h_n} \beta_{3,n}(\theta_0, \sigma) \xrightarrow{n \to \infty} \mathbb{P} \left( \frac{(b(x, \theta_0) - b(x, \theta))^2}{c(x, \sigma)} + e_1(x, \theta, \sigma) - e_1(x, \theta_0, \sigma) \right) \mu_0(dx)
\]

uniformly in \( \alpha \). Finally, since \( d_1 = -e_1 \), the lemma is proved.
We will now proceed with the proof of the asymptotic normality of \( \hat{\alpha}_{p,n} \). For that purpose we first have to compute \( \nabla l_{p,n}(\alpha) \).

If we denote
\[
\xi_{i,1}(\alpha) := \frac{(X_{t_i} - r_0(h_2, X_{t_{i-1}}, \alpha))^2}{h_n c_{i-1}(\sigma)} \left\{ f \frac{k_0}{\sum_{j=1}^{k_0} h^j \partial_0 d_j} \right\} + \frac{k_0}{\sum_{j=1}^{k_0} h^j \partial_0 e(X_{t_{i-1}}, \alpha)} \right\} + \frac{\partial_0 c_{i-1}(\sigma)}{c_{i-1}(\sigma)}
\]
\[
\xi_{i,2}(\alpha) := \frac{(X_{t_i} - r_0(h_2, X_{t_{i-1}}, \alpha))^2}{h_n c_{i-1}(\sigma)} \left\{ \left( 1 + \sum_{j=1}^{k_0} h^j d_j(X_{t_{i-1}}, \alpha) \right) \frac{\partial_0 c_{i-1}(\sigma)}{h_n c_{i-1}(\sigma)} \right\}
\]
\[
\xi_{i,3}(\alpha) := -2 \partial_0 r_0(h_2, X_{t_{i-1}}, \alpha)(X_{t_i} - r_0(h_2, X_{t_{i-1}}, \alpha)) \times \left\{ \left( 1 + \sum_{j=1}^{k_0} h^j d_j(X_{t_{i-1}}, \alpha) \right) \right\}
\]

we have
\[
\frac{\partial l_{p,n}(\alpha)}{\partial \alpha} = \sum_{i=1}^{n} \{ \xi_{i,1}(\alpha) + \xi_{i,2}(\alpha) + \xi_{i,3}(\alpha) \}.
\]

Let us now compute \( \partial l_{p,n}(\alpha)/\partial \theta \). We set
\[
\delta_{i,1}(\alpha) := -2 \partial_0 r_0(h_2, X_{t_{i-1}}, \alpha)(X_{t_i} - r_0(h_2, X_{t_{i-1}}, \alpha)) \times \left\{ \left( 1 + \sum_{j=1}^{k_0} h^j d_j(X_{t_{i-1}}, \alpha) \right) \right\}
\]
\[
\delta_{i,2}(\alpha) := \frac{(X_{t_i} - r_0(h_2, X_{t_{i-1}}, \alpha))^2}{h_n c_{i-1}(\sigma)} \left\{ f \frac{k_0}{\sum_{j=1}^{k_0} h^j \partial_0 d_j} \right\} + \frac{k_0}{\sum_{j=1}^{k_0} h^j \partial_0 e(X_{t_{i-1}}, \alpha)} \right\} + \frac{\partial_0 c_{i-1}(\sigma)}{c_{i-1}(\sigma)}
\]

Thus
\[
\frac{\partial l_{p,n}(\alpha)}{\partial \theta} = \sum_{i=1}^{n} \{ \delta_{i,1}(\alpha) + \delta_{i,2}(\alpha) \}.
\]

We will also need \( \nabla^2 l_{p,n}(\alpha) \).

From the decomposition (4.2) and the expressions (4.3) we easily deduce that
\[
\frac{\partial^2 l_{p,n}(\alpha)}{\partial \alpha^2} = \sum_{i=1}^{n} \frac{(X_{t_i} - r_0(h_2, X_{t_{i-1}}, \alpha)^2}{h_n} \left\{ \left( \partial_0^2 (c^{-1}) \right)_{i-1}(\alpha) + R(\alpha, h_2, X_{t_{i-1}}) \right\}
\]
\[
+ \sum_{i=1}^{n} \frac{(X_{t_i} - r_0(h_2, X_{t_{i-1}}, \alpha)^2}{h_n} \right\} R(\alpha, 1, X_{t_{i-1}})
\]
\[
+ \sum_{i=1}^{n} \{ (\partial_0^2 (\log c))_{i-1}(\alpha) + R(\alpha, h_2, X_{t_{i-1}}) \},
\]

\[
\frac{\partial^2 l_{p,n}(\alpha)}{\partial \theta^2} = \sum_{i=1}^{n} \frac{(X_t^* - r_k(h_n, X_t^{*,-1}, \alpha_0))^2}{c_{i-1}(\sigma)} \{ (\partial^2 d_1)_{i-1}(\alpha) + R(\alpha, h_n, X_t^{*,-1}) \}
\]
\[
+ \sum_{i=1}^{n} \frac{2(X_t^* - r_k(h_n, X_t^{*,-1}, \alpha_0))}{c_{i-1}(\sigma)} \{ -\partial^2 b_{i-1}(\theta) + R(\alpha, h_n, X_t^{*,-1}) \}
\]
\[
+ \sum_{i=1}^{n} h_n \partial^2 p \epsilon_1(X_t^{*,-1}, \alpha)
\]
\[
+ \sum_{i=1}^{n} h_n \left\{ \frac{2 \partial^2 b_{i-1}(\theta)(b_{i-1}(\theta) - b_{i-1}(\theta_0)) + (\partial_0 b_{i-1}(\theta))^2}{c_{i-1}(\sigma)} + R(\alpha, h_n, X_t^{*,-1}) \right\},
\]
\[(4.7)\]

and
\[
\frac{\partial^2 l_{p,n}(\alpha)}{\partial \theta \partial \sigma} = \sum_{i=1}^{n} \frac{(X_t^* - r_k(h_n, X_t^{*,-1}, \alpha_0))^2}{h_n} \{ R(\alpha, h_n, X_t^{*,-1}) \}
\]
\[
+ \sum_{i=1}^{n} \frac{2(X_t^* - r_k(h_n, X_t^{*,-1}, \alpha_0))}{h_n} \{ R(\alpha, h_n, X_t^{*,-1}) \}
\]
\[
+ \sum_{i=1}^{n} R(\alpha, h_n, X_t^{*,-1}).
\]
\[(4.8)\]

The proof of the asymptotic normality goes along a classical route (see for instance sect. 5a of Genon-Catalot & Jacod, 1993). By Taylor’s formula we have, if \( \hat{\alpha}_{p,n} \) is in the interior of \( H \),
\[
\int_0^1 \nabla^2 l_{p,n}(\alpha_0 + u(\hat{\alpha}_{p,n} - \alpha_0)) du \left( \frac{\hat{\theta}_{p,n} - \theta_0}{\sigma_{p,n} - \sigma_0} \right) = -\nabla l_{p,n}(\alpha_0).
\]
\[(4.9)\]
Let us note
\[
S_{p,n} := \left( \frac{\sqrt{n} h_n (\hat{\theta}_{p,n} - \theta_0)}{\sqrt{n} (\hat{\sigma}_{p,n} - \sigma_0)} \right), \quad L_{p,n} := \left( \begin{array}{c} -\frac{1}{\sqrt{n}} h_n \\ -\frac{1}{\sqrt{n}} \frac{\delta \theta l_{p,n}(\alpha_0)}{\delta \sigma_{p,n}(\alpha_0)} \end{array} \right)
\]
and
\[
M := \left( \begin{array}{cc} 1 & 0 \\ 0 & \frac{1}{\sqrt{n}} \end{array} \right).
\]

If we notice that
\[
M \nabla^2 l_{p,n} = \left( \begin{array}{cc} \frac{1}{\sqrt{n} h_n} & \frac{1}{\sqrt{n} h_n} \\ \frac{1}{\sqrt{n} h_n} & \frac{1}{\sqrt{n} h_n} \end{array} \right) \left( \begin{array}{c} \frac{\partial^2 l_{p,n}}{\partial \theta^2} \\ \frac{\partial^2 l_{p,n}}{\partial \theta \partial \sigma} \\ \frac{\partial^2 l_{p,n}}{\partial \theta \partial \sigma} \\ \frac{\partial^2 l_{p,n}}{\partial \sigma^2} \end{array} \right) = C_{p,n}(\alpha) \left( \begin{array}{cc} \sqrt{n} h_n & 0 \\ 0 & \sqrt{n} \end{array} \right)
\]
with
\[
C_{p,n}(\alpha) := \left( \begin{array}{cc} \frac{1}{h_n} \frac{\partial^2 l_{p,n}(\alpha)}{\partial \theta^2} & \frac{1}{n \sqrt{h_n}} \frac{\partial \theta l_{p,n}(\alpha)}{\partial \sigma} \\ \frac{1}{n \sqrt{h_n}} \frac{\partial \theta l_{p,n}(\alpha)}{\partial \sigma} & \frac{1}{n} \frac{\partial^2 l_{p,n}(\alpha)}{\partial \sigma^2} \end{array} \right).
\]

then, multiplying by $M$ in (4.9) gives
\[ \int_0^1 C_{p,n}(\alpha_0 + u(\hat{\alpha}_{p,n} - \alpha_0)) du = L_{p,n}. \] (4.10)

We deduce from this equality that, in order to prove the asymptotic normality of $\hat{\alpha}_{p,n}$ and hence to end the proof of theorem 1, it is enough to prove the following lemmas.

**Lemma 4**
If $h_n \to 0$ and $nh_n \to \infty$,

(a) \[ C_{p,n}(\alpha_0) \xrightarrow{p} \sup_{(x, \alpha_0) \in B} \left( \begin{array}{cc} 2 \left( \frac{\partial \alpha b}{c} \right)^2 (x, \alpha_0) u_0(dx) & 0 \\ 0 & 2 \left( \frac{\partial_2 c}{c} \right) (x, \alpha_0) u_0(dx) \end{array} \right) \]

(b) \[ \sup_{|\alpha| \leq \epsilon_n} |C_{p,n}(\alpha_0 + \alpha) - C_{p,n}(\alpha_0)| \xrightarrow{p} 0, \quad \text{where} \quad \epsilon_n \to 0. \]

**Lemma 5**
If $h_n \to 0$, $nh_n \to \infty$ and $nh_n^p \to 0$,
\[ L_{p,n} \xrightarrow{p} L \in \mathcal{F}(0, K'), \]

where
\[ K' = \left( \begin{array}{cc} 4 \left( \frac{\partial \alpha b}{c} \right)^2 (x, \alpha_0) u_0(dx) & 0 \\ 0 & 2 \left( \frac{\partial_2 c}{c} \right) (x, \alpha_0) u_0(dx) \end{array} \right). \]

**Proof of lemma 4.** We deduce from the equalities (4.6), (4.7) and (4.8) after an application of lemmas 8, 9 and 10, that
\[ \frac{1}{n} \frac{\partial^2 I_{p,n}(\alpha)}{\partial \alpha^2} \xrightarrow{p} \sup_{(x, \alpha_0) \in B} \left( \begin{array}{c} c(x, \sigma_0) \partial_2^2 (c^{-1})(x, \sigma) u_0(dx) + \int \partial_2^2 c \log c(x, \sigma) u_0(dx) \\ 0 \end{array} \right) (x, \alpha_0) u_0(dx) \]
\[ \text{uniformly in } \alpha, \quad (4.11) \]

\[ \frac{1}{nh_n} \frac{\partial^2 I_{p,n}(\alpha)}{\partial \theta^2} \xrightarrow{p} \sup_{(x, \alpha_0) \in B} \left( \begin{array}{c} \partial_\theta^2 d_1 + \partial_\theta^2 c_1 + 2 \left( \frac{\partial \alpha b}{c} \right)^2 (x, \alpha_0) u_0(dx) \\ + \int \partial_\theta^2 d_1 (x, \alpha) \left[ \frac{c(x, \sigma_0)}{c(x, \sigma)} - 1 \right] u_0(dx) \\ + 2 \left( \frac{\partial \alpha b}{c} \right) (x, \alpha) (b(x, \theta) - b(x, \theta_0)) u_0(dx) \end{array} \right) \]
\[ \text{uniformly in } \alpha, \quad (4.12) \]

\[ = \left( \begin{array}{c} 2 \left( \frac{\partial \alpha b}{c} \right)^2 (x, \alpha_0) u_0(dx) \\ + \int \partial_\theta^2 d_1 (x, \alpha) \left[ \frac{c(x, \sigma_0)}{c(x, \sigma)} - 1 \right] + 2 \left( \frac{\partial \alpha b}{c} \right) (x, \alpha) (b(x, \theta) - b(x, \theta_0)) \end{array} \right) u_0(dx). \]
since \(d_1 = -e_1\) and
\[
\frac{1}{n \sqrt{h_n}} \frac{\partial^2 I_{p,n}(\alpha)}{\partial \theta \partial \sigma} P_m^{(n)} n \rightarrow \infty 0 \quad \text{uniformly in } \alpha.
\]

This proves in particular the point (a) of the lemma. Moreover, as a consequence of assumptions A3 and A5, the limits (4.11) and (4.12) are continuous as functions of \(\alpha\), hence we deduce also from these convergences the point (b).

**Proof of Lemma 5.** The result is a consequence of a combination of th. 3.2 and 3.4 of Hall & Heyde (1980), if we prove

\[
\sum_{i=1}^{n} E_n \left[ \left( \frac{1}{\sqrt{n h_n}} (\hat{\delta}_{1,1}(\alpha_0) + \hat{\delta}_{1,2}(\alpha_0)) \right)^2 \right] P_m^{(n)} n \rightarrow \infty 0 \quad \text{(4.13)}
\]

\[
\sum_{i=1}^{n} E_n \left[ \frac{1}{n h_n} (\hat{\delta}_{1,1}(\alpha_0) + \hat{\delta}_{1,2}(\alpha_0))^2 \right] P_m^{(n)} n \rightarrow \infty K_1^{(n)} \quad \text{(4.14)}
\]

\[
\sum_{i=1}^{n} E_n \left[ \frac{1}{n} (\hat{\xi}_{1,1}(\alpha_0) + \hat{\xi}_{1,2}(\alpha_0) + \hat{\xi}_{1,3}(\alpha_0))^2 \right] P_m^{(n)} n \rightarrow \infty K_2^{(n)} \quad \text{(4.15)}
\]

\[
\sum_{i=1}^{n} E_n \left[ \frac{1}{n \sqrt{h_n}} (\hat{\delta}_{1,1}(\alpha_0) + \hat{\delta}_{1,2}(\alpha_0)) \right] P_m^{(n)} n \rightarrow \infty 0 \quad \text{(4.16)}
\]

\[
\sum_{i=1}^{n} E_n \left[ \frac{1}{\sqrt{n h_n}} (\hat{\delta}_{1,1}(\alpha_0) + \hat{\delta}_{1,2}(\alpha_0)) \right]^2 P_m^{(n)} n \rightarrow \infty 0. \quad \text{(4.17)}
\]

Let us begin by proving (4.14) to (4.17).

Since \(\partial_0 r_{k_4}(h_n, X_{t_{i-1}, \alpha}) = h_n \partial_0 b_{i-1}(\theta) + R(\alpha, h_n^2, X_{t_{i-1}, \alpha})\), and thanks to lemma 6, we deduce from the expressions (4.5)

\[
E_n[\hat{\delta}_{1,1}(\alpha_0)^2] P_m^{(n)} n \rightarrow \infty = \frac{4 \partial_0 b_{i-1}(\theta_0)^2}{\sigma_i^{(n)}(\sigma_0)} h_n c_{i-1}(\sigma_0) + R(\alpha, h_n^2, X_{t_{i-1}, \alpha})
\]

\[
E_n[\hat{\delta}_{1,2}(\alpha_0)^2] P_m^{(n)} n \rightarrow \infty = R(\alpha, h_n^2, X_{t_{i-1}, \alpha}).
\]

Hence

\[
\frac{1}{n h_n} \sum_{i=1}^{n} E_n[\hat{\delta}_{1,1}(\alpha_0) + \hat{\delta}_{1,2}(\alpha_0)]^2 P_m^{(n)} n \rightarrow \infty = \frac{4 \sum_{i=1}^{n} \partial_0 b_{i-1}(\theta_0)^2}{c_i^{(n)}(\sigma_0)} + \frac{1}{n} \sum_{i=1}^{n} R(\alpha, h_n, X_{t_{i-1}, \alpha}),
\]

and an application of lemma 8 yields (4.14).

**Proof of (4.15).** Since \(\partial_0 r_{k_4}(h_n, X_{t_{i-1}, \alpha}) = R(\alpha, h_n^2, X_{t_{i-1}, \alpha})\), from lemma 6 we have

\[
E_n[\hat{\xi}_{1,1}(\alpha_0)^2] P_m^{(n)} n \rightarrow \infty = R(\alpha, h_n^2, X_{t_{i-1}, \alpha})
\]

\[
E_n[\hat{\xi}_{1,2}(\alpha_0)^2] P_m^{(n)} n \rightarrow \infty = 2 \left( \frac{\partial_0 c_{i-1}(\sigma_0)}{\sigma_i^{(n)}(\sigma_0)} \right)^2 + R(\alpha, h_n, X_{t_{i-1}, \alpha}).
\]

Hence

\copyright Board of the Foundation of the Scandinavian Journal of Statistics 1997.
\[
\frac{1}{n} \sum_{i=1}^{n} E_{0_n}(\xi_{i,1}(a_0) + \xi_{i,2}(a_0) + \xi_{i,3}(a_0))^2 | \mathcal{F}_{i-1}^{n} = \frac{2}{n} \sum_{i=1}^{n} \left( \frac{\partial \alpha c_{i-1}(\alpha_0)}{c_{i-1}(\alpha_0)} \right)^2 + \frac{1}{n} \sum_{i=1}^{n} R(\alpha, h_n, X_{i-1}^n). \]

(4.15) is proved.

**Proof of (4.16).** From lemma 6,

\[
E_{0_n}(\xi_{i,1}(a_0)\delta_{i,1}(a_0) | \mathcal{F}_{i-1}^{n}) = R(\alpha, h^2_n, X_{i-1}^n),
\]

\[
E_{0_n}(\xi_{i,1}(a_0)\delta_{i,2}(a_0) | \mathcal{F}_{i-1}^{n}) = R(\alpha, h^2_n, X_{i-1}^n),
\]

\[
E_{0_n}(\xi_{i,2}(a_0)\delta_{i,1}(a_0) | \mathcal{F}_{i-1}^{n}) = R(\alpha, h_n, X_{i-1}^n),
\]

\[
E_{0_n}(\xi_{i,2}(a_0)\delta_{i,2}(a_0) | \mathcal{F}_{i-1}^{n}) = R(\alpha, h_n, X_{i-1}^n),
\]

\[
E_{0_n}(\xi_{i,3}(a_0)\delta_{i,1}(a_0) | \mathcal{F}_{i-1}^{n}) = R(\alpha, h^2_n, X_{i-1}^n),
\]

\[
E_{0_n}(\xi_{i,3}(a_0)\delta_{i,2}(a_0) | \mathcal{F}_{i-1}^{n}) = R(\alpha, h^3_n, X_{i-1}^n).
\]

Hence

\[
E_{0_n}(\xi_{i,1}(a_0) + \xi_{i,2}(a_0) + \xi_{i,3}(a_0)) | \mathcal{F}_{i-1}^{n}) = \sum_{i=1}^{n} R(\alpha, h_n, X_{i-1}^n).
\]

**Proof of (4.17).** Once again, lemma 6 yields

\[
E_{0_n}(\xi_{i,4}) | \mathcal{F}_{i-1}^{n}) = R(\alpha, h^4_n, X_{i-1}^n),
\]

\[
E_{0_n}(\xi_{i,4}) | \mathcal{F}_{i-1}^{n}) = R(\alpha, h^4_n, X_{i-1}^n),
\]

\[
E_{0_n}(\xi_{i,4}) | \mathcal{F}_{i-1}^{n}) = R(\alpha, h^4_n, X_{i-1}^n),
\]

Let us now proceed with the proof of (4.13). Recall that, from lemmas 1 and 2,

\[
\begin{align*}
E_{0_n}(X_{n}^n - r_k(h_n, X_{i-1}^n, a_0)) | \mathcal{F}_{i-1}^{n}) &= R(\alpha, h^{k+1}_n, X_{i-1}^n), \\
E_{0_n}(X_{n}^n - r_k(h_n, X_{i-1}^n, a_0)) | \mathcal{F}_{i-1}^{n}) &= h_n c_{i-1}(\alpha_0)(1 + \tilde{r}_{k+1}(h_n, X_{i-1}^n, a_0)) \\
&+ R(\alpha, h^{k+2}_n, X_{i-1}^n).
\end{align*}
\]

Hence

\[
E_{0_n}(\xi_{i,1}(a_0) | \mathcal{F}_{i-1}^{n}) = (1 + \tilde{r}_{k+1}(h_n, X_{i-1}^n, a_0)) \left\{ \sum_{j=1}^{k_n} h^{j}_n \partial \alpha c_j(X_{i-1}^n, a_0) \right\} + \sum_{j=1}^{k_n} h^{j}_n \partial \alpha c_j(X_{i-1}^n, a_0) + R(\alpha, h^{k+1}_n, X_{i-1}^n),
\]

\[
E_{0_n}(\xi_{i,2}(a_0) | \mathcal{F}_{i-1}^{n}) = \frac{\partial \alpha c_{i-1}(\alpha_0)}{c_{i-1}(\alpha_0)} \times \left\{ 1 - (1 + \tilde{r}_{k+1}(h_n, X_{i-1}^n, a_0))(1 + \sum_{j=1}^{k_n} h^{j}_n d_j(X_{i-1}^n, a_0)) \right\}
\]

Thus, by lemmas 11 and 12, 
\[ E_{\alpha_0}[\xi_{i,1}(\alpha_0)\mid \mathcal{G}_{i-1}^n] = R(\alpha, h_n^{k_{i,1}+1}, X_{i}^{c_i}) \]

Let us now consider \( 1/\sqrt{n}h_n \sum_{i=1}^{n} E_{\alpha_0}[\delta_{i,1}(\alpha_0) + \delta_{i,2}(\alpha_0)\mid \mathcal{G}_{i-1}^n] \). We have 
\[ E_{\alpha_0}[\delta_{i,1}(\alpha_0)\mid \mathcal{G}_{i-1}^n] = R(\alpha, h_n^{k_{i,1}+1}, X_{i}^{c_i}) \]

From lemma 12, we obtain 
\[ E_{\alpha_0}[\delta_{i,2}(\alpha_0)\mid \mathcal{G}_{i-1}^n] = R(\alpha, h_n^{k_{i,2}+1}, X_{i}^{c_i}) \]
We deduce that (4.13) holds if \[ \sqrt{n}h_n^{k_{i,1}+1} \to 0 \quad \text{and} \quad \sqrt{n}h_n^{k_{i,2}+1/2} \to 0, \]

i.e. if \[ \sqrt{n}h_n^{k_{i,s}+1/2} \to 0, \] which can be rewritten \[ nh_n^{2k_{i,s}+1} \to 0. \] But 

if \( p \) is even, \( k_0 = \frac{p}{2} \Rightarrow 2k_0 + 1 = p + 1 \)

if \( p \) is odd, \( k_0 = \frac{p - 1}{2} \Rightarrow 2k_0 + 1 = p. \)

Finally, (4.13) holds if \( nh_n^{p} \to 0. \) This ends the proof of lemma 5.

Remark 4. Of course, in a first approach, we began by considering two different orders of expansion for \( m(x, \alpha) \) and \( m_2(x, \alpha) \) in (1.8). Namely, we considered the contrast 
\[ \sum_{i=1}^{n} \frac{(X_i - r_k(h_n, X_{i-1}^{c_i}, \alpha))^2}{h_n c_{i-1}(\alpha)} \left( 1 + \sum_{j=1}^{k_2} h_n^{j} d_j(X_{i-1}^{c_i}, \alpha) \right) \]

\[ + \sum_{j=1}^{k_1} \log c_{i-1}(\alpha) + \sum_{j=1}^{k_2} h_n^{j} c_j(X_{i-1}^{c_i}, \alpha) \] (4.18)

and came up with the smallest values of \( k_1 \) and \( k_2 \) for our estimator to be asymptotically normal. This leads to the specification \( k_1 = k_2 = \lfloor p/2 \rfloor \) and thus to the contrast \( I_{p,x} \) in (3.5). Nevertheless, in the case when \( nh_n^{p} \to 0 \) and because of simplifications in the proof of lemma 5, the contrast 
\[ \tilde{I}_{2,n} = \sum_{i=1}^{n} \frac{(X_i - X_{i-1}^{c_i} - h_n b_{i-1}(\theta))^2}{h_n c_{i-1}(\sigma)} + \sum_{i=1}^{n} \log c_{i-1}(\alpha) \]
(which corresponds to $k_1 = 1$ and $k_2 = 0$ in (4.18)) also yields an asymptotically efficient estimator.

Acknowledgements
The author thanks Prof. J. Jacod for many helpful comments and discussions. This work was supported in part by the EC Human Capital and Mobility program.

References


Received September 1995, in final form June 1996

Mathieu Kessler, Laboratoire de Probabilités (URA 224), Université P. et M. Curie, Tour 56, 4, Place Jussieu, 75252 Paris Cedex 05, France.

Appendix
Lemma 6
For $k \geq 1$ and $t^n_i \geq t \geq t^n_i - 1$,

$$
\mathbb{E}_{\alpha_0}[|X_t - X_{t^n_i}|^k|\mathcal{G}_{t^n_i - 1}] \leq C_k|t - t^n_i|^{k/2}(1 + |X_{t^n_i}|)^C. \quad (A.1)
$$

If $f: \mathbb{R} \times H \to \mathbb{R}$ is of polynomial growth in $x$ uniformly in $\alpha$ then

$$
\mathbb{E}_{\alpha_0}[|f(X_{t^n_i})||\mathcal{G}_{t^n_i - 1}] \leq C_{t^n_i - t^n_i - 1}(1 + |X_{t^n_i}|)^C. \quad (A.2)
$$

Proof. It is easy to obtain

$$
\mathbb{E}_{\alpha_0}[|X_t - X_{t^n_i}|^k|\mathcal{G}_{t^n_i - 1}] \leq C_{t^n_i - t^n_i - 1}(1 + |X_{t^n_i}|)^k + C\mathbb{E}_{\alpha_0}\left[\int_{t^n_i}^t |X_s - X_{t^n_i}|^k \, ds|\mathcal{G}_{t^n_i - 1}\right].
$$

A straightforward application of the Gronwall–Belmann lemma yields (A.1). To prove (A.2), we write

$$
\mathbb{E}_{\alpha_0}[|f(X_t)||\mathcal{G}_{t^n_i - 1}] \leq \mathbb{E}_{\alpha_0}[(1 + |X_t|)^C|\mathcal{G}_{t^n_i - 1}] \leq C(1 + |X_{t^n_i}|)^C + \mathbb{E}_{\alpha_0}[(|X_t - X_{t^n_i}|^C|\mathcal{G}_{t^n_i - 1})].
$$

(A.1) ends the proof.
Lemma 7
For $l \geq 1$,
\begin{align}
    r(h_n, X^*_{t_{i-l}}, \alpha) &= X^*_{t_{i-l}} + h_n b_{i-l}(\theta) + R(\alpha, h_n^2, X^*_{t_{i-l}}) \quad (A.3) \\
    E_\alpha[(X^*_{t_{i-l}} - r(h_n, X^*_{t_{i-l}}, \alpha))^2 | \mathcal{F}_{t_{i-l-1}}] &= h_n c_{i-l}(\sigma_0) + R(\alpha, h_n^2, X^*_{t_{i-l}}) \quad (A.4) \\
    E_\alpha[(X^*_{t_{i-l}} - r(h_n, X^*_{t_{i-l}}, \alpha))^3 | \mathcal{F}_{t_{i-l-1}}] &= R(\alpha, h_n^2, X^*_{t_{i-l}}) \quad (A.5) \\
    E_\alpha[(X^*_{t_{i-l}} - r(h_n, X^*_{t_{i-l}}, \alpha))^4 | \mathcal{F}_{t_{i-l-1}}] &= 3 h_n^2 c_{i-l}(\sigma_0) + R(\alpha, h_n^3, X^*_{t_{i-l}}). \quad (A.6)
\end{align}

Remark. We easily deduce from (A.1) and (A.3) for $k \geq 1$ and $l \geq 0$,
\begin{equation}
    E_\alpha[(X^*_{t_{i-l}} - r(h_n, X^*_{t_{i-l}}, \alpha))^k | \mathcal{F}_{t_{i-l-1}}] \equiv C h_n^{k/2} (1 + |X^*_{t_{i-l}}|)^C. \quad (A.7)
\end{equation}

Proof of Lemma 7. (A.3) and (A.4) were already proved in lemmas 1 and 2. From (A.3) and (A.4), to prove (A.5) and (A.6), it is enough to prove
\begin{align}
    E_\alpha[(X^*_{t_{i-l}} - X^*_{t_{i-l}})^2 | \mathcal{F}_{t_{i-l-1}}] &= R(\alpha, h_n^2, X^*_{t_{i-l}}) \quad (A.8) \\
    E_\alpha[(X^*_{t_{i-l}} - X^*_{t_{i-l}})^3 | \mathcal{F}_{t_{i-l-1}}] &= 3 h_n^2 c_{i-l}(\sigma_0) + R(\alpha, h_n^3, X^*_{t_{i-l}}). \quad (A.9)
\end{align}
We only prove (A.9), the proof of (A.8) being identical and easier. Set $f_3(y) = (y - x)^4$.
According to the lem. 1 from Florens-Zmirou (1989), we have
\begin{align*}
    E_\alpha[(X^*_{t_{i-l}} - X^*_{t_{i-l}})^3 | \mathcal{F}_{t_{i-l-1}}] &= f_{X^*_{t_{i-l}}} (X^*_{t_{i-l}}) + h_n L_0 f_{X^*_{t_{i-l}}} (X^*_{t_{i-l}}) + \frac{h_n^2}{2} L_0^2 f_{X^*_{t_{i-l}}} (X^*_{t_{i-l}}) \\
    &\quad + \int_0^{h_n} \int_0^{h_n} \int_0^{h_n} E_\alpha[L_0 f_{X^*_{t_{i-l}}} (X^*_{t_{i-l}} + u_i)] \mathcal{F}_{t_{i-l-1}} du_1 \ldots du_3.
\end{align*}
But it is easy to check that $f_3(x) = 0$, $L_0 f_3(x) = 0$ and $L_0^2 f_3(x) = 6 c^2(x, \sigma_0)$. Thus, we deduce (A.9) as in the proof of lemma 1.

We state a few limit results in our particular framework, under rather strong assumptions that are satisfied in the situations we consider.

Lemma 8
Assume A1 to A6 and $h_n \to 0$ and $nh_n \to \infty$; let $f: \mathbb{R} \times H \to \mathbb{R}$, be such that $f$ is differentiable with respect to $x$ and $\alpha$, with derivatives of polynomial growth in $x$ uniformly in $\alpha$. Then
\begin{equation}
    \frac{1}{n} \sum_{i=1}^{n} f(X^*_{t_{i-l}}, \alpha) \xrightarrow{P_{\sigma_0}} \int f(x, \alpha) \mu_0(dx) \quad \text{uniformly in } \alpha.
\end{equation}

Proof. We first prove that the convergence holds for all $\alpha$, this point is, in our particular case, an easy consequence of the continuous ergodic theorem. Indeed, we have
\begin{equation}
    \frac{1}{nh_n} \int_0^{nh_n} f(X_s, \alpha) ds \xrightarrow{P_{\sigma_0}} \int f(x, \alpha) \mu_0(dx).
\end{equation}
But
\begin{equation}
    E_\alpha \left[ \frac{1}{nh_n} \sum_{i=1}^{n} \int_{t_{i-l}}^{t_i} |f(X_s, \alpha) - f(X^*_{t_{i-l}}, \alpha)| ds \right] \leq 
\end{equation}
\[
\frac{1}{nh_n} \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \left( \mathbb{E}_{\alpha_0}[(X_s - X_{t_{i-1}}^\ast)^2] \right)^{1/2} \left( \mathbb{E}_{\alpha_0} \left[ \left( \int f \frac{\partial}{\partial x} (X_s + u(X_s - X_{t_{i-1}}^\ast)) \, du \right)^2 \right] \right)^{1/2} \, ds.
\]

Using the assumption on \( \partial f/\partial x \), A4, (A.1) and (A.2), we deduce from this inequality that
\[
\frac{1}{nh_n} \int_{0}^{nh_n} f(X_s, \alpha) \, ds = \frac{1}{n} \sum_{i=1}^{n} f(X_{t_{i-1}}^\ast, \alpha) \overset{P}{\longrightarrow} 0.
\]

In order to prove the uniformity of the convergence, we will prove the tightness of the \((C(H), \| \cdot \|_\infty)\)-valued sequence \((s_n(\cdot)) = 1/n \sum_{i=1}^{n} f_i(\cdot)\). The tightness is implied by \(\sup_n \mathbb{E}_{\alpha_0} [\sup_x |\nabla s_n(x)|] < \infty\). This inequality is straightforward if we use the assumption on \(\nabla \alpha f\) and A5.

**Lemma 9**

Under assumptions A1 to A6, if \(nh_n \to \infty\), \(h_n \to 0\) and if \(f\) satisfies the assumptions of lemma 8, for \(l \geq 0\)
\[
\frac{1}{nh_n} \sum_{i=1}^{n} f_{l-1}(\alpha)(X_{t_i} - r_l(h_n, X_{t_{i-1}}^\ast, \alpha_0))^2 \overset{P}{\longrightarrow} 0
\]
uniformly in \(\alpha\).

**Proof.** Let us prove that the convergence holds for all \(\alpha\). Let
\[
\xi_n^\alpha(\alpha) := \frac{1}{nh_n} f_{l-1}(\alpha)(X_{t_i} - r_l(h_n, X_{t_{i-1}}^\ast, \alpha_0))^2.
\]

Thanks to the lemma 7, we have
\[
\begin{aligned}
\left\{ \sum_{i=1}^{n} \mathbb{E}_{\alpha_0} [\xi_n^\alpha(\alpha) | \mathcal{F}_{n-1}^n] = \frac{1}{n} \sum_{i=1}^{n} (f_{l-1}(\alpha)c_{l-1}(\sigma_0) + R(\alpha, h_n, X_{t_{i-1}}^\ast)) \\
\sum_{i=1}^{n} \mathbb{E}_{\alpha_0} [\xi_n^\alpha(\alpha)^2 | \mathcal{F}_{n-1}^n] = \frac{1}{n} \sum_{i=1}^{n} R(\alpha, 1, X_{t_{i-1}}^\ast).
\end{aligned}
\]

From lemma 8, we deduce
\[
\begin{aligned}
\left\{ \sum_{i=1}^{n} \mathbb{E}_{\alpha_0} [\xi_n^\alpha(\alpha) | \mathcal{F}_{n-1}^n] \overset{P}{\longrightarrow} 0 \right. \right. \mathbb{E}_{\alpha_0} [f(x, \alpha)c(x, \sigma_0)\mu_0(dx) \\
\left. \left. \sum_{i=1}^{n} \mathbb{E}_{\alpha_0} [\xi_n^\alpha(\alpha)^2 | \mathcal{F}_{n-1}^n] \overset{P}{\longrightarrow} 0 \right. \right. \mathbb{E}_{\alpha_0} [f(x, \alpha)c(x, \sigma_0)\mu_0(dx)
\end{aligned}
\]
Lem. 9 from Genon-Catalot & Jacod (1993) applies and the convergence for all \(\alpha\) is proved, the proof of uniformity is the same as for Lemma 8.

**Lemma 10**

Under assumptions A1 to A6, if \(nh_n \to \infty\), \(h_n \to 0\) and if \(f\) satisfies the assumptions of lemma 8 for \(l \geq 1\)
\[
\frac{1}{nh_n} \sum_{i=1}^{n} f_{l-1}(\alpha)(X_{t_i} - r_l(h_n, X_{t_{i-1}}^\ast, \alpha_0)) \overset{P}{\longrightarrow} 0
\]
uniformly in \(\alpha\).
Proof. Let
\[
Z_n(\alpha) = \frac{1}{nh_n} f_{i-1}(\alpha)(X_{t_i} - r_i(h_n, X_{t_i-1}, a_0)).
\]

\[
\sum_{i=1}^{n} E_{a_0} [Z_n^{2}(\alpha) | \mathcal{F}_{i-1}^{n}] = \frac{1}{nh_n} \sum_{i=1}^{n} f_{i-1}(\alpha) R(\alpha, h_n^2, X_{t_i-1})
\]

from (A.3),
\[
\sum_{i=1}^{n} E_{a_0} [Z_n^{2}(\alpha) | \mathcal{F}_{i-1}^{n}] = \frac{1}{n^2 h_n^2} \sum_{i=1}^{n} \left( f_{i-1}^{2}(\alpha) h_n c_{i-1} (\sigma_0) + R(\alpha, h_n^2, X_{t_i-1}) \right)
\]

from (A.4).

Hence lem. 9 from Genon-Catalot & Jacod (1993) proves the convergence for all \( \alpha \). The criterion for tightness we will use is different from the one in lemmas 8 and 9.

According to th. 20 in app. 1 from Ibragimov & Has’minskii (1981), if we prove the following inequalities, we are done.

\[
E_{a_0} \left( \sum_{i=1}^{n} \frac{Z_n^2(\alpha)}{C} \right)^2 \leq C \quad \forall \alpha
\]

(A.10)

\[
E_{a_0} \left( \sum_{i=1}^{n} \frac{Z_n^2(\alpha_1) - Z_n^2(\alpha_2)}{C} \right)^2 \leq C |\alpha_1 - \alpha_2|^2 \quad \forall \alpha_1 \forall \alpha_2.
\]

(A.11)

We will only prove (A.10), the proof of (A.11) is identical.

\[
E_{a_0} \left( \sum_{i=1}^{n} \frac{Z_n^2(\alpha)}{C} \right)^2 = \frac{1}{n^2 h_n^2} \sum_{i=1}^{n} f_{i-1}^{2}(\alpha)(X_{t_i} - r_i(h_n, X_{t_i-1}, a_0))^2 + 2 \frac{1}{n^2 h_n^2} \sum_{i=1}^{n} f_{i-1}(\alpha) f_{j-1}(\alpha)(X_{t_i} - r_i(h_n, X_{t_i-1}, a_0))(X_{t_j} - r_j(h_n, X_{t_j-1}, a_0)).
\]

By (A.4), assumption A4 and the assumption on \( f \) we prove that the first term is a \( O(1/nh_n) \). As for the second term, if we take the conditional expectation with respect to \( \mathcal{F}_{i-1}^{n} \), we obtain

\[
2 \frac{1}{n^2 h_n^2} \sum_{i=1}^{n} \sum_{1 \leq j < i \leq n} f_{i-1}(\alpha) f_{j-1}(\alpha)(X_{t_i} - r_i(h_n, X_{t_i-1}, a_0))(X_{t_j} - r_j(h_n, X_{t_j-1}, a_0)) R(\alpha, h_n^2, X_{t_{i-1}}).
\]

Let us now apply the Cauchy–Schwarz inequality, (A.4) and A4 to obtain

\[
E_{a_0} \left( \sum_{i=1}^{n} \frac{Z_n^2(\alpha)}{C} \right)^2 = O \left( \sqrt{h_n + \frac{1}{nh_n}} \right).
\]

We now state two useful and elementary lemmas.

Recall that \( d_i(x, \alpha) \), resp. \( e_j(x, \alpha) \), is the coefficient of \( h_i^j \) in the Taylor expansion of \( (1 + \Gamma_{k_0+1}(x, \alpha))^{-1} \), resp. \( \log (1 + \Gamma_{k_0+1}(x, \alpha)) \). Since we will consider \( d_i(x, \alpha) \), \( e_j(x, \alpha) \) and \( \Gamma_{k_0+1}(x, \alpha) \) for fixed \( x \) and \( \alpha \), we will write \( d_i, e_j \) and \( \Gamma_{k_0+1} \) respectively.

**Lemma 11**

We consider the polynomial in \( h_n \)

\[
\{1 + \tilde{\Gamma}_{k_0+1}\} \left\{1 + \sum_{j=1}^{k_0} dj_h^l \right\}.
\]

For this polynomial, the coefficients of \( h_n^l \) for \( 1 \leq l \leq k_0 \) are null.

**Proof.** We have
\[
\left\{1 + \sum_{j=1}^{k_0} dj_h^l \right\} = \frac{1}{1 + \tilde{\Gamma}_{k_0+1}} + O(h_n^{k_0+1})
\]

In the same way, we obtain the following lemma.

**Lemma 12**
We consider the polynomial in \( h_n \)
\[
(1 + \tilde{\Gamma}_{k_0+1}) \left\{ \sum_{j=1}^{k_0} \partial d_j h_n^l \right\} + \left\{ \sum_{j=1}^{k_0} \partial c_j h_n^l \right\},
\]
where \( \partial \) denotes the operator \( \partial_0 \) or \( \partial_a \). For this polynomial, the coefficients of \( h_n^l \), \( 0 \leq l \leq k_0 \), are null.

**Proof.** We have \( \partial \log (1 + \tilde{\Gamma}_{k_0+1}) = -(1 + \tilde{\Gamma}_{k_0+1}) \partial (1 + \tilde{\Gamma}_{k_0+1})^{-1} \). On the other hand, under our assumptions, we easily have:
\[
\partial \log (1 + \tilde{\Gamma}_{k_0+1}) = \sum_{j=1}^{k_0} \partial c_j h_n^l + O(h_n^{k_0+1})
\]
\[
\partial (1 + \tilde{\Gamma}_{k_0+1})^{-1} = \sum_{j=1}^{k_0} \partial d_j h_n^l + O(h_n^{k_0+1}).
\]