ESAIM: Probability and Statistics December 2000, Vol. 4, 205–227
URL: http://www.emath.fr/ps/

DISCRETE SAMPLING OF AN INTEGRATED DIFFUSION PROCESS AND PARAMETER ESTIMATION OF THE DIFFUSION COEFFICIENT

Arnaud Gloter[1]

Abstract. Let (X_t) be a diffusion on the interval (l, r) and Δ_n a sequence of positive numbers tending to zero. We define J_i as the integral between $i\Delta_n$ and $(i+1)\Delta_n$ of X_s. We give an approximation of the law of $(J_0, \dots, J_{n-1})$ by means of a Euler scheme expansion for the process (J_i). In some special cases, an approximation by an explicit Gaussian ARMA(1,1) process is obtained. When $\Delta_n = n^{-1}$, we deduce from this expansion estimators of the diffusion coefficient of X based on (J_i). These estimators are shown to be asymptotically mixed normal as n tends to infinity.

AMS Subject Classification. 62F12, 62M09.

Received January 6, 1999. Revised April 5, 2000.

1. Introduction

Consider the stochastic process $I_t = \int_0^t X_s ds$ where X is a one-dimensional diffusion process given by

$$dX_t = a(X_t)dB_t + b(X_t)dt, \quad X_0 = \eta, \tag{1}$$

B is a standard Brownian motion and η is a random variable independent of B. The process I_t appears naturally in many problems studied recently.

First, the two-dimensional process (I_t, X_t) solves the system:

$$\begin{cases} dI_t = X_t \; dt \\ dX_t = a(X_t)dB_t + b(X_t)dt \end{cases} \tag{2}$$

which is a special case of two-dimensional model without noise in the first equation. In Lefebvre [10], the component I_t is used for modelling a non Markovian process.

Second, integrals of stochastic processes play an important role in finance. For instance, in the continuous stochastic volatility models, introduced by Hull and White [6], the logarithm of the stock price, Y_t, is modelled by:

$$\begin{cases} dY_t = \rho(X_t)dt + \sqrt{X_t}dW_t, \\ dX_t = a(X_t)dB_t + b(X_t)dt \end{cases} \tag{3}$$

Keywords and phrases: Diffusion processes, discrete time observation, hidden markov model.

[1] Université de Marne-la-Vallée, Équipe d'Analyse et de Mathématiques Appliquées, 5 boulevard Descartes, Champs-sur-Marne, 77454 Marne-la-Vallée Cedex 2, France; e-mail: gloter@math.univ-mlv.fr

where X_t is a positive diffusion process called the volatility of the stock price. The quadratic variation of Y_t is $I_t = \int_0^t X_s ds$. This integrated volatility plays a crucial role in finance. For instance, to derive option prices formulae, it is necessary to compute the distribution of I_t (see *e.g.* Leblanc [9], see also Genon-Catalot *et al.* [4], Barndorff-Nielsen and Shephard [1]).

Now, the exact distribution of the integrated process (I_t) is generally not explicit except for very few models.

In this paper, our first concern is the study of the distribution of a discrete sampling of (I_t). Then, we have in view statistical applications to the inference of unknown parameters in the coefficients of model (1) from such an observation. Data may be obtained from option prices and their associated implied volatilities (see *e.g.* Pastorello *et al.* [11]).

We focus on the estimation of unknown parameters in the diffusion coefficient without knowledge of the drift in the spirit of Genon-Catalot and Jacod [3]: we shall assume that I is discretely observed on a fixed length time interval with sampling interval tending to 0, and no ergodicity assumptions will be required on model (1).

For $i \geq 0$, let us set $J_i = \int_{i\Delta}^{(i+1)\Delta} X_s ds = I_{(i+1)\Delta} - I_{i\Delta}$. We study here the joint distribution of (J_i).

The case of X a stationary Ornstein-Uhlenbeck process,

$$dX_t = \mu X_t dt + \sigma dB_t$$

has been investigated in a previous work (see Gloter [5]). Explicit computations, for this special model, yield that (J_i) is a Gaussian ARMA(1,1) process.

There are difficulties in dealing with a general case, and our approach, which is now classical in the statistics of diffusion processes from discrete observations, is to obtain an approximation when the sampling interval $\Delta = \Delta_n$ depends on n and tends to 0 as $n \to \infty$.

Indeed, under the assumption $\Delta_n \to 0$, the Euler scheme,

$$X_{(i+1)\Delta_n} \simeq X_{i\Delta_n} + b(X_{i\Delta_n})\Delta_n + a(X_{i\Delta_n})(B_{(i+1)\Delta_n} - B_{i\Delta_n}) \tag{4}$$

provides an approximation of the distribution of $(X_{i\Delta_n}, i \leq n)$: conditionally on $(X_{j\Delta_n}, j \leq i)$, $X_{(i+1)\Delta_n}$ is almost Gaussian with mean $X_{i\Delta_n} + \Delta_n b(X_{i\Delta_n})$ and variance $\Delta_n a^2(X_{i\Delta_n})$. This approximation has been fruitfully used for statistical applications (see *e.g.* Genon-Catalot and Jacod [3] for the estimation of the diffusion coefficient and Kessler [8] for estimation of drift and diffusion coefficients under ergodicity assumptions).

Here, we obtain expansions for $J_i^n = \int_{i\Delta_n}^{(i+1)\Delta_n} X_s ds = I_{(i+1)\Delta_n} - I_{i\Delta_n}$. For simplicity, we omit the superscript n and simply write $J_i = J_i^n$. Noting that $\frac{J_i}{\Delta_n}$ is close to $X_{i\Delta_n}$, we should in particular answer the following question: can we approximate the law of $(J_0, \ldots, J_{n-1})$, by the law of a Markov process? Actually, we prove that $(\frac{J_i}{\Delta_n})$ is different from $(X_{i\Delta_n})$, and this has consequences for the statistical inference.

The paper is organized as follows. Assumptions on the model are presented in Section 2.1.

Then, in Section 2.2, we give our asymptotic expansions. First we compare $\frac{J_i}{\Delta_n}$ and $X_{i\Delta_n}$ (Prop. 2.2): the difference is of order $\Delta_n^{\frac{1}{2}}$ and we give an expansion for the difference $\frac{J_i}{\Delta_n} - X_{i\Delta_n}$.

But, this expansion is not enough for the statistical applications. So, in Theorem 2.3, we give an expansion of $\frac{J_{i+1}}{\Delta_n} - \frac{J_i}{\Delta_n}$:

$$\frac{J_{i+1}}{\Delta_n} - \frac{J_i}{\Delta_n} - b\left(\frac{J_i}{\Delta_n}\right)\Delta_n = a(X_{i\Delta_n})(\xi_{i,n} + \xi'_{i+1,n})\Delta_n^{\frac{1}{2}} + \varepsilon_{i,n}$$

where $\varepsilon_{i,n}$ is a remainder term and the vector $(\xi_{i,n}+\xi'_{i+1,n})_{0\leq i\leq n-1}$ is centered Gaussian with the same covariance matrix as a MA(1) process. Let us notice that the analogous term in (4), $\left(\frac{B_{(i+1)\Delta_n} - B_{i\Delta_n}}{\sqrt{\Delta_n}}\right)_{0\leq i\leq n-1}$ is Gaussian with independent indentically distributed components. In particular, when a is constant, our expansion provides an approximation by a Gaussian ARMA(1,1) process. Hence, $(\frac{J_i}{\Delta_n})$ is really different from $X_{i\Delta_n}$.

Section 3 is devoted to the estimation, based on the observation of $(J_i)_{0\leq i\leq n-1}$, of an unknown parameter σ_0, appearing in the diffusion coefficient $a(x) = a(x, \sigma_0)$.

Recall that when $(X_{\frac{i}{n}})_{i=0,\ldots,n}$ is observed, Dohnal [2] shows that the statistical problem of estimating σ satisfies the LAMN property. More precisely, his result implies that an asymptotically efficient estimator, σ_n^{eff}, of σ_0 must satisfy

$$n^{\frac{1}{2}}(\sigma_n^{\mathrm{eff}} - \sigma_0) \xrightarrow{\mathcal{D}} Z,$$

where Z is conditionally on X centered Gaussian with variance $\frac{1}{2}\left(\int_0^1 \left(\frac{\frac{\partial}{\partial\sigma}a(X_s,\sigma_0)}{a(X_s,\sigma_0)}\right)^2 ds\right)^{-1}$. The construction of such efficient estimators is possible and is essentialy based on an approximation of the quadratic variation of X.

Hence, in Section 3, we choose $\Delta_n = n^{-1}$ and first study the quadratic variation of our observed process (J_i). A surprising consequence of expansions of Section 2.2, is that, $\sum_{i=0}^{n-2}\left(\frac{J_{i+1}}{\Delta_n} - \frac{J_i}{\Delta_n}\right)^2$ is not a consistent estimator of the quadratic variation of (X_t) but converges to $\frac{2}{3}\int_0^1 a^2(X_s)ds$ (see Sect. 3.1).

In Section 3.2, we derive statistical applications for the estimation of the parameter σ_0, based on the observation of (J_i). We introduce the following contrast, modification of the contrast used in Genon-Catalot and Jacod [3],

$$U_n(\sigma) = \frac{3}{2}\sum_{i=0}^{n-2}\frac{(nJ_{i+1} - nJ_i)^2}{a^2(nJ_i,\sigma)} + n^{-1}\sum_{i=0}^{n-2}\log(a^2(nJ_i,\sigma)),$$

and denote by $\overline{\sigma}_n = \operatorname{arginf}_\sigma U_n(\sigma)$ the associated minimum contrast estimator. We show that this estimator is consistent and asymptotically mixed normal:

$$n^{\frac{1}{2}}(\overline{\sigma}_n - \sigma_0) \xrightarrow{\mathcal{D}} S,$$

where S is conditionally on X centered Gaussian with variance $\frac{9}{16}\left(\int_0^1 \left(\frac{\frac{\partial}{\partial\sigma}a(X_s,\sigma_0)}{a(X_s,\sigma_0)}\right)^2 ds\right)^{-1}$.

Section 4 is devoted to the extension of our results when we replace the uniform mean $\frac{1}{\Delta}\int_{i\Delta}^{(i+1)\Delta} X_s ds$ by the more general form $\frac{1}{\Delta}\int_{i\Delta}^{(i+1)\Delta} X_s\phi(\frac{s-i\Delta}{\Delta})ds$ with ϕ a non negative, measurable function defined on $[0,1]$ and such that $\int_0^1 \phi(s)ds = 1$. This correspond to the discrete observation of the convoluted signal $X \star \psi^\Delta$, where $\psi^\Delta(x) = \frac{1}{\Delta}\phi(\frac{-x}{\Delta})$ and thus model the mesurement of X through an instrument.

In Section 5, we give examples of classical models satisfying our set of assumptions.

2. Asymptotic expansions for small sampling interval

2.1. **Assumptions on the diffusion model**

We assume that (X_t) is the one dimensional diffusion process defined by:

$$dX_t = a(X_t)dB_t + b(X_t)dt, \quad X_0 = \eta \tag{5}$$

where $(B_t)_{t\geq 0}$ and η are defined on a probability space $(\Omega, \mathcal{G}, P)$: $(B_t)_{t\geq 0}$ is a standard Brownian motion, η is a random variable independent of (B_t).

Let $-\infty \leq l < r \leq \infty$ and consider the following assumptions:

(A1) Equation (5) admits a unique strong solution taking value in (l,r); a and b are two real valued functions defined on (l,r) with continuous second derivatives on (l,r).

Let us consider two positive measurable functions, $\mathcal{B}_l$ and $\mathcal{B}_r$ defined on (l, r) satisfying the following property: for all five non negative real numbers $\alpha, \beta, \alpha', \beta', p$, there exists a constant c such that for all $x \in (l, r)$:

$$(\mathcal{B}_l^{\alpha}(x) + \mathcal{B}_r^{\beta}(x)) \times (\mathcal{B}_l^{\alpha'}(x) + \mathcal{B}_r^{\beta'}(x)) \leq c(\mathcal{B}_l^{\alpha+\alpha'}(x) + \mathcal{B}_r^{\beta+\beta'}(x))$$
$$(\mathcal{B}_l^{\alpha}(x) + \mathcal{B}_r^{\beta}(x))^p \leq c(\mathcal{B}_l^{p\alpha}(x) + \mathcal{B}_r^{p\beta}(x)).$$

These functions are introduced to bound the growth of other functions near the boundaries l and r. For instance, if $-\infty < l < \infty$ (respectively $-\infty < r < \infty$) we may take $\mathcal{B}_l(x) = 1 + \frac{1}{x-l}$ (respectively $\mathcal{B}_r(x) = 1 + \frac{1}{r-x}$). And if $l = -\infty$ (resp. $r = \infty$), we may take $\mathcal{B}_l(x) = 1 + |x|$ (resp. $\mathcal{B}_r(x) = 1 + |x|$).

(A2) There exist non negative constants c, $\alpha_1, \alpha_2, \beta_1, \beta_2$ such that, for all $x \in (l, r)$,

$$|a(x)| + |b(x)| \leq c(1 + \mathcal{B}_r(x)),$$
$$|a'(x)| \leq c(\mathcal{B}_l^{\alpha_1}(x) + \mathcal{B}_r^{\alpha_1}(x)), \quad |a''(x)| \leq c(\mathcal{B}_l^{\alpha_2}(x) + \mathcal{B}_r^{\alpha_2}(x)),$$
$$|b'(x)| \leq c(\mathcal{B}_l^{\beta_1}(x) + \mathcal{B}_r^{\beta_1}(x)), \quad |b''(x)| \leq c(\mathcal{B}_l^{\beta_2}(x) + \mathcal{B}_r^{\beta_2}(x)).$$

Now, let Δ_n be a sequence of positive numbers with $\Delta_n \to 0$ as $n \to \infty$ and assume (for convenience) that $\Delta_n \leq 1$ for all n.

We set $\mathcal{G}_t = \sigma(B_s,\ s \leq t;\ \eta)$ and $\mathcal{G}_i^n = \mathcal{G}_{i\Delta_n}$. Below, the values of the constant c may change from one line to another but never depends on i or n.

(A3) There exists a positive constant K_l, such that:
$\forall\ k \in [0, K_l)$, $\exists c$, $\forall i, n\ (i \leq n)$, $E\left(\sup_{s\in[i\Delta_n,(i+1)\Delta_n]} \mathcal{B}_l^k(X_s) \mid \mathcal{G}_i^n\right) \leq c\mathcal{B}_l^k(X_{i\Delta_n})$
$\forall\ k \in [0, \infty)$, $\exists c$, $\forall i, n\ (i \leq n)$, $E\left(\sup_{s\in[i\Delta_n,(i+1)\Delta_n]} \mathcal{B}_r^k(X_s) \mid \mathcal{G}_i^n\right) \leq c\mathcal{B}_r^k(X_{i\Delta_n})$.

Assumption (A3) means that the diffusion will not approach too abruptly the end points l and r.

In Section 5, we check this assumption for some classical models. The reason why our condition is not symmetric in l and r appears there. Indeed we shall see that for all models, except one (see Sect. 5.2.2), we have (A3) with $K_l = \infty$.

2.2. **Expansions for means of the diffusion process over small intervals**

$$\text{Let } J_i = J_i^n = \int_{i\Delta_n}^{(i+1)\Delta_n} X_s ds,$$

and consider the following random variables which will appear in our expansions,

$$\xi_{i,n} = \frac{1}{\Delta_n^{\frac{3}{2}}} \int_{i\Delta_n}^{(i+1)\Delta_n} (s - i\Delta_n) dB_s \quad \text{for } i,\ n \geq 0 \tag{6}$$
$$\xi'_{i+1,n} = \frac{1}{\Delta_n^{\frac{3}{2}}} \int_{(i+1)\Delta_n}^{(i+2)\Delta_n} (i\Delta_n + 2\Delta_n - s) dB_s \quad \text{for } i \geq -1,\ n \geq 0. \tag{7}$$

Lemma 2.1. *The r.v.* $\xi_{i,n}$ *and* $\xi'_{i+1,n}$ *are independent and Gaussian;* $\xi_{i,n}$ *is* $\mathcal{G}^n_{i+1}$ *measurable and independent of* $\mathcal{G}^n_i$*;* $\xi'_{i+1,n}$ *is* $\mathcal{G}^n_{i+2}$ *measurable and independent of* $\mathcal{G}^n_{i+1}$*. The following expectations are useful for the sequel:*

$$E\left(\xi_{i,n}\mid \mathcal{G}^n_i\right)=E\left(\xi'_{i+1,n}\mid \mathcal{G}^n_i\right)=0$$
$$E\left(\xi^2_{i,n}\mid \mathcal{G}^n_i\right)=E\left(\xi'^2_{i+1,n}\mid \mathcal{G}^n_i\right)=\frac{1}{3}$$
$$E\left(\left(\xi^2_{i,n}-\frac{1}{3}\right)^2\mid \mathcal{G}^n_i\right)=E\left(\left(\xi'^2_{i+1,n}-\frac{1}{3}\right)^2\mid \mathcal{G}^n_i\right)=\frac{2}{9}$$
$$E\left(\left(\xi^2_{i,n}-\frac{1}{3}\right)\xi'_{i,n}\mid \mathcal{G}^n_i\right)=E\left(\left(\xi'^2_{i+1,n}-\frac{1}{3}\right)\xi'_{i,n}\mid \mathcal{G}^n_i\right)=0$$
$$E\left(\xi_{i,n}\xi'_{i,n}\mid \mathcal{G}^n_i\right)=\frac{1}{6}\,\cdot$$

Proof. Easy computations based on (6) and (7) give the result. For example

$$E\left(\xi_{i,n}\xi'_{i,n}\mid \mathcal{G}^n_i\right)=\frac{1}{\Delta^3_n}\int_{i\Delta_n}^{(i+1)\Delta_n}(s-i\Delta_n)(i\Delta_n+\Delta_n-s)ds=\frac{1}{6}\,\cdot$$

□

Our first result is a first order comparison between $\frac{J_i}{\Delta_n}$ and $X_{i\Delta_n}$.

Proposition 2.2. *Assume that* $2\alpha_1<K_l$*, (with* α_1 *and* K_l *given in (A2, A3)) then:*

$$\frac{J_i}{\Delta_n}-X_{i\Delta_n}=a(X_{i\Delta_n})\Delta_n^{\frac{1}{2}}\xi'_{i,n}+e_{i,n}\tag{8}$$

where,

$$|E\left(e_{i,n}\mid \mathcal{G}^n_i\right)|\le\Delta_n c(1+\mathcal{B}_r(X_{i\Delta_n}))\tag{9}$$
$$E\left(e^2_{i,n}\mid \mathcal{G}^n_i\right)\le\Delta^2_n c(\mathcal{B}_l^{2\alpha_1}(X_{i\Delta_n})+\mathcal{B}_r^{2(1+\alpha_1)}(X_{i\Delta_n})).\tag{10}$$

Moreover, if k *is a real number* ≥ 1*, then for all* i,n $(i\le n-1)$*:*

$$E\left(\left|\frac{J_i}{\Delta_n}-X_{i\Delta_n}\right|^k\mid \mathcal{G}^n_i\right)\le\Delta_n^{\frac{k}{2}}c(1+\mathcal{B}^k_r(X_{i\Delta_n})).\tag{11}$$

Proof. We have:

$$\frac{J_i}{\Delta_n}-X_{i\Delta_n}=\frac{1}{\Delta_n}\int_{i\Delta_n}^{(i+1)\Delta_n}(X_v-X_{i\Delta_n})dv$$

and $X_v-X_{i\Delta_n}=\int_{i\Delta_n}^v b(X_s)ds+\int_{i\Delta_n}^v a(X_s)dB_s$.

So by the Fubini theorem, we get:

$$\frac{J_i}{\Delta_n}-X_{i\Delta_n}=\frac{1}{\Delta_n}a(X_{i\Delta_n})\int_{i\Delta_n}^{(i+1)\Delta_n}((i+1)\Delta_n-v)dB_v+e_{i,n}=\Delta_n^{\frac{1}{2}}a(X_{i\Delta_n})\xi'_{i,n}+e_{i,n}\tag{12}$$

where $e_{i,n} = \alpha_{i,n} + \beta_{i,n}$, and

$$\alpha_{i,n} = \frac{1}{\Delta_n} \int_{i\Delta_n}^{(i+1)\Delta_n} (a(X_v) - a(X_{i\Delta_n}))(i\Delta_n + \Delta_n - v) dB_v \tag{13}$$

$$\beta_{i,n} = \frac{1}{\Delta_n} \int_{i\Delta_n}^{(i+1)\Delta_n} \int_{i\Delta_n}^{v} b(X_s) ds dv. \tag{14}$$

Using Assumption (A2), we get $|\beta_{i,n}| \leq c\Delta_n(1 + \sup_{s\in[i\Delta_n,(i+1)\Delta_n]} \mathcal{B}_r(X_s))$.

Now by Assumption (A3), for all $k \geq 0$

$$E\left(|\beta_{i,n}|^k \mid \mathcal{G}_i^n\right) \leq c\Delta_n^k(1 + \mathcal{B}_r^k(X_{i\Delta_n})). \tag{15}$$

Also $E\left(\alpha_{i,n} \mid \mathcal{G}_i^n\right) = 0$, so we get $|E\left(e_{i,n} \mid \mathcal{G}_i^n\right)| \leq c\Delta_n(1 + \mathcal{B}_r(X_{i\Delta_n}))$.

Now, for $k \geq 2$, applying the Burkholder-Davis-Gundy and the Jensen inequalities yields:

$$\begin{aligned} E\left(\left|\alpha_{i,n}^k\right| \mid \mathcal{G}_i^n\right) &\leq \frac{c}{\Delta_n^k} E\left(\left(\int_{i\Delta_n}^{(i+1)\Delta_n} ((i+1)\Delta_n - v)^2 (a(X_v) - a(X_{i\Delta_n}))^2 dv\right)^{\frac{k}{2}} \mid \mathcal{G}_i^n\right) \\ &\leq c \int_{i\Delta_n}^{(i+1)\Delta_n} \phi_k(v) dv \end{aligned}$$

with $\phi_k(v) = E\left(|a(X_v) - a(X_{i\Delta_n})|^k \mid \mathcal{G}_i^n\right)$. By Proposition A of the Appendix, there exists $c > 0$ such that, for all $v \in [i\Delta_n, (i+1)\Delta_n]$,

$$\phi_k(v) \leq c\Delta_n^{\frac{k}{2}}(\mathcal{B}_l^{\alpha_1 k}(X_{i\Delta_n}) + \mathcal{B}_r^{(\alpha_1+1)k}(X_{i\Delta_n})).$$

Finally,

$$E\left(\left|\alpha_{i,n}^k\right| \mid \mathcal{G}_i^n\right) \leq c\Delta_n^{\frac{k}{2}+1}(\mathcal{B}_l^{\alpha_1 k}(X_{i\Delta_n}) + \mathcal{B}_r^{(\alpha_1+1)k}(X_{i\Delta_n})). \tag{16}$$

So, by (15, 16), with $k = 2$,

$$E\left(e_{i,n}^2 \mid \mathcal{G}_i^n\right) \leq c\Delta_n^2(\mathcal{B}_l^{2\alpha_1}(X_{i\Delta_n}) + \mathcal{B}_r^{2(\alpha_1+1)}(X_{i\Delta_n})).$$

Using Proposition A (in the Appendix), we have

$$E\left(\sup_{s\in[i\Delta_n,(i+1)\Delta_n]} |X_s - X_{i\Delta_n}|^k \mid \mathcal{G}_i^n\right) \leq \Delta_n^{\frac{k}{2}}(1 + \mathcal{B}_r^k(X_{i\Delta_n})).$$

This implies (11). $\square$

For statistical applications, Proposition 2.2 is not sufficient. An approximation of the joint law of $(\frac{J_i}{\Delta_n})$ is required. This is done by the following result:

Theorem 2.3. *We have*

$$\frac{J_{i+1}}{\Delta_n} - \frac{J_i}{\Delta_n} - b\left(\frac{J_i}{\Delta_n}\right)\Delta_n = a(X_{i\Delta_n})(\xi_{i,n} + \xi'_{i+1,n})\Delta_n^{\frac{1}{2}} + \varepsilon_{i,n} \tag{17}$$

where $\varepsilon_{i,n}$ is $\mathcal{G}_{i+2}^n$ measurable, and there exists a constant c such that for all i,n:

$$\text{If } \beta_1 \vee 2\beta_2 \vee 4\alpha_1 < K_l,$$

$$|E(\varepsilon_{i,n} \mid \mathcal{G}_i^n)| \le \Delta_n^2 c(\mathcal{B}_l^{\beta_1\vee\beta_2}(X_{i\Delta_n}) + \mathcal{B}_r^{(1+\beta_1)\vee(2+\beta_2)}(X_{i\Delta_n})) \tag{18}$$

$$\text{If } 4\alpha_1 \vee 2\alpha_2 < K_l,$$

$$E\left(\varepsilon_{i,n}^2 \mid \mathcal{G}_i^n\right) \le \Delta_n^2 c(\mathcal{B}_l^{2\alpha_1\vee\alpha_2}(X_{i\Delta_n}) + \mathcal{B}_r^{3+2\alpha_1+\alpha_2}(X_{i\Delta_n})) \tag{19}$$

$$E\left(\varepsilon_{i,n}^4 \mid \mathcal{G}_i^n\right) \le \Delta_n^4 c(\mathcal{B}_l^{4\alpha_1\vee 2\alpha_2}(X_{i\Delta_n}) + \mathcal{B}_r^{6+4\alpha_1+2\alpha_2}(X_{i\Delta_n})) \tag{20}$$

$$\text{If } 4\alpha_1 \vee \alpha_2 \vee 4\beta_1 < K_l,$$

$$|E(\varepsilon_{i,n}\xi_{i,n} \mid \mathcal{G}_i^n)| \le \Delta_n^{\frac{3}{2}} c(\mathcal{B}_l^{(\alpha_1+\beta_1)\vee\alpha_2}(X_{i\Delta_n}) + \mathcal{B}_r^{(1+\alpha_1+\beta_1)\vee(2+\alpha_2)}(X_{i\Delta_n})) \tag{21}$$

$$\left|E\left(\varepsilon_{i,n}\xi'_{i+1,n} \mid \mathcal{G}_i^n\right)\right| \le \Delta_n^{\frac{3}{2}}(\mathcal{B}_l^{(\alpha_1+\beta_1)\vee\alpha_2)}(X_{i\Delta_n}) + \mathcal{B}_r^{(1+\alpha_1+\beta_1)\vee(2+\alpha_2)}(X_{i\Delta_n})). \tag{22}$$

Futhermore, if k is a real number ≥ 1, then for all i,n $(i \le n-1)$,

$$E\left(\left|\frac{J_{i+1}}{\Delta_n} - \frac{J_i}{\Delta_n}\right|^k \mid \mathcal{G}_i^n\right) \le \Delta_n^{\frac{k}{2}}(1 + \mathcal{B}_r^k(X_{i\Delta_n})). \tag{23}$$

Proof. We integrate, between $i\Delta_n$ and $(i+1)\Delta_n$, the following equality:

$$X_{s+\Delta_n} - X_s = \int_s^{s+\Delta_n} (a(X_v)dB_v + b(X_v)dv).$$

Hence, $J_{i+1} - J_i = A_i + B_i$, with

$$A_i = \int_{i\Delta_n}^{(i+1)\Delta_n} ds \int_s^{s+\Delta_n} a(X_v)dB_v, \quad B_i = \int_{i\Delta_n}^{(i+1)\Delta_n} ds \int_s^{s+\Delta_n} b(X_v)dv.$$

Interchanging the order of integrations, we obtain

$$A_i = \int_{i\Delta_n}^{(i+1)\Delta_n} a(X_v)(v - i\Delta_n)dB_v + \int_{(i+1)\Delta_n}^{(i+2)\Delta_n} a(X_v)((i+2)\Delta_n - v)dB_v.$$

Analogously,

$$B_i = \int_{i\Delta_n}^{(i+1)\Delta_n} b(X_v)(v - i\Delta_n)dv + \int_{(i+1)\Delta_n}^{(i+2)\Delta_n} b(X_v)((i+2)\Delta_n - v)dv.$$

Introducing $a(X_{i\Delta_n})$ in A_i yields

$$\begin{aligned} A_i &= a(X_{i\Delta_n})\left(\int_{i\Delta_n}^{(i+1)\Delta_n} (v - i\Delta_n)dB_v + \int_{(i+1)\Delta_n}^{(i+2)\Delta_n} ((i+2)\Delta_n - v)dB_v\right) + a_{i,n} + a'_{i+1,n} \\ &= a(X_{i\Delta_n})\Delta_n^{\frac{3}{2}}(\xi_{i,n} + \xi'_{i+1,n}) + a_{i,n} + a'_{i+1,n} \end{aligned}$$

with

$$a_{i,n} = \int_{i\Delta_n}^{(i+1)\Delta_n} (v - i\Delta_n)(a(X_v) - a(X_{i\Delta_n}))dB_v \tag{24}$$

$$a'_{i,n} = \int_{(i+1)\Delta_n}^{(i+2)\Delta_n} ((i+2)\Delta_n - v)(a(X_v) - a(X_{i\Delta_n}))dB_v. \tag{25}$$

Analogously, introducing $b(\frac{J_i}{\Delta_n})$ in B_i yields

$$\begin{aligned} B_i &= b\left(\frac{J_i}{\Delta_n}\right)\left(\int_{i\Delta_n}^{(i+1)\Delta_n} (v - i\Delta_n)dv + \int_{(i+1)\Delta_n}^{(i+2)\Delta_n} ((i+2)\Delta_n - v)dv\right) + b_{i,n} + b'_{i+1,n} \\ &= b\left(\frac{J_i}{\Delta_n}\right)\Delta_n^2 + b_{i,n} + b'_{i+1,n} \end{aligned}$$

with

$$b_{i,n} = \int_{i\Delta_n}^{(i+1)\Delta_n} (v - i\Delta_n)\left(b(X_v) - b\left(\frac{J_i}{\Delta_n}\right)\right)dv \tag{26}$$

$$b'_{i,n} = \int_{(i+1)\Delta_n}^{(i+2)\Delta_n} ((i+2)\Delta_n - v)\left(b(X_v) - b\left(\frac{J_i}{\Delta_n}\right)\right)dv. \tag{27}$$

Therefore, we get the expansion

$$\frac{J_{i+1}}{\Delta_n} - \frac{J_i}{\Delta_n} - b\left(\frac{J_i}{\Delta_n}\right)\Delta_n = a(X_{i\Delta_n})(\xi_{i,n} + \xi'_{i+1,n})\Delta_n^{\frac{1}{2}} + \varepsilon_{i,n}$$

with

$$\varepsilon_{i,n} = \frac{a_{i,n}}{\Delta_n} + \frac{a'_{i+1,n}}{\Delta_n} + \frac{b_{i,n}}{\Delta_n} + \frac{b'_{i+1,n}}{\Delta_n}\,. \tag{28}$$

• Let us prove (18). By (26), we have

$$E\left(\frac{b_{i,n}}{\Delta_n} \mid \mathcal{G}_i^n\right) = \frac{1}{\Delta_n}\int_{i\Delta_n}^{i\Delta_n+\Delta_n} (v - i\Delta_n)E\left(b(X_v) - b(J_i) \mid \mathcal{G}_i^n\right)dv.$$

Since we know by Ito's formula, Assumptions (A3) and (A2), that

$$\sup_{v\in[i\Delta_n,\ (i+1)\Delta_n]} \left|E\left(b(X_v) - b(X_{i\Delta_n}) \mid \mathcal{G}_i^n\right)\right| \le \Delta_n c(\mathcal{B}_l^{\beta_1\vee\beta_2}(X_{i\Delta_n}) + \mathcal{B}_r^{(1+\beta_1)\vee(2+\beta_2)}(X_{i\Delta_n}))$$

an application of Taylor's formula and Proposition 2.2, yields

$$\left|E\left(b\left(\frac{J_i}{\Delta_n}\right) - b(X_{i\Delta_n}) \mid \mathcal{G}_i^n\right)\right| \le \Delta_n c(\mathcal{B}_l^{\beta_1\vee\beta_2}(X_{i\Delta_n}) + \mathcal{B}_r^{(1+\beta_1)\vee(2+\beta_2)}(X_{i\Delta_n})).$$

Hence

$$\sup_{v\in[i\Delta_n,\ (i+1)\Delta_n]} \left|E\left(b(X_v) - b\left(\frac{J_i}{\Delta_n}\right) \mid \mathcal{G}_i^n\right)\right| \le \Delta_n c(\mathcal{B}_l^{\beta_1\vee\beta_2}(X_{i\Delta_n}) + \mathcal{B}_r^{(1+\beta_1)\vee(2+\beta_2)}(X_{i\Delta_n})).$$

Hence

$$\left| E\left(\frac{b_{i,n}}{\Delta_n} \mid \mathcal{G}_i^n\right)\right| \leq \Delta_n^2 c(\mathcal{B}_l^{\beta_1\vee\beta_2}(X_{i\Delta_n}) + \mathcal{B}_r^{(1+\beta_1)\vee(2+\beta_2)}(X_{i\Delta_n})).$$

In a similar way, we bound $\left|E\left(\frac{b'_{i+1,n}}{\Delta_n} \mid \mathcal{G}_i^n\right)\right|$ (see (27)).

Now (see (24, 25)), $E\left(a_{i,n} \mid \mathcal{G}_i^n\right) = E\left(a'_{i+1,n} \mid \mathcal{G}_i^n\right) = 0$, so (18) is proved.

- We now prove (19) and (20).

Using the Cauchy-Schwarz inequality it is enough to show (20).

By (A2), we write (see (26, 27)):

$$\left|\frac{b_{i,n}}{\Delta_n}\right| \leq \Delta_n 2c \sup_{s\in[i\Delta_n, i\Delta_n+\Delta_n]}(1+\mathcal{B}_r(X_s)), \quad \left|\frac{b'_{i+1,n}}{\Delta_n}\right| \leq \Delta_n 2c \sup_{s\in[i\Delta_n+\Delta_n, i\Delta_n+2\Delta_n]}(1+\mathcal{B}_r(X_s)).$$

Using Assumption (A3), we get

$$E\left(\left|\frac{b_{i,n}}{\Delta_n}\right|^4 + \left|\frac{b'_{i+1,n}}{\Delta_n}\right|^4 \mid \mathcal{G}_i^n\right) \leq \Delta_n^4 c(1+\mathcal{B}_r^4(X_{i\Delta_n})).$$

To end the proof, we have to bound $E\left(\left|\frac{a_{i,n}}{\Delta_n}\right|^4 \mid \mathcal{G}_i^n\right)$. Using the Burkholder-Davis-Gundy inequality we obtain:

$$E\left(\left|\frac{a_{i,n}}{\Delta_n}\right|^4 \mid \mathcal{G}_i^n\right) = \frac{c}{\Delta_n^4} E\left(\left(\int_{i\Delta_n}^{i\Delta_n+\Delta_n} (a(X_v)-a(X_{i\Delta_n}))^2 (v-i\Delta)^2 dv\right)^2 \mid \mathcal{G}_i^n\right). \tag{29}$$

But using the Ito formula, and Assumption (A2) we can write: $(a(X_v)-a(X_{i\Delta_n}))^2 = M_v + A_v$, where

$$M_v = 2\int_{i\Delta_n}^{v} \psi_{i,n}(s)dB_s, \quad \text{with } \psi_{i,n}(s) = (a(X_s)-a(X_{i\Delta_n}))a'(X_s)a(X_s) \tag{30}$$

and:

$$\sup_{v\in[i\Delta_n,\ (i+1)\Delta_n]} |A_v| \leq c\Delta_n \sup_{s\in[i\Delta_n,\ (i+1)\Delta_n]} (\mathcal{B}_r^{2\alpha_1\vee\alpha_2}(X_s) + \mathcal{B}_l^{(2+2\alpha_1)\vee(3+\alpha_2)}(X_s)). \tag{31}$$

So, replacing in (29), after some easy computations, and applying (A3) to the right hand side of (31), we get

$$E\left(\left|\frac{a_{i,n}}{\Delta_n}\right|^4 \mid \mathcal{G}_i^n\right) \leq \frac{2}{\Delta_n^4} E\left(\gamma_{i,n}^2 \mid \mathcal{G}_i^n\right) + \Delta_n^4 (B_l^{4\alpha_1\vee 2\alpha_2}(X_{i\Delta_n}) + B_r^{6+4\alpha_1+2\alpha_2} X_{i\Delta_n}))$$

with $\gamma_{i,n} = \int_{i\Delta_n}^{(i+1)\Delta_n} M_v(v-i\Delta_n)^2 dv$ (see (30)).

It remains to bound $E\left(\gamma_{i,n}^2 \mid \mathcal{G}_i^n\right)$. Using the Fubini theorem, we have

$$\gamma_{i,n} = \int_{i\Delta_n}^{(i+1)\Delta_n} \psi_{i,n}(s)\left(\int_s^{i\Delta_n+\Delta_n}(v-i\Delta_n)^2 dv\right) dB_s.$$

Hence,

$$E\left(\gamma_{i,n}^2 \mid \mathcal{G}_i^n\right) \leq \int_{i\Delta_n}^{(i+1)\Delta_n} E\left(\psi_{i,n}^2(s) \mid \mathcal{G}_i^n\right)\Delta_n^6 ds. \tag{32}$$

But by the Cauchy-Schwarz inequality,

$$E\left(\psi_{i,n}^2(s) \mid \mathcal{G}_i^n\right) \leq E\left((a(X_s)-a(X_{i\Delta_n}))^4 \mid \mathcal{G}_i^n\right)^{\frac{1}{2}} E\left(a'^4(X_s)a^4(X_s) \mid \mathcal{G}_i^n\right)^{\frac{1}{2}}.$$

Then using (A2, A3) and Proposition A of the Appendix, we obtain for $s \in [i\Delta_n, (i+1)\Delta_n]$:

$$E\left(\psi_{i,n}^2(s) \mid \mathcal{G}_i^n\right) \leq \Delta_n c(\mathcal{B}_l^{4\alpha_1}(X_{i\Delta_n}) + \mathcal{B}_r^{4+2\alpha_1}(X_{i\Delta_n})).$$

Replacing the last inequality in (32), we get $E\left(\gamma_{i,n}^2 \mid \mathcal{G}_i^n\right) \leq \Delta_n^8 c(\mathcal{B}_l^{4\alpha_1}(X_{i\Delta_n}) + \mathcal{B}_r^{4+2\alpha_1})$. We obtain a similar bound for $E\left(\left|\frac{a'_{i+1,n}}{\Delta_n}\right|^4 \mid \mathcal{G}_i^n\right)$, and hence (20) is proved.

- To show (21, 22), it is enough to obtain the following inequalities (recall (28)):

$$\begin{aligned}
|E\left(a_{i,n}\xi_{i,n} \mid \mathcal{G}_i^n\right)| &\leq \Delta_n^{\frac{5}{2}} c(\mathcal{B}_l^{\alpha_1\vee\alpha_2}(X_{i\Delta_n}) + \mathcal{B}_r^{(1+\alpha_1)\vee(2+\alpha_1)}(X_{i\Delta_n}))\\
\left|E\left(a'_{i+1,n}\xi_{i,n} \mid \mathcal{G}_i^n\right)\right| &= 0\\
|E\left(b_{i,n}\xi_{i,n} \mid \mathcal{G}_i^n\right)| &\leq \Delta_n^{\frac{5}{2}} c(\mathcal{B}_l^{\alpha_1+\beta_1}(X_{i\Delta_n}) + \mathcal{B}_r^{1+\alpha_1+\beta_1}(X_{i\Delta_n}))\\
\left|E\left(b'_{i+1,n}\xi_{i,n} \mid \mathcal{G}_i^n\right)\right| &\leq \Delta_n^{\frac{5}{2}} c(\mathcal{B}_l^{\alpha_1+\beta_1}(X_{i\Delta_n}) + \mathcal{B}_r^{1+\alpha_1+\beta_1}(X_{i\Delta_n})).
\end{aligned}$$

These inequalities follow from the expressions of $a_{i,n}$, $a'_{i+1,n}$, $b_{i,n}$, $b'_{i+1,n}$.

- By a straightforward modification of Proposition A, we show that for all $k \geq 0$, there exists c such that for all i, n $(i \leq n-1)$,

$$E\left(\sup_{s\in[i\Delta_n,(i+2)\Delta_n]} |X_s - X_{i\Delta}|^k \mid \mathcal{G}_i^n\right) \leq c(1+\mathcal{B}_r^k(X_{i\Delta_n})).$$

This implies (23). □

Remark 2.4. 1. *In the case where X is a stationary Ornstein-Uhlenbeck process $dX_t = \mu X_t dt + \sigma dB_t$ and for a fixed sampling interval Δ, we have an exact formula, analogous to our expansion* (17) *(see Gloter [5]).*

$$J_{i+1} - e^{\mu\Delta}J_i = \frac{\sigma}{\mu}\int_{i\Delta}^{(i+1)\Delta}(e^{\mu\Delta} - e^{\mu((i+1)\Delta-s)})dB_s + \frac{\sigma}{\mu}\int_{(i+1)\Delta}^{(i+2)\Delta}(e^{\mu((i+2)\Delta-s)}-1)dB_s.$$

Furthermore, in this case, the covariance structure of (J_i) is the one of an ARMA(1,1) process.

2. $(U_i)_{i=0,\dots,n-1} = (\xi_{i,n} + \xi'_{i+1,n})_{i=0,\dots,n-1}$ *is a Gaussian vector with covariance function:* *Var* $U_i = \frac{2}{3}$, *Cov* $(U_i, U_{i+1}) = \frac{1}{6}$ *and Cov* $(U_i, U_{i+k}) = 0$ *if* $k \geq 2$.

This is the covariance function of an MA(1) vector. Therefore, through expansion (17) *we do not recover a Markovian property for* $\frac{J_i}{\Delta_n}$. *In the special case where a is constant, the expansion means that the process* $(\frac{J_i}{\Delta_n})$ *may be approximated by an ARMA(1,1) process.*

3. Statistical applications: Estimation of the diffusion coefficient

For the estimation of the diffusion coefficient, we suppose that the integrated process I is observed on a time interval with finite length (which we arbitrarily suppose to be equal to 1). So, $\Delta_n = n^{-1}$ and $J_i = J_i^n = \int_{\frac{i}{n}}^{\frac{i+1}{n}} X_s ds$.

We recall that, when $(X_{\frac{i}{n}})_{i=0,\ldots,n-1}$ is observed, estimation of the diffusion coefficient is based on the quadratic variation of X (see Genon-Catalot and Jacod [3]). Hence, in Section 3.1, we study the quadratic variation of the process (J_i). In Section 3.2, where the coefficient a depends on an unknown parameter σ, we deduce an estimator of this parameter.

3.1. Approximation of the quadratic variation

We show that we can not replace $X_{\frac{i}{n}}$ by nJ_i in the classical approximation of the quadratic variation of X:

$$\sum_{i=0}^{n-2}(X_{\frac{i+1}{n}} - X_{\frac{i}{n}})^2 \xrightarrow{n\to\infty} \int_0^1 a^2(X_s)ds. \tag{33}$$

Theorem 3.1. *Let $(X_t)_{t\in[0,1]}$ be a diffusion satisfying (A1–A3) with $K_l = \infty$. Furthermore assume that:*

(A4) $\forall\ k \geq 0$, $E\left(\mathcal{B}_l^k(X_0)\right) < \infty$, $E\left(\mathcal{B}_r^k(X_0)\right) < \infty$.

Let $f \in \mathcal{C}^1(l, r)$ be a function with real values and suppose that there exists $c \geq 0$ such that

$$|f(x)| + |f'(x)| \leq c(\mathcal{B}_l^c(x) + \mathcal{B}_r^c(x)). \tag{34}$$

Then,

$$\sum_{i=0}^{n-2} (nJ_{i+1} - nJ_i)^2 f(nJ_i) \xrightarrow[\mathbf{P}]{n\to\infty} \frac{2}{3}\int_0^1 a^2(X_s)f(X_s)ds.$$

Proof. First, we remark that by Assumptions (A3) and (A4), we have

$$\forall k \geq 0, \quad \sup_{t\in[0,1]} E\left(\mathcal{B}_l^k(X_t)\right) < \infty \quad \text{and} \quad \sup_{t\in[0,1]} E\left(\mathcal{B}_r^k(X_t)\right) < \infty. \tag{35}$$

By the continuity of a, f and X, $\frac{2}{3}\sum_{i=0}^{n-2} a^2(X_{\frac{i}{n}})f(X_{\frac{i}{n}})n^{-1} \xrightarrow[\mathbf{P}]{n\to\infty} \frac{2}{3}\int_0^1 a^2(X_s)f(X_s)ds$. Hence it is enough to prove:

$$\sum_{i=0}^{n-2}\left\{(nJ_{i+1} - nJ_i)^2 f(nJ_i) - \frac{2}{3n}a^2(X_{\frac{i}{n}})f(X_{\frac{i}{n}})\right\} \xrightarrow[\mathbf{P}]{n\to\infty} 0.$$

For this, we use the expansion (17):

$$\sum_{i=0}^{n-2}\left\{(nJ_{i+1} - nJ_i)^2 f(nJ_i) - \frac{2}{3n}a^2(X_{\frac{i}{n}})f(X_{\frac{i}{n}})\right\} = D_n^{(1)} + D_n^{(2)} + D_n^{(3)} + D_n^{(4)},$$

with:

$$D_n^{(1)} = \sum_{i=0}^{n-2} (nJ_{i+1} - nJ_i)^2 (f(nJ_i) - f(X_{\frac{i}{n}}))$$

$$D_n^{(2)} = \sum_{i=0}^{n-2} n^{-1} \left\{ (\xi_{i,n} + \xi'_{i+1,n})^2 - \frac{2}{3} \right\} a^2(X_{\frac{i}{n}}) f(X_{\frac{i}{n}})$$

$$D_n^{(3)} = \sum_{i=0}^{n-2} (b(nJ_i)n^{-1} + \varepsilon_{i,n})^2 f(X_{\frac{i}{n}})$$

$$D_n^{(4)} = 2\sum_{i=0}^{n-2} n^{-\frac{1}{2}} (\xi_{i,n} + \xi'_{i+1,n})(b(nJ_i)n^{-1} + \varepsilon_{i,n}) a(X_{\frac{i}{n}}) f(X_{\frac{i}{n}}).$$

First, we write $f(nJ_i) - f(X_{\frac{i}{n}}) = f'(T_{i,n})(nJ_i - X_{\frac{i}{n}})$, where $T_{i,n} \in [nJ_i, X_{\frac{i}{n}}]$. Using Cauchy-Schwarz's inequality (34, A3) and (11) we obtain

$$E\left(\left| f'(T_{i,n})(nJ_i - X_{\frac{i}{n}}) \right|^2 \mid \mathcal{G}_i^n \right) \le cn^{-1} (\mathcal{B}_l^c(X_{\frac{i}{n}}) + \mathcal{B}_r^c(X_{\frac{i}{n}})) \text{ with } c \text{ a positive constant.}$$

Again Cauchy-Schwarz's inequality, with (23) yields

$$E\left((nJ_{i+1} - nJ_i)^2 \left| f(nJ_i) - f\left(X_{\frac{i}{n}}\right) \right| \mid \mathcal{G}_i^n \right) \le cn^{-\frac{3}{2}} \left(\mathcal{B}_l^c\left(X_{\frac{i}{n}}\right) + \mathcal{B}_r^c\left(\frac{i}{n}\right) \right).$$

Hence, by (35),

$$E\left(\left| D_n^{(1)} \right| \right) \le cn^{-\frac{1}{2}} \sup_{t\in[0,1]} E\left(\mathcal{B}_l^c(X_t) + \mathcal{B}_r^c(X_t) \right).$$

Now, we bound, using (A2, 34, A3) and (19):

$$E\left((b(nJ_i)n^{-1} + \varepsilon_{i,n})^2 \left| f\left(X_{\frac{i}{n}}\right) \right| \mid \mathcal{G}_i^n \right) \le cn^{-2} \left(\mathcal{B}_l^c\left(X_{\frac{i}{n}}\right) + \mathcal{B}_r^c\left(X_{\frac{i}{n}}\right) \right) \text{ with } c \text{ a positive constant.}$$

Hence, by (35),

$$E\left(\left| D_n^{(3)} \right| \right) \le cn^{-1} \sup_{t\in[0,1]} E\left(\mathcal{B}_l^c(X_t) + \mathcal{B}_r^c(X_t) \right).$$

We deduce $D_n^{(3)} \xrightarrow[\mathbf{P}]{n\to\infty} 0$.

Analogously: $E\left(\left| D_n^{(4)} \right| \right) \le cn^{-\frac{1}{2}} \sup_{t\in[0,1]} E\left(\mathcal{B}_l^c(X_t) + \mathcal{B}_r^c(X_t) \right)$, with $c > 0$. Hence $D_n^{(4)} \xrightarrow[\mathbf{P}]{n\to\infty} 0$.

Using that, by Lemma 2.1, $E\left((\xi_{i,n} + \xi'_{i+1,n})^2 - \frac{2}{3} \mid \mathcal{G}_i^n \right) = 0$ and then Cauchy-Schwarz's inequality we compute:

$$\begin{aligned} E\left({D_n^{(2)}}^2 \right) &= \sum_{\substack{0\le i,j\le n-2 \\ |i-j|\le 1}} n^{-2} E\left(a^2(X_{i\Delta_n}) \left\{ (\xi_{i,n} + \xi'_{i+1,n})^2 - \frac{2}{3} \right\}^2 a^2(X_{j\Delta_n}) \left\{ (\xi_{j,n} + \xi'_{j+1,n})^2 - \frac{2}{3} \right\}^2 \right) \\ &\le cn^{-1} \sup_{t\in[0,1]} E(1 + \mathcal{B}_r^8(X_t))^{\frac{1}{2}}. \end{aligned}$$

Therefore, $D_n^{(2)}$ converges to zero in $\mathbf{L}^2$ sense. $\square$

Remark 3.2. *In particular, for $f = 1$, we see that if we replace $X_{\frac{i}{n}}$ by nJ_i in (33), then we underestimate the quadratic variation of X by a factor $\frac{2}{3}$.*

3.2. Parametric estimation of the diffusion coefficient

We suppose that $a(x, \sigma)$ is a real function defined on $(l, r) \times \Theta$ where Θ is a compact interval of $\mathbb{R}$. The diffusion X satisfies (1) with $a(x) = a(x, \sigma_0)$, where $\sigma_0 \in \overset{\circ}{\Theta}$. Our goal is to estimate the true value σ_0 of the parameter with the observation $(J_0, \ldots, J_{n-1})$. The function b may be known or unknown and since our observation of I is restricted to the finite time interval $[0, 1]$, we cannot estimate b (see Genon-Catalot and Jacod [3]).

We suppose that $a(x, \theta)$ is the restriction of a function defined on an open subset $\mathcal{O}$ of $\mathbb{R}^2$, such that $(l, r) \times \Theta \subset \mathcal{O}$ and $\frac{\partial^{i+j}}{\partial^i x \partial^j \sigma} a(x, \sigma)$ exists and are continuous on $\mathcal{O}$.

We need some assumptions for statistical purposes:

(S1) There exists c such that for all $x \in (l, r)$,

$$\sup_{\sigma \in \Theta} \sup_{(i,j) \in \{1,2,3\}} \left| \frac{\partial^{i+j}}{\partial^i x \partial^j \sigma} a(x, \sigma) \right| \leq c(\mathcal{B}_l^c(x) + \mathcal{B}_r^c(x))$$
$$\sup_{\sigma \in \Theta} \left| a^{-1}(x, \sigma) \right| \leq c(\mathcal{B}_l^c(x) + \mathcal{B}_r^c(x)).$$

Our identification assumption is the following

(S2) For almost all realisation of the path X, if $\sigma \neq \sigma_0$ then the two functions on $[0, 1]$ $t \to a^2(X_t, \sigma)$ and $t \to a^2(X_t, \sigma_0)$ are not a.e. equal.

Taking into account Theorem 3.1, we modify the contrast based on $X_{\frac{i}{n}}$, used for instance in Genon-Catalot and Jacod [3]. This yields the definition of the following contrast,

$$U_n(\sigma) = \frac{3}{2} \sum_{i=0}^{n-2} \frac{(nJ_{i+1} - nJ_i)^2}{a^2(nJ_i, \sigma)} + n^{-1} \sum_{i=0}^{n-2} \log(a^2(nJ_i, \sigma)). \tag{36}$$

Let $\overline{\sigma}_n = \operatorname{arginf}_{\sigma \in \Theta} U_n(\sigma)$ be the associated minimum contrast estimator.

Theorem 3.3. *Assume (A1–A4) with $K_l = \infty$ and (S1, S2), then $\overline{\sigma}_n \xrightarrow[\mathbf{P}]{n \to \infty} \sigma_0$.*

Proof. We follow the standard proof of consistency of minimum contrast estimators: we easily check, using Assumption (S2), that the function $\sigma \mapsto K(\sigma, \sigma_0) := \int_0^1 \left\{ \frac{a^2(X_s, \sigma_0)}{a^2(X_s, \sigma)} + \log a^2(X_s, \sigma) \right\} ds$ admits a unique minimum for $\sigma = \sigma_0$. Hence to show the consistency it suffices to prove

$$U_n(\sigma) \xrightarrow{n \to \infty} K(\sigma, \sigma_0) \text{ uniformly in } \sigma \text{ in probability.}$$

The pointwise convergence in σ is clear by Theorem 3.1, and since we can show, using (11), that $n^{-1} \sum_{i=0}^{n-2} \left\{ \log(a^2(nJ_i, \sigma)) - \log(a^2(X_{\frac{i}{n}}, \sigma)) \right\} \xrightarrow[\mathbf{P}]{n \to \infty} 0$.

To get the uniformity, by (S1, 23) and (35), we deduce

$$E\left(\sup_{\sigma \in \Theta} \left| (nJ_{i+1} - nJ_i)^2 \frac{\partial}{\partial \sigma} a^{-2}(nJ_i, \sigma) \right| \right) \leq cn^{-1}.$$

Hence $sup_{n \geq 0} E\left(\sup_{\sigma \in \Theta} \left| \frac{\partial}{\partial \sigma} U_n(\sigma) \right| \right) < \infty$. Since Θ is compact, this bound is sufficient to imply the uniformity in the convergence. $\square$

Theorem 3.4. *Assume (A1–A4) with $K_l = \infty$ and (S1, S2),then*

$$n^{\frac{1}{2}}(\overline{\sigma}_n - \sigma_0) \xrightarrow[\mathcal{D}]{n\to\infty} S, \tag{37}$$

where S is a variable defined on an extension of the space $(\Omega, \mathcal{G}_1, P)$ and such that conditionally on $\mathcal{G}_1$, S is centered Gaussian with variance $\frac{9}{16}\left(\int_0^1 \left(\frac{\frac{\partial}{\partial\sigma}a(X_s,\sigma_0)}{a(X_s,\sigma_0)}\right)^2 ds\right)^{-1}$.

Proof. By the classical scheme of proof for studying the asymptotic law of contrast based estimators, we see that it is enough to show the two following properties:

1) The convergence in probability $\int_0^1 \frac{\partial^2}{\partial\sigma^2}U_n(\sigma_0 + (\overline{\sigma}_n - \sigma_0)s)ds \xrightarrow[\mathbf{P}]{n\to\infty} 4\int_0^1 \left(\frac{\dot{a}(X_s,\sigma_0)}{a(X_s,\sigma_0)}\right)^2 ds$ (where $\dot{f} = \frac{\partial}{\partial\sigma}f$ for a function f).
2) The stable convergence in law $n^{\frac{1}{2}}\frac{\partial}{\partial\sigma}U_n(\sigma_0) \xrightarrow[\mathcal{D}]{n\to\infty} Z$, where Z is a variable defined on an extension of the space $(\Omega, \mathcal{G}_1, P)$ and such that conditionally on $\mathcal{G}_1$, Z is centered Gaussian with variance $9\int_0^1 \left(\frac{\dot{a}(X_s,\sigma_0)}{a(X_s,\sigma_0)}\right)^2 ds$. By stable convergence, we mean that for any variable Y, $\mathcal{G}_1$-measurable, we have convergence in law $(Y, n^{\frac{1}{2}}\frac{\partial}{\partial\sigma}U_n(\sigma_0)) \xrightarrow[\mathcal{D}]{n\to\infty} (Y, Z)$. Let us point out that we need this stable convergence since the limit in 1) is not deterministic and that this difficulty appears as soon as mixed normality is involved (see Genon-Catalot and Jacod [3] for instance).

We start by proving 2). We compute

$$\frac{\partial}{\partial\sigma}U_n(\sigma_0) = -3\sum_{i=0}^{n-2}(nJ_{i+1} - nJ_i)^2\frac{\dot{a}(nJ_i,\sigma_0)}{a^3(nJ_i,\sigma_0)} + 2n^{-1}\sum_{i=0}^{n-2}\frac{\dot{a}(nJ_i,\sigma_0)}{a(nJ_i,\sigma_0)},$$

and set

$$V_n = n^{-1}\sum_{i=0}^{n-2}\{-3(\xi_{i,n} + \xi'_{i+1,n})^2 + 2\}\frac{\dot{a}(X_{\frac{i}{n}},\sigma_0)}{a(X_{\frac{i}{n}},\sigma_0)}\,.$$

We write

$$n^{\frac{1}{2}}\left(\frac{\partial}{\partial\sigma}U_n(\sigma_0) - V_n\right) = \sum_{i=0}^{n-2}\phi_{i,n},$$

where $\phi_{i,n} = \left\{3n^{-1}(\xi_{i,n} + \xi'_{i+1,n})^2\frac{\dot{a}(X_{\frac{i}{n}},\sigma_0)}{a(X_{\frac{i}{n}},\sigma_0)} - 3(nJ_{i+1} - nJ_i)^2\frac{\dot{a}(nJ_i,\sigma_0)}{a^3(nJ_i,\sigma_0)}\right\} + 2n^{-1}\left\{\frac{\dot{a}(nJ_i,\sigma_0)}{a(nJ_i,\sigma_0)} - \frac{\dot{a}(X_{\frac{i}{n}},\sigma_0)}{a(X_{\frac{i}{n}},\sigma_0)}\right\}$ is $\mathcal{G}^n_{i+2}$-measurable.

Now, we show after some easy but tedious computations based on Proposition 2.2 and Theorem 2.3 that:

$$|E(\phi_{i,n} \mid \mathcal{G}^n_i)| \le cn^{-\frac{3}{2}}(\mathcal{B}^c_l(X_{\frac{i}{n}}) + \mathcal{B}^c_r(X_{\frac{i}{n}})), \quad E\left(\phi^2_{i,n} \mid \mathcal{G}^n_i\right) \le cn^{-2}(\mathcal{B}^c_l(X_{\frac{i}{n}}) + \mathcal{B}^c_r(X_{\frac{i}{n}})).$$

These two inequalities implies, with (35), $E\left(\left(\sum_{i=0}^{n-2}\phi_{i,n}\right)^2\right) \le cn^{-\frac{1}{2}}$. Hence it is enough to study the asymptotic law of $n^{\frac{1}{2}}V_n$.

For this, we reorder terms in V_n to deal with a triangular array of martingale increments and obtain

$$n^{\frac{1}{2}}V_n = \sum_{i=1}^{n-2}\chi_{i,n} + o_{\mathbf{P}}(1)$$

where,

$$\chi_{i,n} = n^{-\frac{1}{2}} \left\{ -3\left(\xi_{i,n}^2 - \frac{1}{3}\right)\frac{\dot a(X_{\frac{i}{n}},\sigma_0)}{a(X_{\frac{i}{n}},\sigma_0)} - 3\left(\xi_{i,n}'^2 - \frac{1}{3}\right)\frac{\dot a(X_{\frac{i-1}{n}},\sigma_0)}{a\left(X_{\frac{i-1}{n}},\sigma_0\right)} - 6\xi_{i-1,n}\xi_{i,n}'\frac{\dot a\left(X_{\frac{i-1}{n}},\sigma_0\right)}{a(X_{\frac{i-1}{n}},\sigma_0)} \right\}$$

and $o_{\mathbf{P}}(1)$ stands for $-3n^{-\frac{1}{2}}(\xi_{0,n}^2 - \frac{1}{3})\frac{\dot a(X_0,\sigma_0)}{a(X_0,\sigma_0)} - 3n^{-\frac{1}{2}}\frac{\dot a(X_{\frac{n-2}{n}},\sigma_0)}{a(X_{\frac{n-2}{n}},\sigma_0)}\left\{(\xi_{n-1,n}'^2 - \frac{1}{3}) + 2\xi_{n-2,n}\xi_{n-1,n}'\right\}$.

If suffices to prove that $\sum_{i=1}^{n-2}\chi_{i,n}$ converges stably in law to a variable Z, as recalled in 2). For this, we follow the proof of Genon-Catalot and Jacod [3] (pp. 138–139) (see Jacod [7] Th. 3.2, too): using that the variable $\chi_{i,n}$ is $\mathcal{G}_{i+1}^n$-measurable, it is sufficient to show the four following properties.

$$\sup_{t\in[0,1]}\left|\sum_{i=1}^{[(n-2)t]} E\left(\chi_{i,n} \mid \mathcal{G}_i^n\right)\right| \xrightarrow[\mathbf{P}]{n\to\infty} 0 \tag{38}$$

$$\forall t\in[0,1],\quad \sum_{i=1}^{[(n-2)t]} E\left(\chi_{i,n}^2 \mid \mathcal{G}_i^n\right) \xrightarrow[\mathbf{P}]{n\to\infty} 9\int_0^t \left(\frac{\dot a(X_s,\sigma_0)}{a(X_s,\sigma_0)}\right)^2 ds \tag{39}$$

$$\forall t\in[0,1],\quad \sum_{i=1}^{[(n-2)t]} E\left(\chi_{i,n}^4 \mid \mathcal{G}_i^n\right) \xrightarrow[\mathbf{P}]{n\to\infty} 0 \tag{40}$$

$$\forall t\in[0,1],\quad \sum_{i=1}^{[(n-2)t]} E\left(\chi_{i,n}(B_{\frac{i+1}{n}} - B_{\frac{i}{n}}) \mid \mathcal{G}_i^n\right) \xrightarrow[\mathbf{P}]{n\to\infty} 0. \tag{41}$$

The convergence (38) is immediate, since by Lemma 2.1, $E\left(\chi_{i,n} \mid \mathcal{G}_i^n\right) = 0$.

For (39), by Lemma 2.1 again,

$$E\left(\chi_{i,n}^2 \mid \mathcal{G}_i^n\right) = n^{-1}\left\{2\left(\frac{\dot a(X_{\frac{i}{n}},\sigma_0)}{a(X_{\frac{i}{n}},\sigma_0)}\right)^2 + 2\left(\frac{\dot a(X_{\frac{i-1}{n}},\sigma_0)}{a(X_{\frac{i-1}{n}},\sigma_0)}\right)^2 + \frac{\dot a(X_{\frac{i}{n}},\sigma_0)}{a(X_{\frac{i}{n}},\sigma_0)}\frac{\dot a(X_{\frac{i-1}{n}},\sigma_0)}{a(X_{\frac{i-1}{n}},\sigma_0)} \right.$$
$$\left. +12\left(\frac{\dot a(X_{\frac{i-1}{n}},\sigma_0)}{a(X_{\frac{i-1}{n}},\sigma_0)}\right)^2 \xi_{i-1,n}^2 \right\}.$$

Hence, the only difficult part, to establish the convergence (39) is to show

$$\sum_{i=1}^{[(n-2)t]} u_{i,n} \xrightarrow[\mathbf{P}]{n\to\infty} 4\int_0^t \left(\frac{\dot a(X_s,\sigma_0)}{a(X_s,\sigma_0)}\right)^2 ds, \quad \text{with } u_{i,n} = 12n^{-1}\left(\frac{\dot a(X_{\frac{i-1}{n}},\sigma_0)}{a(X_{\frac{i-1}{n}},\sigma_0)}\right)^2 \xi_{i-1,n}^2. \tag{42}$$

Remarking that $u_{i,n}$ is $\mathcal{G}_i^n$-measurable and that,

$$E(u_{i,n} \mid \mathcal{G}_{i-1}^n) = 4n^{-1}\left(\frac{\dot a(X_{\frac{i-1}{n}},\sigma_0)}{a(X_{\frac{i-1}{n}},\sigma_0)}\right)^2, \quad E(u_{i,n}^2 \mid \mathcal{G}_{i-1}^n) = 32n^{-2}\left(\frac{\dot a(X_{\frac{i-1}{n}},\sigma_0)}{a(X_{\frac{i-1}{n}},\sigma_0)}\right)^4,$$

we use Lemma B of the Appendix and obtain (42).

The convergence (40) is clear since $E\left(\left|E\left(\chi_{i,n}^4 \mid \mathcal{G}_i^n\right)\right|\right) \le cn^{-2}$.

Remarking that $E\left(\xi'_{i,n}(B_{\frac{i+1}{n}} - B_{\frac{i}{n}}) \mid \mathcal{G}_i^n\right) = \frac{1}{2}n^{-\frac{1}{2}}$, we get

$$E\left(\chi_{i,n}(B_{\frac{i+1}{n}} - B_{\frac{i}{n}}) \mid \mathcal{G}_i^n\right) = -3n^{-1}\frac{\dot{a}(X_{\frac{i-1}{n}},\sigma_0)}{a(X_{\frac{i-1}{n}},\sigma_0)}\xi_{i-1,n}.$$

Since $E(\xi_{i-1,n} \mid \mathcal{G}_{i-1}^n) = 0$, an appliciation of Lemma B of the Appendix yields (41).

Hence the property 2) is proved.

Now, we compute the second derivative of $U_n(\sigma)$ (see (36)) and then apply Theorem 3.1 to easily obtain the convergence, for all $\sigma \in \Theta$, in probability

$$\frac{\partial^2}{\partial\sigma^2}U_n(\sigma) \xrightarrow[\mathbf{P}]{n\to\infty} c(\sigma,\sigma_0), \tag{43}$$

with

$$c(\sigma,\sigma_0) = -2\int_0^1 a^2(X_s,\sigma_0)\left\{\frac{\ddot{a}(X_s,\sigma)}{a^3(X_s,\sigma)} - 3\frac{\dot{a}^2(X_s,\sigma)}{a^4(X_s,\sigma)}\right\}ds + 2\int_0^1\left\{\frac{\ddot{a}(X_s,\sigma)}{a(X_s,\sigma)} - \frac{\dot{a}^2(X_s,\sigma)}{a^2(X_s,\sigma)}\right\}ds.$$

Recalling the consistency of $\overline{\sigma}_n$, to obtain 1), it is enough to show that the convergence in (43) is uniform with respect to $\sigma \in \Theta$ and that $c(\sigma,\sigma_0) \xrightarrow[\mathbf{P}]{\sigma\to\sigma_0} 4\int_0^1 \frac{\dot{a}^2(X_s,\sigma_0)}{a^2(X_s,\sigma_0)}ds$.

The uniformity in (43) is obtain by an argument similar to the one used in the proof of Theorem 3.3.

By (35), we deduce that almost surely, $\sup_{s\in[0,1]}\mathcal{B}_l(X_s) < \infty$ and $\sup_{s\in[0,1]}\mathcal{B}_r(X_s) < \infty$. Hence, we can apply, almost surely, by $(S1)$, the Lebesgue dominated convergence theorem, and obtain the almost sure convergence (and hence the convergence in probability):

$$c(\sigma,\sigma_0) \xrightarrow[\mathbf{P}]{\sigma\to\sigma_0} c(\sigma_0,\sigma_0) = 4\int_0^1 \frac{\dot{a}^2(X_s,\sigma_0)}{a^2(X_s,\sigma_0)}ds. \tag{44}$$

So 1) is proved. □

Remark 3.5. *1) In the case of a multiplicative dependence for the parameter, $a(x,\sigma) = \sigma a(x)$, we have an explicit expression for our estimator of σ^2:*

$$\overline{\sigma_n^2} = \frac{3}{2}\sum_{i=0}^{n-2}\frac{(nJ_{i+1} - nJ_i)^2}{a^2(nJ_i)}.$$

Furthermore, this estimator is asymptotically Gaussian, since, in this case (37) reduces to $n^{\frac{1}{2}}(\overline{\sigma_n^2} - \sigma_0^2) \xrightarrow{n\to\infty} \mathcal{N}\left(0, \frac{9}{4}\sigma_0^4\right)$.

2) Recall that an asymptotically efficient estimator based on $X_{\frac{i}{n}}$ is defined as:

$$\sigma_n^* = \underset{\sigma\in\Theta}{\operatorname{arginf}}\, U_n^*(\sigma), \qquad \text{with } U_n^*(\sigma) = \sum_{i=0}^{n-1}\frac{\left(X_{\frac{i+1}{n}} - X_{\frac{i}{n}}\right)^2}{a^2\left(X_{\frac{i}{n}},\sigma\right)} + n^{-1}\sum_{i=0}^{n-1}\log a^2\left(X_{\frac{i}{n}},\sigma\right).$$

It is known that $n^{\frac{1}{2}}(\sigma_n^* - \sigma_0)$ converges in law to S', where S' is defined on an extension $(\Omega,\mathcal{G}_1,P)$ and is, conditionally on $\mathcal{G}_1$, centered Gaussian with variance $\frac{1}{2}\left(\int_0^1\left(\frac{\frac{\partial}{\partial\sigma}a(X_s,\sigma_0)}{a(X_s,\sigma_0)}\right)^2 ds\right)^{-1}$ (see e.g. Genon-Catalot and Jacod [3]). Hence, the asymptotic conditional variance of our estimator $\overline{\sigma}_n$ is slightly bigger than the one of σ_n^*.

3) Our identifiability Assumption (S2) is the same one as when $(X_{\frac{i}{n}})_{i=0,\dots,n-1}$ is observed (see Genon-Catalot and Jacod [3]).

4) For simplification of the proof of Theorem 3.4 we have supposed that $K_l = \infty$, but in Section 5 we shall see that K_l is finite when X is a Cox-Ingersoll-Ross process. In this case, a direct study of the estimator based on expansions of Section 2 yields the result (see Sect. 5.2.2).

4. Extension

In Sections 2 and 3 we gave results for the process $(\frac{J_0}{\Delta_n}, \dots, \frac{J_{n-1}}{\Delta_n})$, where $\frac{J_i}{\Delta_n}$ is the uniform mean, in the interval $[i\Delta_n, (i+1)\Delta_n]$ of X. Here, we consider more general means. We suppose that ϕ is a measurable, bounded, non-negative function defined on $[0,1]$, such that $\int_0^1 \phi(s)ds = 1$, and we define

$$\overline{X}_i(\phi) = \Delta_n^{-1} \int_{i\Delta_n}^{(i+1)\Delta_n} X_s \phi\left(\frac{s - i\Delta_n}{\Delta_n}\right) ds.$$

Remark that $(\overline{X}_i(\phi))_{i\geq 0}$ may be interpreted as the discrete sampling with step Δ_n of the convoluted function $X \star \psi^{\Delta_n}$, where $\psi^{\Delta_n}(x) = \Delta_n^{-1}\phi(-\frac{x}{\Delta_n})$.

In this section, we extend results of Sections 2 and 3 to $(\overline{X}_i(\phi))$. We do not give proofs for these results, since they are analogous to those of Sections 2 and 3.

We define:

$$\xi_{i,n}(\phi) = \frac{1}{\Delta_n^{\frac{3}{2}}} \int_{i\Delta_n}^{(i+1)\Delta_n} \int_{i\Delta_n}^{s} \phi\left(\frac{v - i\Delta_n}{\Delta_n}\right) dv dB_s \quad \text{for } i,\ n \geq 0$$

$$\xi'_{i+1,n}(\phi) = \frac{1}{\Delta_n^{\frac{3}{2}}} \int_{(i+1)\Delta_n}^{(i+2)\Delta_n} \int_{s-\Delta_n}^{(i+1)\Delta_n} \phi\left(\frac{v - i\Delta_n}{\Delta_n}\right) dv dB_s \quad \text{for } i \geq -1,\ n \geq 0.$$

Proposition 4.1. *Assume (A1–A3) and $2\alpha_1 < K_l$, then:*

$$\overline{X}_i(\phi) - X_{i\Delta_n} = a(X_{i\Delta_n})\Delta_n^{\frac{1}{2}}\xi'_{i,n}(\phi) + e_{i,n}(\phi)$$

where $e_{i,n}(\phi)$ satisfies inequalities analogous to (9, 10). Moreover, if k is a real number ≥ 1, for all i, n:

$$E\left(\left|\overline{X}_i(\phi) - X_{i\Delta_n}\right|^k \mid \mathcal{G}_i^n\right) \leq \Delta_n^{\frac{k}{2}} c(1 + \mathcal{B}_r^k(X_{i\Delta_n})).$$

Theorem 4.2. *Assume (A1–A3). We have*

$$\overline{X}_{i+1}(\phi) - \overline{X}_i(\phi) - b(\overline{X}_i(\phi))\Delta_n = a(X_{i\Delta_n})(\xi_{i,n}(\phi) + \xi'_{i+1,n}(\phi))\Delta_n^{\frac{1}{2}} + \varepsilon_{i,n}(\phi)$$

where $\varepsilon_{i,n}(\phi)$ is $\mathcal{G}_{i+2}^n$ measurable and satisfies inequalities analogous to (18–22).

Suppose, now, that $\Delta_n = n^{-1}$ and define

$$v = E((\xi_{i,n}(\phi))^2) = \int_0^1 \left(\int_0^s \phi(u)du\right)^2 ds$$

$$v' = E((\xi'_{i,n}(\phi))^2) = \int_0^1 \left(\int_s^1 \phi(u)du\right)^2 ds$$

$$k = E(\xi_{i,n}(\phi)\xi'_{i,n}(\phi)) = \int_0^1 \left(\int_0^s \phi(u)du \int_s^1 \phi(u)du\right) ds.$$

Theorem 4.3. *Assume (A1–A4) with $K_l = \infty$ and let f be as in Theorem 3.1. Then,*

$$\sum_{i=0}^{n-2} \left(\overline{X}_{i+1}(\phi) - \overline{X}_i(\phi)\right)^2 f(\overline{X}_i(\phi)) \xrightarrow[\mathbf{P}]{n\to\infty} (v+v') \int_0^1 a^2(X_s) f(X_s) ds.$$

If a depends on a unknown parameter σ, we introduce the contrast

$$U_n(\phi,\sigma) = (v+v')^{-1} \sum_{i=0}^{n-2} \frac{(\overline{X}_{i+1}(\phi) - \overline{X}_i(\phi))^2}{a^2(\overline{X}_i(\phi),\sigma)} + n^{-1} \sum_{i=0}^{n-2} \log(a^2(\overline{X}_i(\phi),\sigma)).$$

Let $\overline{\sigma}_n(\phi) = \operatorname{arginf}_{\sigma\in\Theta} U_n(\phi,\sigma)$ be the associated minimum contrast estimator.

Theorem 4.4. *Assume (A1–A4) with $K_l = \infty$ and (S1, S2),then*

$$n^{\frac{1}{2}}(\overline{\sigma}_n(\phi) - \sigma_0) \xrightarrow[\mathcal{D}]{n\to\infty} S(\phi),$$

where $S(\phi)$ is a variable defined on an extension of the space $(\Omega, \mathcal{G}_1, P)$ and such that conditionally on $\mathcal{G}_1$, $S(\phi)$ is centered Gaussian with variance $\left(\frac{1}{2} + \frac{k^2}{(v+v')^2}\right)\left(\int_0^1 \left(\frac{\frac{\partial}{\partial\sigma} a(X_s,\sigma_0)}{a(X_s,\sigma_0)}\right)^2 ds\right)^{-1}$.

5. Examples

5.1. **Diffusion on $\mathbb{R}$**

Here, $(l,r) = (-\infty,\infty)$ and we set $\mathcal{B}_{-\infty}(x) = 1$, $\mathcal{B}_{\infty}(x) = 1 + |x|$ (by this choice, we decide, for simplification, to use the same function, $\mathcal{B}_\infty$, to bound other functions near the two different bounds $-\infty$, ∞). Let X be the solution of $dX_t = a(X_t)dB_t + b(X_t)dt$, and assume that there exists c such that for all $x \in \mathbb{R}$, $|a(x)| + |b(x)| \le c(1+|x|)$; a and b are twice continuously diffentiable and their second derivatives have polynomial growth.

Then, it is immediate to check Assumptions (A1) and (A2). By the following proposition, Assumption (A3) is satisfied with $K_l = \infty$.

Proposition 5.1. *Assume that X is the solution of (5) and that (A1) holds. Furthermore, a and b are supposed such that for all $x \in \mathbb{R}$, $|a(x)| + |b(x)| \le c(1+|x|)$. Then, for all integer $k \ge 1$, there exists a constant $c(k)$ depending only on c and k such that:*

$$E\left(\sup_{s\in[t,t+1]} |X_s|^k \mid \mathcal{G}_t\right) \le c(k)(1+|X_t|^k). \tag{45}$$

Proof. We can suppose (by the Hölder inequality) that $k \ge 2$. For $s \in [t,t+1]$ we write:

$$X_s = X_t + \int_t^s a(X_v)dB_v + \int_t^s b(X_v)dv. \tag{46}$$

Using the Burkhölder inequality, we get ($c(k)$ may change from a line to another):

$$E\left(\sup_{u\in[t,s]} \left|X_u^k\right| \mid \mathcal{G}_t\right) \le c(k)\,|X_t|^k + c(k)E\left(\left(\int_t^s a^2(X_v)dv\right)^{\frac{k}{2}} \mid \mathcal{G}_t\right) + c(k)E\left(\left(\int_t^s |b|\,(X_v)dv\right)^k \mid \mathcal{G}_t\right).$$

Using that $k \geq 2$:

$$E\left(\sup_{u\in[t,s]}\left|X_u^k\right| \mid \mathcal{G}_t\right) \leq c(k)\left|X_t\right|^k + c(k)E\left(\int_t^s |a|^k(X_v) + |b|^k(X_v)dv \mid \mathcal{G}_t\right).$$

Using the bound on $|a|+|b|$, and the Fubini theorem:

$$E\left(\sup_{u\in[t,s]}\left|X_u^k\right| \mid \mathcal{G}_t\right) \leq c(k)\left|X_t\right|^k + c(k)\int_t^s E\left((1+|X_v|^k) \mid \mathcal{G}_t\right)dv. \tag{47}$$

Hence we have, if we set $\phi(s) = \sup_{u\in[t,s]} E\left(\left|X_u^k\right| \mid \mathcal{G}_t\right)$:

$$\phi(s) \leq c(k)(1+|X_t|^k) + c(k)\int_t^s E\left(|X_v|^k) \mid \mathcal{G}_t\right)dv \leq c(k)(1+|X_t|^k) + c(k)\int_t^s \phi(v)dv.$$

By (A1), $\phi(s)$ is almost surely finite, and we may apply Gronwall's Lemma to obtain:

$$\phi(s) \leq c(1+|X_t|^k)e^{c(k)(s-t)}.$$

Using that $s \in [t, t+1]$, we deduce:

$$\sup_{u\in[t,t+1]} E\left(\left|X_u^k\right| \mid \mathcal{G}_t\right) \leq c(k)(1+|X_t|^k).$$

Reporting the last inequation in (47) gives (45). □

As example, we may consider the case of X an Ornstein-Uhlenbeck process solution of $dX_t = -X_t dt + \sigma_0 dB_t,\ X_0 = 0$. Results of numerical simulations, given in Table 1, show that our estimator $\overline{\sigma^2}_n$ performs as well as the classical estimator, σ_n^{2*} based on $X_{\frac{i}{n}}$ (see Rem. 3.5 2)).

Table 1. (Mean; Standard deviation) of the estimator for $n = 100$ and different values of σ_0^2.

	$\sigma_0^2 = 0.01$	$\sigma_0^2 = 0.1$	$\sigma_0^2 = 1$	$\sigma_0^2 = 2$
$\overline{\sigma_n^2}$	0.0094; 0.0014	0.094; 0.013	0.93; 0.15	1.87; 0.27
σ_n^{2*}	0.0095; 0.0014	0.094; 0.013	0.94; 0.13	1.84; 0.26

5.2. **Positive diffusions**

In these models $l = 0$ and $r = \infty$.

5.2.1. *Exponential of a diffusion on* $\mathbb{R}$

Here $(l, r) = (0, \infty)$ and we set $\mathcal{B}_0(x) = 1 + \frac{1}{x}$ and $\mathcal{B}_\infty(x) = 1 + x^2$.

We assume that $X_t = \exp(Z_t)$, where Z_t is a diffusion on $\mathbb{R}$ defined as the solution of the equation: $dZ_t = \tilde{a}(Z_t)dB_t + \tilde{b}(Z_t)dt$ and $Z_0 = \eta$, is independent of B. Consider the following assumptions.

Functions $\tilde{a}$ and $\tilde{b}$ are defined on $\mathbb{R}$; $\tilde{a}$ is bounded; $\limsup_{z\to\infty}\tilde{b}(z) < \infty$; $\liminf_{z\to-\infty}\tilde{b}(z) > -\infty$; there exists c such that: for all $z \in \mathbb{R}$, $\left|\tilde{b}(z)\right| \leq c(1+|z|)$; $\tilde{a}$ and $\tilde{b}$ are twice continuously differentiable; their second derivatives have polynomial growth and $\tilde{a'}$ is bounded.

Then diffusion X satisfies the stochastic differential equation:

$$dX_t = a(X_t)dB_t + b(X_t)dt, \tag{48}$$

with $a(x) = x\tilde{a}(\ln x)$ and $b(x) = x\tilde{b}(\ln x) + \frac{1}{2}x\tilde{a}^2(\ln x)$. By assumptions, a' is bounded, so a has linear growth.

Proposition 5.2. *The diffusion (X_t) satisfies Assumptions (A1–A3), with $K_l = \infty$.*

Proof. (A1) and (A2) are clear. For (A3), we first show the inequality with $\mathcal{B}_\infty$. Using the Markov property for X, it suffices to show that, for all $k \geq 0$, and Y solution of (48) starting with $y_0 > 0$, there exists c such that:

$$E\left(\sup_{t\in[0,1]} Y_t^k\right) \leq c(1 + y_0^k). \tag{49}$$

Since $\tilde{a}$ is bounded and $\limsup_{x\to\infty} \tilde{b}(x) < \infty$ there exists a constant M such that $\forall\, x > 0,\ b(x) \leq M(x+1)$. We define $\hat{Y}$ as the strong solution, with the same Brownian motion B, of $d\hat{Y}_t = a(\hat{Y}_t)dB_t + M(\hat{Y}_t + 1)dt,\ \hat{Y}_0 = y_0$. Using a comparison theorem (see Revuz-Yor [12], p. 375), we get:

$$\forall t \geq 0, \quad Y_t \leq \hat{Y}_t \text{ a.s.} \tag{50}$$

Now, $\hat{Y}$ solves a stochastic differential equation with coefficients having at most a linear growth, so by Proposition 5.1:

$$E\left(\sup_{s\in[0,1]} \left|\hat{Y}_s\right|^k\right) \leq c(1 + y_0^k). \tag{51}$$

Then (50) and (51) imply (49), hence we get the first inequality of (A3).

Noticing that $\frac{1}{X} = \exp(-Z)$, and $-Z$ has the same properties as Z, we obtain the second inequality (with $\mathcal{B}_0$) analogously. □

Remark 5.3. *Hence, our results are valid for the exponantial of a Brownian motion or the exponential of an Ornstein-Uhlenbeck process.*

5.2.2. *Cox-Ingersoll-Ross process*

Again, here, $(l, r) = (0, \infty)$ and we set $\mathcal{B}_0(x) = 1 + \frac{1}{x}$ and $\mathcal{B}_\infty(x) = 1 + x$.

Let X be given by:

$$dX_t = (\alpha X_t + \beta)dt + \sigma\sqrt{X_t}dB_t, \quad X_0 = \eta, \tag{52}$$

with $\alpha < 0$, $\sigma, \beta > 0$ and η is a positive random variable independent of (B_t), and $\frac{2\beta}{\sigma^2} > 1$.

Assumptions (A1) and (A2) hold. Combining Propositions 5.1 with the following proposition, we see that (X_t) satisfies Assumption (A3), for $K_0 = \frac{2\beta}{\sigma^2} - 1$.

Proposition 5.4. *Assume that X is the solution of $dX_t = (\alpha X_t + \beta)dt + \sigma\sqrt{X_t}dB_t$, $X_0 = \eta > 0$ with $\alpha < 0$, $\beta > 0$, $\sigma > 0$ and $\frac{2\beta}{\sigma^2} > 1$. Let $k \in [0, \frac{2\beta}{\sigma^2} - 1)$, then $\exists c$ such that $\forall\ t \geq 0$:*

$$E\left(\sup_{s\in[t,t+1]} \left(\frac{1}{X_s}\right)^k \mid \mathcal{G}_t\right) \leq c\left(\frac{1}{X_t}\right)^k.$$

Proof. Using the Markov property, we get for $x>0$:

$$\mathbb{1}_{\{X_t=x\}} E\left(\sup_{s\in[t,t+1]}\left(\frac{1}{X_s}\right)^k \mid \mathcal{G}_t\right) = E\left(\sup_{s\in[0,1]}\left(\frac{1}{\tilde{X}_s}\right)^k\right),$$

where $\tilde{X}$ solves the same stochastic differential equation as X, with the initial condition $\tilde{X}_0 = x$.

Now, we use that the process $\tilde{X}$ can be represented as (see Leblanc [9]): $\tilde{X}_s = e^{\alpha s} R_{\tau(s)}$, where (R_s) is the square of a Bessel process of dimension δ, with $\delta = \frac{4\beta}{\sigma^2}$, starting from $\tilde{X}_0 = x$, and $\tau(s) = \frac{\sigma^2}{4}\frac{e^{|\alpha|s}-1}{|\alpha|}$ is a deterministic change of time.

We deduce that $\inf_{s\in[0,1]} \tilde{X}_s \geq e^{\alpha} \inf_{s\geq 0} R_s$, and hence:

$$E\left(\sup_{s\in[0,1]}\left(\frac{1}{\tilde{X}_s}\right)^k\right) \leq e^{|\alpha|} E\left(\frac{1}{\inf_{s\geq 0} R_s^k}\right).$$

But we know that, since $\delta > 2$, the law of $\inf_{s\geq 0} R_s$ is $xU^{\frac{2}{\delta-2}}$ where U is uniformly distributed on $[0,1]$ (see Revuz-Yor [12], p. 430 Ex. 1.18). So,

$$E\left(\frac{1}{\inf_{s\geq 0} R_s^k}\right) = \frac{1}{x^k} E\left(U^{\frac{-2k}{\delta-2}}\right) = \frac{1}{x^k}\int_0^1 u^{\frac{-2k}{\delta-2}} du \leq c\frac{1}{x^k}$$

since the integral above is finite for $\frac{2k}{\delta-2} < 1$ *i.e.* for $k < \frac{\delta}{2} - 1 = \frac{2\beta}{\sigma^2} - 1$. □

Since $K_0 < \infty$, we cannot apply results of Section 3. But, here, the estimator is explicit (Rem. 3.5 1)) and a direct study based on the expansions of Section 2 shows that, if $K_0 > 4$, $\overline{\sigma^2}_n \xrightarrow[\mathbf{P}]{n\to\infty} \sigma_0^2$ and that if $K_0 > 6$, $n^{\frac{1}{2}}(\overline{\sigma^2}_n - \sigma_0^2) \xrightarrow[\mathcal{D}]{n\to\infty} \mathcal{N}(0, \frac{9}{4}\sigma_0^4)$.

5.2.3. *Bilinear diffusion*

We set $\mathcal{B}_0(x) = 1 + \frac{1}{x}$ and $\mathcal{B}_\infty(x) = 1 + x$.

We suppose that:

$$dX_t = (\alpha X_t + \beta)dt + \sigma X_t dB_t, \quad X_0 = \eta, \tag{53}$$

with $\alpha < 0$, $\sigma, \beta > 0$ and η is a positive and independent of (B_t).

Using Propositions 5.1 and the following proposition, we get that the diffusion (X_t) satisfies Assumptions (A1–A3) with $K_l = \infty$.

Proposition 5.5. *Assume that X solves the equation*

$$dX_t = (\alpha X_t + \beta)dt + \sigma X_t dB_t,\ X_0 > 0 \tag{54}$$

with $\alpha < 0$, β, $\sigma > 0$.

Then $\forall k \geq 0$, $\exists c$ such that for all t:

$$E\left(\sup_{s\in[t,t+1]} \frac{1}{X_s^k} \mid \mathcal{G}_t\right) \leq c\left(1 + \frac{1}{X_t^k}\right).$$

Proof. Again by the Markov property of X, for $x > 0$,

$$\mathbb{1}_{\{X_t=x\}} E\left(\sup_{s\in[t,t+1]} \frac{1}{X_s^k} \mid \mathcal{G}_t\right) = E\left(\sup_{s\in[0,1]} \frac{1}{\tilde{X}_s^k}\right) \tag{55}$$

where $\tilde{X}$ is the solution of (54) starting with $\tilde{X}_0 = x$.

We set $Z_s = \frac{1}{\tilde{X}_s}$, then Z solves $dZ_s = -\sigma Z_s dB_s + \{(\sigma^2 - \alpha)Z_s - \beta Z_s^2\}ds$, with $Z_0 = \frac{1}{x}$. Now, we define Z' as the solution: $dZ'_s = -\sigma Z'_s dB_s + (\sigma^2 - \alpha)Z'_s ds$, with $Z'_0 = \frac{1}{x} = Z_0$.

Since $\beta < 0$, we can use a comparison theorem (see Revuz-Yor [12], p. 375) to obtain:

$$\forall s, \quad Z_s \le Z'_s \quad a.s. \tag{56}$$

We apply Proposition 5.1 to Z':

$$E\left(\sup_{s\in[0,1]} {Z'_s}^k\right) \le c(1 + \left|{Z'_0}^k\right|) = c\left(1 + \frac{1}{x^k}\right) \cdot \tag{57}$$

The definition of Z (55, 56) and (57) yield the result. □

6. Appendix

Proposition A. *Let $f \in \mathcal{C}^1(l,r)$ satisfy:*

$$\exists c \ \forall x \in (l,r) \ \ |f'(x)| \le c(\mathcal{B}_l^\gamma(x) + \mathcal{B}_r^\gamma(x))$$

then, for all integer $k \ge 1$, such that $k\gamma < K_l$:

$$E\left(\sup_{v\in[i\Delta_n,(i+1)\Delta_n]} |f(X_v) - f(X_{i\Delta_n})|^k \mid \mathcal{G}_i^n\right) \le c\Delta_n^{\frac{k}{2}}(\mathcal{B}_l^{k\gamma}(X_{i\Delta_n}) + \mathcal{B}_r^{k(1+\gamma)}(X_{i\Delta_n})). \tag{58}$$

Proof. We start with $f(x) = x$. Let

$$\delta_{i,n} = \sup_{v\in[i\Delta_n,(i+1)\Delta_n]} |X_v - X_{i\Delta_n}| \,. \tag{59}$$

Using (46) and the Burkholder inequality, we get:

$$E\left(\delta_{i,n} \mid \mathcal{G}_i^n\right) \le cE\left(\left(\int_{i\Delta_n}^{(i+1)\Delta_n} a^2(X_v)dv\right)^{\frac{k}{2}} \mid \mathcal{G}_i^n\right) + cE\left(\left(\int_{i\Delta_n}^{(i+1)\Delta_n} |b|\,(X_v)dv\right)^k \mid \mathcal{G}_i^n\right).$$

Hence,

$$E\left(\delta_{i,n}^k \mid \mathcal{G}_i^n\right) \le c\Delta_n^{\frac{k}{2}} E\left(\sup_{s\in[i\Delta_n,(i+1)\Delta_n]} |a|^k\,(X_v) \mid \mathcal{G}_i^n\right) + c\Delta_n^k E\left(\sup_{s\in[i\Delta_n,(i+1)\Delta_n]} |b|^k\,(X_v) \mid \mathcal{G}_i^n\right).$$

Using Assumption (A2), we get $|a|^k + |b|^k \le c\mathcal{B}_r^k$, so Assumption (A3) yields:

$$E\left(\delta_{i,n}^k \mid \mathcal{G}_i^n\right) \le c\Delta_n^{\frac{k}{2}}\mathcal{B}_r^k(X_{i\Delta_n}). \tag{60}$$

Now for a general f, we set

$$\delta_{i,n}(f) = \sup_{v\in[i\Delta_n,(i+1)\Delta_n]} |f(X_v) - f(X_{i\Delta_n})| \tag{61}$$

and write, using the bound on f':

$$\delta_{i,n}(f) \le c \sup_{v\in[i\Delta_n,(i+1)\Delta_n]} (\mathcal{B}_l^\gamma(X_v) + \mathcal{B}_r^\gamma(X_v))\delta_{i,n}.$$

We choose p and q such that $\frac{1}{p} + \frac{1}{q} = 1$ and $pk\gamma < K_l$ and apply Hölder's inequality (see (59–61)):

$$E\left(\delta_{i,n}(f)^k \mid \mathcal{G}_i^n\right) \le cE\left(\sup_{v\in[i\Delta_n,(i+1)\Delta_n]} (\mathcal{B}_l^{\gamma kp}(X_s) + \mathcal{B}_r^{\gamma kp}(X_s)) \mid \mathcal{G}_i^n\right)^{\frac{1}{p}} E\left(\delta_{i,n}^{qk} \mid \mathcal{G}_i^n\right).$$

Now, Assumption (A3) and (60) yield:

$$E\left(\delta_{i,n}(f)^k \mid \mathcal{G}_i^n\right) \le c(\mathcal{B}_l^{\gamma k}(X_{i\Delta_n}) + \mathcal{B}_r^{\gamma k}(X_{i\Delta_n}))(1 + \mathcal{B}_r^k(X_{i\Delta_n})).$$

So, we have the result. □

We recall, the useful lemma which is given in Genon-Catalot and Jacod [3].

Lemme B. *Let χ_i^n, U be random variables, with χ_i^n being $\mathcal{G}_i^n$-measurable. The following two conditions imply $\sum_{i=1}^n \chi_i^n \xrightarrow{\mathbf{P}} U$:*

$$\sum_{i=1}^n E(\chi_i^n \mid \mathcal{G}_{i-1}^n) \xrightarrow{\mathbf{P}} U$$
$$\sum_{i=1}^n E((\chi_i^n)^2 \mid \mathcal{G}_{i-1}^n) \xrightarrow{\mathbf{P}} 0.$$

References

[1] O.E. Barndorff-Nielsen and N. Shephard, *Aggregation and model construction for volatility models*. Working paper series No. 10. Center for Analytical Finance, University of Aarhus (1998).

[2] G. Dohnal, On estimating the diffusion coefficient. *J. Appl. Probab.* **24** (1987) 105–114.

[3] V. Genon-Catalot and J. Jacod, On the estimation of the diffusion coefficient for multidimensional diffusion processes. *Ann. Inst. H. Poincaré Probab. Statist.* **29** (1993) 119–151.

[4] V. Genon-Catalot, T. Jeantheau and C. Laredo, Limit theorems for discretely observed stochastic volatility models. *Bernoulli* **4** (1998) 283–303.

[5] A. Gloter, Parameter estimation for a discrete sampling of an integrated Ornstein–Uhlenbeck process. *Statistics* (to appear).

[6] J. Hull and A. White, The pricing of options on assets with stochastic volatilities. *J. Finance* **42** (1987) 281–300.

[7] J. Jacod, On continuous conditional Gaussian martingales and stable convergence in law. *Séminaire de Probabilités XXXI.* **1655**. Springer, Berlin, *Lectures Notes in Math.* (1997) 232–246.

[8] M. Kessler, Estimation of an ergodic diffusion from discrete observations. *Scand. J. Statist.* **24** (1997) 211–229.

[9] B. Leblanc, *Modélisation de la Volatilité d'un Actif Financier et Applications.* Thèse, Université Paris 7 (1997).

[10] M. Lefebvre, On the inverse of the first hitting time problem for bidimensional processes. *J. Appl. Probab.* **34** (1997) 610–622.

[11] S. Pastorello, E. Renault and N. Touzi, Statistical inference for random variance option pricing. *Southern European Economics Discussion Series*, D.P.136 (1994).

[12] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion.* Springer–Verlag, Berlin Heidelberg, second edition (1994).