ESAIM: PS
August 2004, Vol. 8, p. 132–149
DOI: 10.1051/ps:2004005

ESAIM: Probability and Statistics

ASYMPTOTICS FOR THE L^p-DEVIATION OF THE VARIANCE ESTIMATOR UNDER DIFFUSION *

Paul Doukhan[1] and José R. León[2]

Abstract. We consider a diffusion process X_t smoothed with (small) sampling parameter ε. As in Berzin, León and Ortega (2001), we consider a kernel estimate $\widehat{\alpha}_\varepsilon$ with window $h(\varepsilon)$ of a function α of its variance. In order to exhibit global tests of hypothesis, we derive here central limit theorems for the L^p deviations such as

$$\frac{1}{\sqrt{h}}\left(\frac{h}{\varepsilon}\right)^{\frac{p}{2}}\left(\|\widehat{\alpha}_\varepsilon-\alpha\|_p^p-\mathbb{E}\,\|\widehat{\alpha}_\varepsilon-\alpha\|_p^p\right).$$

Mathematics Subject Classification. 60F05, 60F25, 60J60, 60H05, 62M02, 62M05.

Received October 22, 2002. Revised October 21, 2003.

1. Introduction and main results

Recently, the problem of statistical inference for integrated diffusions attracted a certain attention. This type of situation appears, for example, when a realization of a process is observed after passage through an electronic filter. Ditlevsen *et al.* [4] studied the ice core records from Greenland, which can also be modeled as an integrated diffusion process. Gloter [9] considers the integrated Ornstein-Uhlenbeck process to make inference about its parameters. Integrated processes are also important in the so-called realized volatility in finance. In all these works the observations are assumed to be $Y_i=\int_{(i-1)\Delta}^{i\Delta}X_s\varphi_\Delta(s-(i-1)\Delta)\,\mathrm{d}s$ where X_s is a diffusion process and φ_Δ is a density in the interval $[0,\Delta]$. Setting $\varphi_\Delta(s)=\frac{1}{\Delta}\varphi(\frac{s}{\Delta})$, we have $Y_i=\int_0^1\varphi(s)X_{s\Delta+(i-1)\Delta}\,\mathrm{d}s$.

In this work we observed a smoothed and continuous time process defined as

$$X_t^\varepsilon=\frac{1}{\varepsilon}\int_{-\infty}^{\infty}\varphi\left(\frac{t-u}{\varepsilon}\right)X_u\,\mathrm{d}u=\int_{-1/2}^{1/2}\varphi(s)X_{t-\varepsilon s}\,\mathrm{d}s, \tag{1}$$

with X_t a diffusion process, φ is a density in $[-\frac{1}{2},\frac{1}{2}]$ and ε is interpreted as a sampling parameter. Our goal is to make non parametric inference for the diffusion coefficient. Let us briefly introduce our framework.

Keywords and phrases. Variance estimator, kernel, L^p-deviation, central limit theorem.

* *This work was supported by the Project Agenda Petróleo: Modelaje Estocástico Aplicado of FONACIT Venezuela.*

[1] Laboratoire de Statistiques LS-CREST, ENSAE, 3 rue Pierre Larousse, France; e-mail: doukhan@ensae.fr

[2] Universidad Central de Venezuela, Escuela de Matemática, Facultad de Ciencias, AP. 47197, Los Chaguaramos Caracas 1041-A, Venezuela; e-mail: jleon@euler.ciens.ucv.ve

Let $(W_t)_{t\geq 0}$ be a standard Brownian motion, and X_t be defined by the equation

$$\mathrm{d}X_t = \sigma(t)\,\mathrm{d}W_t + b(X_t)\,\mathrm{d}t, \quad \sigma > 0. \tag{2}$$

To assure the existence and uniqueness of the solution X_t, $\sigma : \mathbb{R} \to \mathbb{R}$ and $b : \mathbb{R} \to \mathbb{R}$ are assumed to satisfy the assumptions (1) and (2) of Theorem 1 of [7] p. 40, adapted to our case, *i.e.*

$$|b(x) - b(y)| \leq K|x - y| \text{ and } |b(x)| + \sigma(x) \leq K^2\left(1 + |x|^2\right). \tag{3}$$

Additional conditions concerning σ will be given later. This shows however that explosive variances cannot be considered in the present work.

In this work we consider the estimation of the function $\sigma(t)$. Process X_t is not directly available, we assume that we observe X_t^ε as defined in (1). The case in which function $\sigma(\cdot)$ only depends on X_t has been studied previously by Perera and Wschebor in [12] and [13].

As in [2], we consider a function $G \in L^2(\phi)$ with $\phi(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}$ together with the continuous and symmetric densities φ and K with support in $[-\frac{1}{2}, \frac{1}{2}]$. If φ is a differentiable function we define $\dot{X}_\varepsilon(t) = \frac{\mathrm{d}}{\mathrm{d}t}X_\varepsilon(t)$.

For any $q \geq 1$, we define $\|f\|_q = \left(\int_{-\infty}^{\infty} |f(t)|^q\,\mathrm{d}t\right)^{\frac{1}{q}}$. Let $h = h(\varepsilon) \to 0$ as $\varepsilon \to 0$ (the dependence of h on ε is implicit throughout the paper) and we set

$$\widehat{\alpha}_\varepsilon(t) = \frac{1}{h}\int_{-\infty}^{\infty} K\left(\frac{t-u}{h}\right) G\left(\frac{\sqrt{\varepsilon}}{\|\varphi\|_2}\dot{X}_\varepsilon(u)\right)\mathrm{d}u. \tag{4}$$

So $\widehat{\alpha}_\varepsilon(t)$ is the non-parametric kernel estimate of the parameter

$$\alpha(t) = \mathbb{E}[G(\sigma(t)Z)], \qquad t \in [0,1] \tag{5}$$

where $Z \sim \mathcal{N}(0,1)$ will denote a standard normal random variable throughout the paper. Berzin *et al.* [2] note several interesting special cases:

- if $G(x) = x^2$ then $\alpha(t) = \sigma^2(t)$ (recall that $\mathbb{E}|Z|^2 = 1$);
- if $G(x) = \sqrt{\frac{\pi}{2}}|x|$ then $\alpha(t) = \sigma(t)$ (recall that $\mathbb{E}|Z| = \sqrt{\frac{2}{\pi}}$);
- if $G(x) = \log|x| - 2\gamma$ then $\alpha(t) = \log\sigma(t)$. For this, note that the constant γ can be written as $\gamma = \int_0^\infty \log x\ \phi(x)\,\mathrm{d}x = 0.57721566\ldots$

By using stable convergence, as in [2], we can deduce our results from the case $b \equiv 0$. In this particular case our process is a time-changed Brownian motion.

Define

$$\beta_\varepsilon(t) = \sqrt{h/\varepsilon}(\widehat{\alpha}_\varepsilon(t) - \mathbb{E}\widehat{\alpha}_\varepsilon(t)). \tag{6}$$

A pointwise central limit theorem (CLT)

$$\beta_\varepsilon(t) \xrightarrow{\mathcal{D}}_{\varepsilon\to 0} \mathcal{N}\left(0, \Sigma^2(t)\right)$$

is proved in [2], where $\Sigma^2(t)$ is defined by equation (11). Alternative estimation techniques and some CLT are proposed in Soulier [17], Genon-Catalot *et al.* [6] and in Brugière [3] under close settings.

Another expression will also be useful

$$\widehat{\beta}_\varepsilon(t) = \sqrt{h/\varepsilon}(\widehat{\alpha}_\varepsilon(t) - \alpha(t)). \tag{7}$$

If $\alpha \in C^2$ (twice continuously differentiable) then the bias term verifies $\mathbb{E}(\widehat{\alpha}_\varepsilon(t)) - \alpha(t) = O(h^2)$. In this case the optimal window size is $h = \varepsilon^{1/5}$. Replacing $\mathbb{E}(\widehat{\alpha}_\varepsilon(t))$ by $\alpha(t)$ in (6) we get again a CLT with no zero mean (resp. zero mean) in the optimal case (resp. in the non optimal case). In Proposition 1 below we will precise the asymptotic bias behavior.

In the present paper, our aim is to provide global estimation results for parameter α in L^p for $p \geq 1$. So, we consider the L^p deviations

$$D_{p,\varepsilon} = \frac{1}{\sqrt{h}} \left(\|\beta_\varepsilon\|_p^p - \mathbb{E}\|\beta_\varepsilon\|_p^p \right), \quad \text{and} \tag{8}$$

$$\mathcal{D}_{p,\varepsilon} = \frac{1}{\sqrt{h}} \left(\|\widehat{\beta}_\varepsilon\|_p^p - \mathbb{E}\|\widehat{\beta}_\varepsilon\|_p^p \right). \tag{9}$$

In Theorems 1 and 2 we show that both expressions are asymptotically normal. These results are used to design global tests of hypotheses for the diffusion's variance in the forthcoming section. That test appears to be of special interest in problems in finance. Using a Poissonization argument, Beirlant and Mason [1] obtained analogous results for the case of kernel density and regression estimates based on independent samples. Soulier [17] proves a CLT for the case $p = 2$ for a wavelet-based estimator of the diffusion coefficient.

Remark. The asymptotic behavior of $\mathcal{D}_{p,\varepsilon}$, more convenient for a test of hypothesis, is an easy consequence of the behavior for $D_{p,\varepsilon}$ in the sub optimal window case, nevertheless needs a more detailed analysis in the optimal one (see remark and proof of Th. 2).

Let us expand the (even) function $g_t(x) = G(\sigma(t)x)$ in terms of Hermite polynomials

$$g_t(x) = \sum_{n=0}^{\infty} a_{2n}(t) H_{2n}(x), \text{ with } a_{2n}(t) = \frac{1}{(2n)!} \mathbb{E}G(\sigma(t)Z) \cdot H_{2n}(Z). \tag{10}$$

We will often use Mehler's formula: $\mathbb{E}(H_n(X)H_m(Y)) = n!\rho^n \delta_{n,m}$, where (X, Y) a two-dimensional standard Gaussian vector having correlation ρ, which is a special case of the Diagram formula, see [11].

Let $f \star g$ stand for the convolution of f and g. For $t \in [0, 1]$ and $w \in [-1, 1]$, we define

$$\Sigma^2(t) = \|K\|_2^2 \sum_{n=1}^{\infty} a_{2n}^2(t)\ (2n)! \int_{-1}^{1} \left(\frac{\varphi \star \varphi(z)}{\|\varphi\|_2^2} \right)^{2n} \mathrm{d}z, \qquad \text{and} \tag{11}$$

$$\Gamma(w) = \frac{K \star K(w)}{\|K\|_2^2} \in [-1, 1]. \tag{12}$$

Let (Z_1, Z_2) be a standard $(0, I_2)$ normal vector, so we define

$$\Sigma_p^2 = \int_{-1}^{1} \mathrm{Cov}\ \left(\left| \sqrt{1 - \Gamma^2(w)} Z_1 + \Gamma(w) Z_2 \right|^p, |Z_2|^p \right) \mathrm{d}w \cdot \int_0^1 \Sigma^{2p}(t)\, \mathrm{d}t. \tag{13}$$

We have the following result

Theorem 1. *Assume that the diffusion (2) is such as the function σ is continuous and $\sigma > 0$ over the compact set $[0, 1]$, there exists some $q \geq 4$ such as $\mathbb{E}|G(\sigma(t)Z)|^{pq} < \infty$ and $\lim_{\varepsilon \to 0} h = \lim_{\varepsilon \to 0} \frac{\varepsilon}{h^{2(1-1/q)}} = 0$. Then*

$$D_{p,\varepsilon} \xrightarrow{\mathcal{D}}_{\varepsilon \to 0} \mathcal{N}\left(0, \Sigma_p^2\right).$$

Remarks. Using Lemma 5 below proves that the same CLT holds for

$$\widetilde{D}_{p,\varepsilon} = \frac{1}{\sqrt{h}} \left(\|\beta_\varepsilon\|_p^p - \mathbb{E}|Z|^p \int_0^1 (\Sigma(t))^p \, \mathrm{d}t \right),$$

where $\Sigma^2(t)$ is defined in equation (11) and it is also the limiting variance in the CLT as proved in [2].

Let $b_{2k} = \frac{1}{(2k)!}\mathbb{E}|Z|^p H_{2k}(Z)$, then we write

$$\Sigma_p^2 = \sum_{k=1}^{\infty} b_{2k}^2\,(2k)! \int \Gamma^{2k}(w)\,\mathrm{d}w \int \Sigma^{2p}(t)\,\mathrm{d}t.$$

Inspired by Jacod [10], the proof of Theorem 1 is divided in two steps: first one assumes that $b \equiv 0$, which mean that X_t is a time-changed Brownian motion, and shows stable convergence. Secondly, using Girsanov's formula we consider the case $b \neq 0$.

We first provide the asymptotic behaviour of the bias:

Proposition 1. *Assume that the even function G is a.s. twice differentiable and assume that $\sigma > 0$ is a C^2-function. Let $b_\varepsilon(t) = \mathbb{E}\widehat{\alpha}_\varepsilon(t) - \alpha(t)$. and $\lim_{\varepsilon\to 0} h = \lim_{\varepsilon\to 0} \frac{\varepsilon}{h} = 0$, then*

$$\lim_{\varepsilon\to 0} h^{-2} \sup_{t\in[0,1]} \left| b_\varepsilon(t) - \frac{h^2}{2}\alpha''(t) \int s^2 K(s)\,\mathrm{d}s \right| = 0.$$

Besides, $\lim_{\varepsilon\to 0} \frac{\varepsilon^2}{h^3} = 0$ and the functions G, σ are C^3, imply that the norming factor h^{-2} may be replaced by h^{-3}.

Remark. As usual, the use of kernels K with higher order[1] yields $b_\varepsilon(t) = c\alpha^{(r)}(t)h^r + o(h^r)$ where o is uniform with respect to t. Moreover the remainder term is $O(h^{r+\delta})$ if the more restrictive assumption $|\alpha^{(r)}(s) - \alpha^{(r)}(t)| \leq c_\delta |s-t|^\delta$ is assumed.

We now turn to the asymptotic behaviour of $\mathcal{D}_{p,\varepsilon}$. Assuming that the functions σ, G are *a.s.* twice differentiable, then the suboptimal window case, $\lim_{\varepsilon\to 0} h^5/\varepsilon = 0$ leads to the same result as Theorem 1.

$$\mathcal{D}_{p,\varepsilon} \xrightarrow{\mathcal{D}}_{\varepsilon\to 0} \mathcal{N}(0, \Sigma_p^2) \qquad \text{if} \qquad \lim_{\varepsilon\to 0} \frac{h^5}{\varepsilon} = 0. \tag{14}$$

Examining the optimal window case $h = \lambda\varepsilon^{\frac{1}{5}}$, we get:

Theorem 2. *Assume that the function $\sigma > 0$ is C^2 (twice continuously differentiable), and that G is a.s. twice differentiable and has a second order bounded derivative set $h = \lambda\varepsilon^{\frac{1}{5}}$ for some constant $\lambda > 0$. If $\mathbb{E}|G(\sigma(t)Z)|^{2p} < \infty$ then*

$$\mathcal{D}_{p,\varepsilon} \xrightarrow{\mathcal{D}}_{\varepsilon\to 0} \mathcal{N}(0, \tau_p^2),$$

where, as in Theorem 1,

$$\tau_p^2 = \int\int \Theta(w,t)\ \Sigma^{2p}(t)\,\mathrm{d}w\,\mathrm{d}t \quad \textit{and}\ c(t) = \frac{1}{2}\lambda^{\frac{5}{2}}\alpha''(t)\int s^2 K(s)\,\mathrm{d}s,$$
$$\Theta(w,t) = \mathrm{Cov}\left(\left|\sqrt{1-\Gamma^2(w)}Z_1 + \Gamma(w)Z_2 + c(t)\right|^p, |Z_2 + c(t)|^p \right).$$

Remark. The statistic $\mathcal{D}_{p,\varepsilon}$ is not well adapted to make a hypothesis test. It can be modified as

$$\mathcal{D}_{p,\varepsilon,so} = \frac{1}{\sqrt{h}} \left(\|\widehat{\beta}_\varepsilon\|_p^p - \mathbb{E}|Z|^p \int_0^1 (\Sigma(t))^p\,\mathrm{d}t \right),$$

[1]This means that $\int t^k K(t)\,\mathrm{d}t$ is 1 for $k = 0$, vanishes for $0 < k < r$ and is well defined for $k = r$; usually one considers bounded and compactly supported kernels and a standard construction of such kernel consists to search polynomials P such as those conditions hold for the kernel $K = P \cdot u$ where u denotes a bounded and compactly supported non-negative density.

under the suboptimal window case and as

$$\mathcal{D}_{p,\varepsilon,o} = \frac{1}{\sqrt{h}} \left(\left\| \widehat{\beta}_\varepsilon \right\|_p^p - \int_0^1 |\Sigma(t) Z + c(t)|^p \, \mathrm{d}t \right),$$

if $h = \lambda \varepsilon^{\frac{1}{5}}$. In Lemma 5 below we will show that these two statistics have the same asymptotic behaviour that $\mathcal{D}_{p,\varepsilon}$ in each case.

Examples. In some special cases of interest, the function G is homogeneous $G(\sigma x) = \sigma^r G(x)$ for $\sigma > 0$, hence $\Sigma^2(t) = A\sigma^{2r}(t)$ for a suitable constant $A > 0$ only depending on ϕ and on G and, this makes much simpler the expressions of Σ_p^2 and τ_p^2. Examples of these situations $G(x) = \sqrt{\frac{\pi}{2}}|x|$ and $G(x) = x^2$ have already been sketched.

Analogous considerations are valid for the function $G(x) = \log|x| - \gamma$ for which only $a_0(t) = \log\sigma(t) - \gamma$ really depends on t while $a_{2n}(t) = a_{2n} = \frac{1}{(2n)!} \mathbb{E} \log|Z| H_{2n}(Z)$ for $n > 0$, and $\Sigma^2(t) = \Sigma_\varphi^2$ only depends on φ. Note that Σ_p^2 does not depend on the function $\sigma(\cdot)$; this however does not hold for the companion variance τ_p^2.

This paper is organized as follows, Section 1 introduces the problem and gives the main results. Section 2 is devoted to provide an explicit expression for a test of hypothesis useful for various applications. Section 3 is devoted to a series of technical lemmas useful in the proof of the main results. The main results are proved in Section 4, while the proof of the preliminary lemmas is given in Section 5.

2. Application to a test of hypothesis

Assume that we want to provide a test for hypothesis $H_0 : \alpha = a$ against a family of contiguous alternatives $\alpha = a + \delta_\epsilon A$ where the functions $a, A \in L^p$ are given, $\delta_\epsilon \downarrow 0$ as $\epsilon \downarrow 0$, and $\|A\|_p > 0$.

From the examples of functions G, it is clear that such tests can be transferred to σ, testing now $\sigma = s$ against $\sigma = s + \delta_\epsilon S$ where $a(t) = \mathbb{E}G(s(t)Z)$ and $A(t) = S(t)\mathbb{E}G'(s(t)Z)$ in the case of a differentiable function G. The interesting cases $G(x) = x^2$ and $\sqrt{\frac{\pi}{2}}|x|$ are straightforward.

We now set

$$\beta_\epsilon^a(t) = \sqrt{\frac{h}{\epsilon}}(\alpha_\epsilon(t) - a(t)).$$

Under the null hypothesis, $\alpha = a$ the remark following Theorem 2 implies

$$\mathcal{D}_{p,\varepsilon,so} = \frac{1}{\sqrt{h}} \left(\|\beta_\epsilon^a\|_p^p - \mathbb{E}|Z|^p \int_0^1 (\Sigma(t))^p \, \mathrm{d}t \right) \to \mathcal{N}\left(0, \Sigma_p^2\right),$$

if $\lim_{\epsilon\to 0} h^5/\epsilon = 0$. This gives a level for a test, provided we have estimated both expressions

$$\Sigma_p^2, \qquad \int_0^1 \Sigma^p(t) \, \mathrm{d}t$$

through empirical standard procedures. To this aim we only make use of the classical plug-in principle.

Passing now to the alternatives and setting

$$\gamma_\epsilon = \delta_\epsilon \sqrt{\frac{h}{\epsilon}}, \qquad \text{with} \qquad \delta_\epsilon = \left(\frac{\epsilon}{h^{(1-1/p)}}\right)^{1/2}.$$

We have $\beta_\epsilon^a(t) = \beta_\epsilon(t) + \gamma_\epsilon A(t) + \mathcal{O}_{L^p}\left(h^2\sqrt{\frac{h}{\epsilon}}\right)$ and $\gamma_\epsilon = h^{\frac{1}{2p}}$. Obtaining

$$\mathcal{D}_{p,\varepsilon,so} = \frac{1}{\sqrt{h}} \left(\|\beta_\varepsilon\|_p^p - \mathbb{E}|Z|^p \int_0^1 (\Sigma(t))^p \, \mathrm{d}t \right) + \|A\|_p^p + \frac{1}{\sqrt{h}} p \gamma_\epsilon^{p-1} \int \beta_\epsilon(t) |A(t)|^{p-1} \mathrm{sign}(A(t)) \, \mathrm{d}t + o(1).$$

We observe that
$$\frac{1}{\sqrt{\epsilon}}\int(\alpha_\epsilon(t)-\mathbb{E}\alpha_\epsilon(t))|A(t)|^{p-1}\mathrm{sign}(A(t))\,\mathrm{d}t \to_{\epsilon\to 0}\mathcal{N}(0,\sigma^2_{a,A}),$$
for a suitable constant $\sigma^2_{a,A}>0$. All this entails
$$\mathcal{D}_{p,\varepsilon,so}\to\mathcal{N}\left(\|A\|_p^p,\Sigma_p^2\right),$$
as is usual under contiguous alternatives. This provides a control of the local power for this procedure.

Even in the case where one considers the optimal window $h=\lambda\varepsilon^{\frac{1}{5}}$ it is possible to work out a test of hypothesis. In this particular case we must use $\mathcal{D}_{p,\varepsilon,o}$ instead of $\mathcal{D}_{p,\varepsilon,so}$, obtaining a similar result. This is based on Lemma 5 and the remark located after Theorem 2.

3. Collecting some facts in the case $b\equiv 0$

The following simple facts are essentially collected from [2]. Set
$$\dot\sigma^2_\varepsilon(t)=\mathrm{Var}\ \dot X_\varepsilon(t). \tag{15}$$

Lemma 1. *We have*
$$\begin{aligned}\mathrm{Cov}\ (\dot X_\varepsilon(s),\dot X_\varepsilon(t)) &= \frac{1}{\varepsilon^2}\int\varphi\left(\frac{s-u}{\varepsilon}\right)\varphi\left(\frac{t-u}{\varepsilon}\right)\sigma^2(u)\,\mathrm{d}u\\ &= \frac{1}{\varepsilon}\int\varphi(x)\varphi\left(x+\frac{t-s}{\varepsilon}\right)\sigma^2(t-\varepsilon x)\,\mathrm{d}x.\end{aligned}$$

This expression vanishes if $|t-s|>2\varepsilon$ *and we have* $\sqrt{\varepsilon}\dot\sigma_\varepsilon(t)\to\|\varphi\|_2\sigma(t)$ *as* $\varepsilon\to 0$ *where the previous convergence holds uniformly on* $[0,1]$.

We often work with the following "almost" white noise process which we shall denote for simplicity's sake
$$Z_\varepsilon(t)=\frac{\dot X_\varepsilon(t)}{\dot\sigma_\varepsilon(t)}\sim\mathcal{N}(0,1). \tag{16}$$

Setting $\rho_\varepsilon(s,t)=\mathrm{Cov}\ (Z_\varepsilon(s),Z_\varepsilon(t))$, note that the previous lemma implies
$$\rho_\varepsilon(s,t)=\frac{\int\varphi(x)\varphi\left(x+\frac{t-s}{\varepsilon}\right)\sigma^2(t-\varepsilon x)\,\mathrm{d}x}{\sqrt{\int\varphi^2(x)\sigma^2(s-\varepsilon x)\,\mathrm{d}x\ \int\varphi^2(x)\sigma^2(t-\varepsilon x)\,\mathrm{d}x}},$$
which yields
$$\mathbb{E}(\beta_\varepsilon(t))^2\sim\frac{h}{\varepsilon}\int\int K(u)K(v)\,\mathrm{Cov}\ \big(G(\sigma(t)Z_\varepsilon(t-uh)),G(\sigma(t)Z_\varepsilon(t-vh))\big)\,\mathrm{d}u\,\mathrm{d}v.$$

The above covariance is a function of t and of $\rho_\varepsilon(t-uh,t-vh)$. Using Mehler's formula we prove that
$$\mathbb{E}(\beta_\varepsilon(t))^2\sim\frac{h}{\varepsilon}\iint K(u)K(v)\sum_{n=1}^{\infty}a_{2n}(t-uh)a_{2n}(t-vh)(2n)!\left(\frac{\int\varphi(x)\varphi(x+\frac{h(v-u)}{\varepsilon})\sigma^2(t-uh-\varepsilon x)\,\mathrm{d}x}{\int\varphi^2(x)\sigma^2(t-uh-\varepsilon x)\mathrm{d}x}\right)^{2n}\mathrm{d}u\,\mathrm{d}v.$$

Finally, the change of variable $z=\frac{h(v-u)}{\varepsilon}$ implies $\mathbb{E}(\beta_\varepsilon(t))^2\to\Sigma^2(t)$ as $\varepsilon\to 0$ where equation (11) defines $\Sigma(t)$.

The process $\beta_\varepsilon(t)$ is not Gaussian (even if asymptotically Gaussian) its L^p-norm cannot be computed using Mehler's formula. Another way to proceed is as used in Giné *et al.* [8] using a Gaussian approximation. We first note that $\beta_\varepsilon(t)$ may be rewritten as the partial sum of 1-dependent random variables.

Lemma 2. *Set* $N = 2\left[\frac{h}{2\varepsilon}\right]$*, then we have* $\beta_\varepsilon(t) = \sum_{k=1}^{N} \zeta_{k,\varepsilon}(t)$ *with*

$$\zeta_{k,\varepsilon}(t) = \int_{-\frac{1}{2}+\frac{k-1}{N}}^{-\frac{1}{2}+\frac{k}{N}} K(u)\left(G\left(\frac{\sqrt{\varepsilon}\dot{\sigma}_\varepsilon(t-uh)}{\|\varphi\|_2}\cdot Z_\varepsilon(t-uh)\right) - \mathbb{E}G\left(\frac{\sqrt{\varepsilon}\dot{\sigma}_\varepsilon(t-uh)}{\|\varphi\|_2}\cdot Z_\varepsilon(t-uh)\right)\right)\mathrm{d}u,$$

and the N *random variables* $\zeta_{k,\varepsilon}(t)$ *are 1-dependent for* $k = 1, \ldots, N$*.*

Given that the random variables $Z_\varepsilon(t)$ and $Z_\varepsilon(s)$ are independent when $|s-t| > 2\varepsilon$. We obtain this lemma from relation $h/N > \varepsilon$, which also yields

Lemma 3. *Set* $M = \left[\frac{1}{2h}\right]$*, then* $\|\beta_\varepsilon\|_p^p = \sum_{\ell=1}^{M} Y_{\ell,\varepsilon}$ *with*

$$Y_{\ell,\varepsilon} = \left(\frac{h}{\varepsilon}\right)^{p/2} \int_{(\ell-1)/M}^{\ell/M} \left| \int K(u)\left(G\left(\frac{\sqrt{\varepsilon}\dot{\sigma}_\varepsilon(t-uh)}{\|\varphi\|_2}\cdot Z_\varepsilon(t-uh)\right) - \mathbb{E}G\left(\frac{\sqrt{\varepsilon}\dot{\sigma}_\varepsilon(t-uh)}{\|\varphi\|_2}\cdot Z_\varepsilon(t-uh)\right)\right)\mathrm{d}u\right|^p \mathrm{d}t,$$

and the M *random variables* $Y_{\ell,\varepsilon}$ *are* 2*-dependent for* $\ell = 1, \ldots, M$ *if we assume that* $h > 2\varepsilon$*.*

Hence, the technique of proof of the main theorem will be based on a Lindeberg central limit theorem for m-dependent random variables. The two first moments of the above random variable are difficult to calculate directly. Thus, in order to avoid this problem we shall proceed as in Giné *et al.* [8]: by using a Gaussian approximation of the previous sums $\beta_\varepsilon(t)$.

The proof of the main theorem will be based on the following lemmas which will provide (in particular) the asymptotic L^2 behaviour of $\|\beta_\varepsilon\|_p^p$.

Lemma 4 (approximating expectations)**.** *Let* $d \in \mathbb{N}$ *and* $x_{1,n}, \ldots, x_{n,n} \in \mathbb{R}^d$ *be centered at expectation,* m*-dependent for some integer* $m \geq 0$*. Denoting by* $\mathrm{Var}\left(\sum_{k=1}^n x_{k,n}\right)$ *the variance covariance matrix of the vector* $\sum_{k=1}^n x_{k,n}$*, supposing that for some definite* $d \times d$ *covariance matrix* V *we have*

$$\mathrm{Var}\left(\sum_{k=1}^n x_{k,n}\right) \to_{n\to\infty} V, \text{ and}$$

$$\sum_{k=1}^n \mathbb{E}\|x_{k,n}\|^{3\vee(dpq)} \to_{n\to\infty} 0.$$

Denoting $x_{j,n} = (x_{j,n}^{(\ell)})_{1\leq\ell\leq d}$*. Then, if* $\mathbf{Z} = (Z^{(1)}, \ldots, Z^{(d)}) \sim \mathcal{N}_d(0, V)$*, there exists a constant* c *(only depending on* d *and on the norm* $\|\cdot\|$ *on* $\mathbb{R}^d$*) such as*

$$\left|\mathbb{E}\prod_{\ell=1}^d \left|\sum_{k=1}^n x_{k,n}^{(\ell)}\right|^p - \mathbb{E}\prod_{\ell=1}^d \left|Z^{(\ell)}\right|^p\right| \leq c\left(\sum_{k=1}^n \mathbb{E}\|x_{k,n}\|^3\right)^\delta,$$

where $\delta = 1 - \frac{1}{q}$ *if* $d = 1$ *and* $\delta = \frac{1}{4}\left(1 - \frac{pd+d-1}{pqd}\right)$ *if* $d \geq 2$*.*

Lemma 5. *Assume that* $\lim_{\varepsilon\to 0} h = \lim_{\varepsilon\to 0} \frac{\varepsilon}{h^2} = 0$. *Using notation (11), we have*

$$\mathbb{E}\|\beta_\varepsilon\|_p^p = \mathbb{E}|Z|^p \int_0^1 (\Sigma(t))^p \, \mathrm{d}t + o\left(\frac{1}{\sqrt{h}}\right), \qquad \text{as } \varepsilon \to 0.$$

If we choose the optimal window $h = \lambda\varepsilon^{\frac{1}{5}}$ *and considering* $\|\hat{\beta}_\varepsilon\|_p^p$ *we have*

$$\mathbb{E}\|\hat{\beta}_\varepsilon\|_p^p = \int_0^1 |\Sigma(t)Z + c(t)|^p \, \mathrm{d}t + o\left(\frac{1}{\sqrt{h}}\right), \qquad \text{as } \varepsilon \to 0.$$

In order to provide the asymptotic variance of $D_{p,\varepsilon}$ we precise the second order properties of the random process $(\beta_\varepsilon(t))_{t\in[0,1]}$. We set

$$\widetilde{\beta}_\varepsilon(t) = \frac{\beta_\varepsilon(t)}{\sqrt{\operatorname{Var}\beta_\varepsilon(t)}}. \tag{17}$$

To obtain the asymptotic behaviour of $\operatorname{Var} D_{p,\varepsilon}$, we shall need the asymptotic behaviour of $\operatorname{Cov}(\widetilde{\beta}_\varepsilon(s), \widetilde{\beta}_\varepsilon(t))$, easily deduced from the following lemma.

Lemma 6. *Assume that* $\lim_{\varepsilon\to 0} h = \lim_{\varepsilon\to 0} \frac{\varepsilon}{h} = 0$, *then*

$$\begin{aligned}\operatorname{Cov}(\beta_\varepsilon(s), \beta_\varepsilon(t)) \sim & \int_{-\frac{1}{2}}^{\frac{1}{2}} K(u) K\left(u + \frac{t-s}{h}\right) \mathrm{d}u \\ & \times \int_{-1}^{1} \sum_{n=1}^{\infty} a_{2n}(s) a_{2n}(t) (2n)! \left(\frac{\sigma(t)}{\sigma(s)\|\varphi\|_2^2} \int_{-1}^{1} \varphi(x)\varphi(x+z)\, \mathrm{d}x\right)^{2n} \mathrm{d}z.\end{aligned}$$

Mehler's formula allows computing moments of non linear functionals of a Gaussian process. Hence if process β_ε was Gaussian then we should be able to derive the asymptotic behaviour of $D_{p,\varepsilon}$, but this is not the case. Using a Gaussian approximation of β_ε, the following lemma indicates what the asymptotic behaviour of $\operatorname{Var} D_{p,\varepsilon}$ would be. Thus, we consider the centered Gaussian process $(B_\varepsilon(t))_{t\in[0,1]}$, such as

$$\operatorname{Cov}(B_\varepsilon(s), B_\varepsilon(t)) = \operatorname{Cov}(\beta_\varepsilon(s), \beta_\varepsilon(t)), \qquad \forall s, t \in [0,1].$$

Lemma 7. *Using notations (11)–(12), we assume that* $\lim_{\varepsilon\to 0} h = \lim_{\varepsilon\to 0} \frac{\varepsilon}{h} = 0$.
Let $b_{2k} = \frac{1}{2k!}\mathbb{E}|Z|^p H_{2k}(Z)$, *then*

$$\operatorname{Var}\|B_\varepsilon\|_p^p \sim h \sum_{k=1}^{\infty} b_{2k}\,(2k)! \int \Gamma^{2k}(w)\, \mathrm{d}w \cdot \int \Sigma^{2p}(t)\, \mathrm{d}t.$$

Remark. Let (Z_1, Z_2) be a standard $(0, I_2)$ normal vector, then the previous expression can be written as

$$\operatorname{Var}\|B_\varepsilon\|_p^p \sim h \int \operatorname{Cov}\left(\left|\sqrt{1-\Gamma^2(w)}Z_1 + \Gamma(w)Z_2\right|^p, |Z_2|^p\right) \mathrm{d}w \cdot \int \Sigma^{2p}(t)\, \mathrm{d}t.$$

4. Proofs of the theorems

4.1. Proof of Theorem 1: case $b \equiv 0$

As quoted in Lemma 3, $D_{p,\varepsilon}$ is a sum of the 2-dependent random variables $(Y_{k,\varepsilon} - \mathbb{E}Y_{k,\varepsilon})_{1\le k\le M}$ with $M = M_\varepsilon = \left[\frac{1}{2h}\right]$.

Now let $s,t\in[0,1]$ be such as $|s-t|\le 2\varepsilon$, then it is simple to deduce from Lemma 2 that the random variable $(\beta_\varepsilon(s),\beta_\varepsilon(t))\in\mathbb{R}^2$ can also be written as the sum of 4-dependent vectors, $x_1+\cdots+x_N$. In case $d=2$ Lemma 4 implies

$$|\mathbb{E}|\beta_\varepsilon(s)|^p|\beta_\varepsilon(t)|^p-\mathbb{E}|B_\varepsilon(s)|^p|B_\varepsilon(t)|^p|\le\left(\frac{\varepsilon}{h}\right)^{\frac{\delta}{2}}.$$

Again applying Lemma 4 with $d=1$ allows us to subtract expectations which finally yields

$$|\mathrm{Cov}\,(|\beta_\varepsilon(s)|^p,|\beta_\varepsilon(t)|^p)-\mathrm{Cov}\,(|B_\varepsilon(s)|^p,|B_\varepsilon(t)|^p)|\le\left(\frac{\varepsilon}{h}\right)^{\frac{\delta}{2}}+o(1),$$

where δ is provided in Lemma 4, which is 1 for $d=2$. In order to compute an approximation of Var $D_{p,\varepsilon}$ we first expand

$$\mathrm{Var}\ D_{p,\varepsilon}=\frac{1}{h}\int\int\mathrm{Cov}\,(|\beta_\varepsilon(s)|^p,|\beta_\varepsilon(t)|^p)\,\mathrm{d}s\,\mathrm{d}t,$$

where $s,t\in[0,1]$ and using Lemma 3, we check that this is enough to assume $|s-t|\le 2\varepsilon$. We get the bound

$$\left|\mathrm{Var}\ D_{p,\varepsilon}-\Sigma_p^2\right|\le\sqrt{\frac{\varepsilon}{h}}\cdot o(1),$$

which is enough for our purpose.

If $q>1+\frac{1}{2p}$, then Lemma 7 yields

$$\mathrm{Var}\,(D_{p,\varepsilon})\longrightarrow_{\varepsilon\to 0}\Sigma_p^2.$$

The CLT will follow from the Lindeberg condition

$$\eta_\varepsilon=\sum_{k=1}^{M_\varepsilon}\mathbb{E}\left|\frac{1}{\sqrt{h}}(Y_{k,\varepsilon}-\mathbb{E}Y_{k,\varepsilon})\right|^4\longrightarrow_{\varepsilon\to 0}0.$$

Using again Lemmas 4–6 we prove that if $q\ge 4$

$$\mathbb{E}\,|Y_{k,\varepsilon}-\mathbb{E}Y_{k,\varepsilon}|^4=\mathcal{O}(h^4),$$

because this is the expectation of a quadruple integral on a set with volume M_ε^{-4} and the integrated function has an expectation uniformly bounded by $2^4\sup_{t\in[0,1]}\mathbb{E}G^{4p}(\sigma(t)Z)$. This yields

$$\eta_\varepsilon=\mathcal{O}\left(h^{-3}\cdot h^4\right).$$

Remark. We set

$$D_{p,\varepsilon,t}=\frac{1}{\sqrt{h}}\int_0^t\left(|\beta_\varepsilon(s)|^p-\mathbb{E}|\beta_\varepsilon(s)|^p\right)\mathrm{d}s. \tag{18}$$

The previous proof provides a Donsker type invariance principle (for m-dependent sequences, again). Sketching the expression in Theorem 1, we set

$$\Sigma_p^2(t)=\int_{-1}^1\Theta(w)\,dw\cdot\int_0^t\Sigma^{2p}(s)\,\mathrm{d}s,$$

and $\widetilde{D}_{p,t}=\displaystyle\int_0^t\Sigma_p(s)\,\mathrm{d}\widetilde{W}_s$ for a standard Brownian motion $(\widetilde{W}_t)_{t\in[0,1]}$, then

$$D_{p,\varepsilon,t}\xrightarrow{\mathcal{D}}_{\varepsilon\to 0}\widetilde{D}_{p,t},\ \text{in the space } C([0,1]). \tag{19}$$

4.2. Proof of Theorem 1: the general case

Notations. To simplify the notation we add the drift parameter as an index in the underlying probability law which we now denote $\mathbb{P}^{(b)}$ and expectations $\mathbb{E}^{(b)}$. Hence the expression relative to the time changed Brownian motion (*i.e.* $b \equiv 0$) can be written as $E_\varepsilon^{(0)}(t)$, and $\mathbb{E}^{(0)}$, respectively.

An essential lemma links the expectations relative to $\mathbb{E}^{(b)}$ and $\mathbb{E}^{(0)}$. Under the conditions Theorem 1 p. 40 of [7]: (3) we have

Lemma 8 (Girsanov formula, *e.g.*, in [7] p. 90)**.** *Let $H : \mathbb{R} \to \mathbb{R}$ be continuous and bounded then:*

$$\mathbb{E}^{(b)} H(D_{p,\varepsilon}) = \mathbb{E}^{(0)} \left\{ H(D_{p,\varepsilon}) \exp\left(\int_0^1 \frac{b(X_s)}{\sigma(s)} \, \mathrm{d}X_s - \frac{1}{2} \int_0^1 \frac{b^2(X_s)}{\sigma^2(s)} \, \mathrm{d}s \right) \right\}.$$

Remark. The conditions that must verify b are restrictive see (3), however if we deal only with weak solutions we can weaken our assumptions and the Girsanov formula will still holds.

An independence argument called stable convergence, that was developed in [10], entails the convergence in distribution of $D_{p,\varepsilon}$ under the general law $\mathbb{P}^{(b)}$ with the help of the Cameron-Martin formula (see [7] p. 82), which states that

$$\mathbb{E}^{(0)} \exp\left(\int_0^1 \frac{b(X_s)}{\sigma(s)} \, \mathrm{d}X_s - \frac{1}{2} \int_0^1 \frac{b^2(X_s)}{\sigma^2(s)} \, \mathrm{d}s \right) = 1.$$

We thus have to prove that the couple $\big((X_t)_{t\in[0,1]}, D_{p,\varepsilon}\big)$ converges in $C([0,1]) \times \mathbb{R}$ (under the distribution $\mathbb{P}^{(0)}$) to $\big((X)_{t\in[0,1]}, \Sigma_p Z\big)$ where the Brownian motion with a time change $(X)_{t\in[0,1]}$ is independent of the standard normal Z.

From now, we will only work under the probability distribution $\mathbb{P}^{(0)}$. Thus the previous asymptotic independence holds if the process

$$(E_{\varepsilon,t})_{t\in[0,1]} \equiv (X_t, D_{p,\varepsilon,t})_{t\in[0,1]}$$

(with values in $\mathbb{R}^2$) converges to a process $(E_t)_{t\in[0,1]} \equiv (X_t, D_{p,t})_{t\in[0,1]}$ as $\varepsilon \to 0$ such as $(X_t)_{t\in[0,1]}$ is independent of $D_{p,1}$ (we shall prove it for $(D_{p,t})_{t\in[0,1]}$).

As the family of distributions $(D_{p,\varepsilon,t})_{t\in[0,1]}$ converges under the probability distribution $\mathbb{P}^{(0)}$ as $\varepsilon \to 0$, this implies its tightness in $C([0,1])$ hence the process $(E_{\varepsilon,t})_{t\in[0,1]}$ is also tight in $C([0,1], \mathbb{R}^2)$. Let us consider now any limit point $(E_t)_{t\in[0,1]} \equiv (X_t, D_{p,t})_{t\in[0,1]}$ (in distribution) of this family, as $\varepsilon \to 0$. The random vector $(D_{p,\varepsilon,t} - D_{p,\varepsilon,s}, X_t - X_s)$ is independent (always under $\mathbb{P}^{(0)}$) from $D_{p,\varepsilon,t'} - D_{p,\varepsilon,s'}$ (if $s \le t \le s' \le t'$ satisfy $s' - t' > 2\varepsilon$) and it is also independent of $X_{t'} - X_{s'}$ because the intervals $[s,t]$ and $[s',t']$ do not overlap. This implies that the process $(E_t)_{t\in[0,1]}$ has independent increments. This process is also a second order process because each of its coordinates has this property. Hence $(E_t)_{t\in[0,1]}$ is a Gaussian process.

Independence of E_t coordinates now relies on their orthogonality. The only point we need to prove is thus that under the probability distribution $\mathbb{P}^{(0)}$, we have

$$\mathrm{Cov}\ (X_s, D_{p,\varepsilon,t}) \longrightarrow_{\varepsilon\to 0} 0, \qquad \forall s,t \in [0,1].$$

Note that

$$\frac{1}{\sqrt{h}} \mathrm{Cov}\ \left(X_s, \int_0^t |\beta_{\varepsilon,u}|^p \, \mathrm{d}u \right) = \frac{1}{\sqrt{h}} \int_0^t \mathrm{Cov}\ (X_s, |\beta_{\varepsilon,u}|^p) \, \mathrm{d}u. \tag{20}$$

In order to proceed we first write $A_\varepsilon(s,u) = \mathrm{Cov}\ (X_s, |\beta_{\varepsilon,u}|^p) = 0$ if $u > s + \varepsilon$. Now we deduce that $\frac{1}{\sqrt{h}} \int_{s-2\varepsilon}^{s+\varepsilon} A_\varepsilon(s,u) \, \mathrm{d}u = O\left(\frac{\varepsilon}{\sqrt{h}} \right)$ from the relation $\sup_{t\in[0,1]} \mathbb{E}|\beta_\varepsilon(t)|^{2p} < \infty$.

We thus only need to consider $\frac{1}{\sqrt{h}}\int_0^{s-2\varepsilon} A_\varepsilon(s,u)\,\mathrm{d}u$. Recall that

$$\beta_{\varepsilon,u} = \sqrt{\frac{h}{\varepsilon}}(\hat{\alpha}_\varepsilon(u) - E(\hat{\alpha}_\varepsilon(u))),$$

where

$$\hat{\alpha}_\varepsilon(u) = \int_{-1/2}^{1/2} K(u) G\left(\frac{\sqrt{\varepsilon}\dot{\sigma}_\varepsilon(u-vh)}{||\varphi||_2}\cdot Z_\varepsilon(u-vh)\right)\mathrm{d}v.$$

By making the change of variable $v = \frac{\varepsilon}{h}w$, we get

$$\frac{\varepsilon}{h}\int_{-1/2}^{1/2} K\left(\frac{\varepsilon}{h}w\right) G\left(\frac{\sqrt{\varepsilon}\dot{\sigma}_\varepsilon(u-\varepsilon w)}{||\varphi||_2}\cdot Z_\varepsilon(u-\varepsilon w)\right)\mathrm{d}w := \frac{\varepsilon}{h}\tilde{\alpha}_\varepsilon(u).$$

Thus, we can write

$$\beta_{\varepsilon,u} = \sqrt{\frac{\varepsilon}{h}}(\tilde{\alpha}_\varepsilon(u) - E(\tilde{\alpha}_\varepsilon(u))).$$

Conditioning w.r.t $X_s = x$ we obtain:

$$Z_\varepsilon(t-\varepsilon w) = \sqrt{\varepsilon}\nu_{t,w,\varepsilon}x + H_{t,w,\varepsilon},$$

for some uniformly bounded and deterministic $\nu_{u,w,\varepsilon}$ and some Gaussian $H_{u,w,\varepsilon}$, because $\mathrm{Cov}\,(X_s, Z(u-wh)) = O(\sqrt{\varepsilon})$.

We now define

$$\tilde{\tilde{\alpha}}(u) = \int_{-1/2}^{1/2} K\left(\frac{\varepsilon}{h}w\right) G\left(\frac{\sqrt{\varepsilon}\dot{\sigma}_\varepsilon(u-\varepsilon w)}{||\varphi||_2}\cdot H_{u,w,\varepsilon}\right)\mathrm{d}w,$$

and

$$\tilde{\beta}_{\varepsilon,u} = \sqrt{\frac{\varepsilon}{h}}\left(\tilde{\tilde{\alpha}}_\varepsilon(u) - E\left(\tilde{\tilde{\alpha}}_\varepsilon(u)\right)\right).$$

We are interested in obtaining the asymptotic behavior of

$$\frac{1}{\sqrt{h}}\left|\mathrm{Cov}\;(X_s, |\beta_{\varepsilon,u}|^p) - \mathrm{Cov}\;\left(X_s, \left|\tilde{\beta}_{\varepsilon,u}\right|^p\right)\right|,$$

for $t < s - 2\epsilon$. Note that the second term in the above difference is 0. Moreover

$$\frac{1}{\sqrt{h}}\left|\mathrm{Cov}\;(X_s, |\beta_{\varepsilon,u}|^p) - \mathrm{Cov}\;\left(X_s, \left|\tilde{\beta}_{\varepsilon,u}\right|^p\right)\right| = \frac{1}{\sqrt{h}}\left|E\left[X_s\left(|\beta_{\varepsilon,u}|^p - \left|\tilde{\beta}_{\varepsilon,u}\right|^p\right)\right]\right|. \tag{21}$$

The inequality $||x+y|^p - |x|^p| \le p|y|(|x|^{p-1} + |y|^{p-1})$, entails

$$\begin{aligned}(21) &\le \frac{p}{\sqrt{h}}E|X_s|\left[\left|\beta_{\varepsilon,u} - \tilde{\beta}_{\varepsilon,u}\right|\left(|\beta_{\varepsilon,u}|^{p-1} + \left|\beta_{\varepsilon,u} - \tilde{\beta}_{\varepsilon,u}\right|^{p-1}\right)\right]\\ &\le \frac{p}{\sqrt{h}}\left(E|X_s|^2\left|\beta_{\varepsilon,u} - \tilde{\beta}_{\varepsilon,u}\right|^2\right)^{1/2} O(1).\end{aligned}$$

Hence we get

$$\frac{p}{\sqrt{h}}\left(E|X_s|^2\left|\beta_{\varepsilon,u} - \tilde{\beta}_{\varepsilon,u}\right|^2\right)^{1/2} \le \frac{\sqrt{2}p}{\sqrt{h}}\left(\frac{\varepsilon}{h}\right)^{p/2}\left(E\left[|X_s|^2\left|\tilde{\alpha}_\varepsilon(u) - \tilde{\tilde{\alpha}}_\varepsilon(u)\right|^2\right] + \left(E|X_s|^2 E\left|\tilde{\alpha}_\varepsilon(u) - \tilde{\tilde{\alpha}}_\varepsilon(u)\right|^2\right)^{1/2}\right).$$

But we have

$$\tilde{\alpha}_\varepsilon(u) - \tilde{\tilde{\alpha}}_\varepsilon(u) = \sqrt{\varepsilon}\int_{-1/2}^{1/2} K\left(\frac{\varepsilon}{h}v\right)\nu_{u,v,\varepsilon}X_s G'\left(\lambda_1\sqrt{\varepsilon}\nu_{u,v,\varepsilon}X_s + \lambda_2 H_{u,v,\varepsilon}\right)\mathrm{d}v.$$

This yields

$$E\left[|X_s|^2\left|\tilde{\alpha}_\varepsilon(u) - \tilde{\tilde{\alpha}}_\varepsilon(u)\right|^2\right] \le \varepsilon O(1),$$

by using the assumption about $G'(x)$. Therefore

$$\frac{1}{\sqrt{h}}\left|\text{Cov }(X_s, |\beta_{\varepsilon,u}|^p) - \text{Cov }\left(X_s, \left|\tilde{\beta}_{\varepsilon,u}\right|^p\right)\right| \le \left(\frac{\varepsilon}{h}\right)^{\frac{p+1}{2}} O(1).$$

This last term tends to 0 when $\lim\limits_{\varepsilon\to 0}\dfrac{\varepsilon}{h} = 0$. All this implies

$$\frac{1}{\sqrt{h}}\int_0^{s-2\varepsilon} A_\varepsilon(s,u)\,\mathrm{d}u \to 0.$$

4.3. **Proof of Proposition 1**

We write

$$b_\varepsilon(t) = \int K(s)\mathbb{E}G\left(\frac{\sqrt{\varepsilon}\dot{\sigma}(t-hs)}{\|\varphi\|_2}Z\right)\mathrm{d}s - \alpha(t).$$

Using Lemma 1, setting $\theta = \sigma^2$, we obtain the following uniform estimates

$$\left(\frac{\sqrt{\varepsilon}\dot{\sigma}(t-hs)}{\|\varphi\|_2}\right)^2 = \theta(t) - sh\theta'(t) + \frac{1}{2}s^2h^2\theta''(t) + o\big(h^2\big).$$

Consider the function $g(x) = G(\sqrt{|x|})$, thus g is also *a.s.* twice differentiable and

$$b_\varepsilon(t) = \int K(s)\mathbb{E}g\left(\left(\frac{\sqrt{\varepsilon}\dot{\sigma}(t-hs)}{\|\varphi\|_2}\right)^2 Z^2\right)\mathrm{d}s - \mathbb{E}g\big(\theta(t)Z^2\big).$$

Use of Taylor formula yields

$$b_\varepsilon(t) = \mathbb{E}\int K(s)\left(\left(-sh\theta'(t) + \frac{1}{2}s^2h^2\theta''(t)\right)Z^2 g'\left(\theta(t)Z^2\right) + \frac{1}{2}s^2h^2\theta'^2(t)Z^4 g''\left(\theta(t)Z^2\right)\right)\mathrm{d}s + o(h^2).$$

Using symmetries yields with the relation $g(u) = G(u^2)$,

$$b_\varepsilon(t) = \frac{h^2}{2}\int s^2K(s)\,\mathrm{d}s\cdot\mathbb{E}\left(\sigma''(t)ZG'(\sigma(t)Z) + \sigma'^2(t)Z^2G''(\sigma(t)Z)\right) + o\big(h^2\big).$$

The remark concerning the C^3 case follows from careful statements of the above relation with the bound $\varepsilon^2 = o(h^3)$.

4.4. **Proof of Theorem 2**

As done previously, we make use of the stable convergence argument in order to deal only with the simpler case $b \equiv 0$. **We assume below that** $b \equiv 0$**.**

We write $\widehat{\beta}_\varepsilon(t) = \beta_\varepsilon(t) + c_\varepsilon(t)$, with

$$c_\varepsilon(t) = \sqrt{\frac{h}{\varepsilon}} b_\varepsilon(t) = \sqrt{\frac{h^5}{\varepsilon}} \left(a(t) \int s^2 K(s)\,\mathrm{d}s + o(1) \right). \tag{22}$$

Thus

$$\mathcal{D}_{p,\varepsilon} = \frac{1}{\sqrt{h}} \int_0^1 \left(|\beta_\varepsilon(t) + c_\varepsilon(t)|^p - \mathbb{E}\,|\beta_\varepsilon(t) + c_\varepsilon(t)|^p \right) \mathrm{d}t.$$

Proof of relation (14). Using the bound

$$\Big| |\beta + c|^p - |\beta|^p \Big| \le p|c| \left(|\beta|^{p-1} + |c|^{p-1} \right),$$

we write the following integral (still with $|s-t| \le 2\varepsilon$)

$$\mathrm{Var}\,(\mathcal{D}_{p,\varepsilon} - D_{p,\varepsilon}) \le \frac{1}{h} \int\!\!\int \mathbb{E}\Big| (|\widehat{\beta}_\varepsilon(s)|^p - |\beta_\varepsilon(s)|^p)(|\widehat{\beta}_\varepsilon(t)|^p - |\beta_\varepsilon(t)|^p) \Big|\,\mathrm{d}s\,\mathrm{d}t.$$

Assuming $\lim_{\varepsilon\to 0} h^5/\varepsilon = 0$ we obtain $\lim_{\varepsilon\to 0} \mathrm{Var}\,(\mathcal{D}_{p,\varepsilon} - D_{p,\varepsilon}) = 0$. The following facts: $\sup_{s\in[0,1]} \|\beta_\varepsilon(s)\|_{2p-1} \le \sup_{s\in[0,1]} \|\beta_\varepsilon(s)\|_{2p} < \infty$ and $\varepsilon/h \to 0$, imply (14). □

Proof of Theorem 2. Replacing β by $\beta + c$ and c by $c_\varepsilon - c$, we prove as above that

$$\lim_{\varepsilon\to 0} \mathrm{Var}\, \left(\mathcal{D}_{p,\varepsilon} - \widehat{\mathcal{D}}_{p,\varepsilon} \right) = 0,$$

where

$$\widehat{\mathcal{D}}_{p,\varepsilon} = \frac{1}{\sqrt{h}} \int_0^1 \left(|\beta_\varepsilon(t) + c(t)|^p - \mathbb{E}|\beta_\varepsilon(t) + c(t)|^p \right)\,\mathrm{d}t.$$

The proof of Theorem 2 follows the same lines as that of Theorem 1 up to very simple changes in Lemma 4. This lemma was indeed dedicated to the approximation of $\mathbb{E}f(x_1 + \cdots + x_n)$ for m-dependent vector valued sequences and for the special function $f(x_1, \ldots, x_d) = \prod_{\ell=1}^d |x_\ell|^p$. Very small changes entail the same result with $f(x_1, \ldots, x_d) = \prod_{\ell=1}^d |x_\ell + c_\ell|^p$ for fixed real numbers $c_1, \ldots, c_d$. Indeed, one may easily rewrite a version of this lemma for which the measurable function f only satisfies $|f(x_1, \ldots, x_d)| \le \prod_{\ell=1}^d |x_\ell|^p \vee 1$. □

5. Proofs of the lemmas in Section 3

5.1. **Proof of Lemma 4**

The proofs are different for $d = 1$ and $d \ge 2$.

Case $d = 1$. . Shergin [16] (Th. 1) proved that

$$\Delta_n = \sup_{x\in\mathbb{R}} \left| \mathbb{P}\left(\sum_{k=1}^n x_{k,n} \le x \right) - \mathbb{P}(Z \le x) \right| \le c \sum_{k=1}^n \mathbb{E}\|x_{k,n}\|^3.$$

Recall that the following relation holds for each random variable in L^p

$$\mathbb{E}|X|^p = p \int_0^\infty x^{p-1} \mathbb{P}(|X| > x)\,\mathrm{d}x,$$

hence the difference of expectations to approximate is an integral over $\mathbb{R} = (-\infty, \infty)$. Divide it for $|x| \le M \equiv \Delta_n^{-\frac{1}{pq}}$ and $|x| > M$. Rosenthal inequality [15] up to order pq (this also holds with m-dependent sequences since sums may be rewritten as m sum of independent variables) and Markov inequality provide a bound for the the second term while the first one is bounded by using the previous result in [16].

Case $d \ge 2$. In order to handle the same technique as above, we need to develop a bound analogue as that in [16].

Lemma 9. *Assuming the assumptions in Lemma 4, then*

$$\Delta_n = \sup_{x \in \mathbb{R}^+} \left| \mathbb{P}\left(\left\| \sum_{k=1}^n x_{k,n} \right\| \le x \right) - \mathbb{P}(\|\mathbf{Z}\| \le x) \right| \le c \left(\sum_{k=1}^n \mathbb{E}\|x_{k,n}\|^3 \right)^{\frac{1}{4}}.$$

Notation. For simplicity's sake, set

$$\Delta_n(x) = \left| \mathbb{P}\left(\left\| \sum_{k=1}^n x_{k,n} \right\| \le x \right) - \mathbb{P}(\|\mathbf{Z}\| \le x) \right|.$$

The proof of Lemma 4 now follows the same lines as for $d = 1$ up to the following expressions

$$\mathbb{E}|X_1 \cdots X_d|^p = p^d \int_0^\infty \cdots \int_0^\infty |x_1 \cdots x_d|^{p-1} \left(1 - \mathbb{P}(X_1 \le x_1, \ldots, X_d \le x_d)\right) \mathrm{d}x_1 \cdots \mathrm{d}x_d.$$

For example using $\|(x_1, \ldots, x_d)\| = \max\{|x_1|, \ldots, |x_d|\}$ implies that the difference of product moments to bound is bounded above by

$$c_p \int_0^\infty x^{pd+d-1} \Delta_n(x)\, \mathrm{d}x,$$

for a constant $c_p > 0$ only depending on p. Using the same techniques as above, yields the result, by truncating at a level $M > 0$ such as

$$M^{-4pqd} = \sum_{k=1}^n \mathbb{E}\|x_{k,n}\|^3.$$

5.2. Proof of Lemma 9

The proof will use the following lemma which is an easy extension of [14] to a vector valued case.

Lemma 10 (Lindeberg-Rio for m-dependent sequences)**.** *Let $d \in \mathbb{N}$. Let $x_1, \ldots, x_n \in \mathbb{R}^d$ be centered at expectation, m-dependent and such as $\mathbb{E}\|x_k\|^3 < \infty$ for $k = 1, \ldots, n$. Then there exists an independent succession $y_1, \ldots, y_n$ of centered d-dimensional random vectors with the following property. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a C^3-function with bounded partial derivatives of order 3 (write $\|f'''\|_\infty = \sup_{\{s, \|h_i\| \le 1; i=1,2,3\}} \|f^{'''}(s)(h_1, h_2, h_3)\|$), then if we consider*

$$\Delta_n(f) = \mathbb{E}(f(x_1 + \cdots + x_n) - f(y_1 + \cdots + y_n)),$$

there exists a constant $c > 0$ such as

$$|\Delta_n(f)| \le c\|f'''\|_\infty \sum_{k=1}^n \mathbb{E}\|x_k\|^3.$$

Remarks. In view of the theorem relative to the equivalence of the norms in the d-dimensional space we may choose any norm on $\mathbb{R}^d$ and the constant c only depends on this norm and on m.

A simple use of Taylor formula at the origin and with order 3 proves that expression $\Delta_n(f)$ is well defined.

With Lemma 10 we consider a C^3-function $g_{\delta,u}$ such as $g_{\delta,u}(x) \in [0,1]$ for each $x \in \mathbb{R}^d$ and $g_{\delta,u}(x) = 1$ if $\|x\| \le u$, $g_{\delta,u}(x) = 0$ if $\|x\| \ge u + \delta$. It is possible to construct such functions satisfying $\|g'''_{\delta,u}\|_\infty \le c\delta^{-3}$. Now let $\delta^4 = \sum_{k=1}^n \mathbb{E}\|x_k\|^3$, then the result follows in a standard way (see, *e.g.*, [5]).

5.3. **Proof of Lemma 10**

Notations. The second derivative of f at point s is a (symmetric) bilinear form on $\mathbb{R}^d$. It will also be considered as a (symmetric) $d \times d$ matrix and we shall denote

$$f''(s) \bullet v = \sum_{i,j=1}^d \frac{\partial^2}{\partial x_i \partial x_j} f(s) \cdot v_{i,j} \text{ if } v = (v_{i,j})_{1 \le i,j \le d}.$$

For simplicity we shall handle only the case $m = 1$. We construct independent Gaussian random variables $y_1, \ldots, y_n$ independent of $(x_1, \ldots, x_n)$, such as $y_k \sim \mathcal{N}(0, v_k)$ with $v_k = \mathbb{E}x_k^t x_k + \mathbb{E}x_{k-1}^t x_k + \mathbb{E}x_k^t x_{k-1}$ for $k = 1, \ldots, n$, where we set $x_0 = 0$.

Remark. In order to complete the proof in the general m-dependent case we should define

$$v_k = \mathbb{E}x_k^t x_k + \sum_{\ell=1}^m \left(\mathbb{E}x_{k-\ell}^t x_k + \mathbb{E}x_k^t x_{k-\ell} \right),$$

for $k = 1, \ldots, n$, where $x_0 = \cdots = x_{1-m} = 0$.

Set $s_k = x_1 + \cdots + x_k$, $t_k = y_{k+1} + \cdots + y_n$ if $k = 0, \ldots, n$, with $s_0 = 0$, $t_n = 0$.

As in [14] (Def. 3) we decompose

$$\begin{aligned}
\Delta_n(f) &= \mathbb{E}(f(s_n) - f(t_0)) = \sum_{k=1}^n \left(\Delta_{1,k}(f_k) - \Delta_{2,k}(f_k) \right), \quad \text{with} \\
\Delta_{1,k}(g) &= \mathbb{E}g(s_k) - g(s_{k-1}) - \frac{1}{2} g''(s_{k-1}) \bullet v_k, \text{ and} \\
\Delta_{2,k}(g) &= \mathbb{E}g(s_{k-1} + y_k) - g(s_{k-1}) - \frac{1}{2} g''(s_{k-1}) \bullet v_k, \text{ where} \\
f_k(x) &= \mathbb{E}f(x + t_k), \text{ hence } \|f_k'''\|_\infty \le \|f'''\|_\infty.
\end{aligned}$$

In the above display $g : \mathbb{R}^d \to \mathbb{R}$ denotes any C^3-function with third order bounded partial derivatives. The bound

$$|\Delta_{2,k}(g)| \le c\|g'''\|_\infty \left(\mathbb{E}\|x_k\|^3 + \|x_{k-1}\|^3 \right), \tag{23}$$

follows from Taylor formula

$$\|g(s+y) - g(s) - g'(s)(y) - \frac{1}{2} g''(s)(y,y)\| \le \frac{1}{6} \|g'''\|_\infty \mathbb{E}\|y\|^3,$$

applied with $s = s_{k-1}$ and $y = y_k$ and the independence properties of $y_1, \ldots, y_n$, for a suitable constant c. In order to let himself be persuaded, the reader may restate the formula

$$\mathbb{E}g''(s_{k-1})(y_k, y_k) = \mathbb{E}g''(s_{k-1}) \bullet v_k.$$

The terms $\Delta_{1,k}(g)$ are more delicate to expand. Again using the previous Taylor expansion (now $y = x_k$) we see that, up to a term bounded as in equation (23), we only need to consider the expectation of

$$g'(s_{k-1})(x_k) + \frac{1}{2} g''(s_{k-1})(x_k, x_k) - \frac{1}{2} g''(s_{k-1}) \bullet v_k = \delta_1 + \delta_2,$$

with

$$\begin{aligned}\delta_1 &= g'(s_{k-1})(x_k) - \frac{1}{2}g''(s_{k-1}) \bullet \left(\mathbb{E}x_{k-1}^t x_k + \mathbb{E}x_k^t x_{k-1}\right),\\ \delta_2 &= g''(s_{k-1})(x_k, x_k) - \frac{1}{2}g''(s_{k-1}) \bullet \mathbb{E}x_k^t x_k.\end{aligned}$$

Rewriting

$$\delta_2 = \frac{1}{2}(g''(s_{k-1}) - g''(s_{k-2}))(x_k, x_k) + \frac{1}{2}g''(s_{k-2})(x_k, x_k) - \frac{1}{2}g''(s_{k-1}) \bullet \mathbb{E}x_k^t x_k.$$

As before using the independence of x_k and s_{k-2}, a first order Taylor expansion yields

$$\begin{aligned}\mathbb{E}\delta_2 &= \frac{1}{2}\left(\mathbb{E}g''(s_{k-2})(x_k, x_k) - g''(s_{k-1}) \bullet \mathbb{E}x_k^t x_k\right)\\ &= \frac{1}{2}\left(\mathbb{E}g''(s_{k-2}) - g''(s_{k-1})\right) \bullet \mathbb{E}x_k^t x_k.\end{aligned}$$

The mean value theorem provides now the expected bound for $\mathbb{E}\delta_2$, analogous to that in equation (23).

The other term considered can be written

$$\delta_1 = g''(s_{k-2})(x_{k-1}, x_k) - \frac{1}{2}g''(s_{k-1}) \bullet \left(\mathbb{E}x_{k-1}^t x_k + \mathbb{E}x_k^t x_{k-1}\right) + R$$

where $|R| \le c\|g'''\|\mathbb{E}\|x_k\|\|x_{k-1}\|^2$ is bounded as in the above equation (23). Hence, using again the mean value theorem, we obtain Lemma 3.

5.4. **Proof of Lemma 5**

We set $d = 1$. Fix $t \in [0, 1]$. We apply the approximation Lemma 4 to the random variables $x_k = \zeta_{k,\varepsilon}(t)$ for $1 \le k \le N$; for simplicity we also define $x_0 = x_{N+1} = 0$. Then setting $y_1, \ldots, y_N$, a sequence of independent and centered random variables such as, $\mathbb{E}y_k^2 = \mathbb{E}x_{k-1}x_k + \mathbb{E}x_k^2 + \mathbb{E}x_k x_{k+1}$ we deduce, using Lemma 4 with $d = 1$ and $q = 2$, that for a suitable constant $c > 0$,

$$\left|\mathbb{E}|\beta_\varepsilon(t)|^p - \mathbb{E}\,|B_\varepsilon(t)|^p\,\right| \le c\left(\frac{\varepsilon}{h}\right)^{\frac{1}{2}}. \tag{24}$$

And since $\frac{\varepsilon}{h^2} \to 0$, we also deduce from (11) that

$$\frac{1}{\sqrt{h}}\left(\mathbb{E}\|\beta_\varepsilon\|_p^p - \mathbb{E}|Z|^p \int_0^1 (\Sigma(t))^p \,\mathrm{d}t\right) \to_{\varepsilon\to 0} 0.$$

To proof the second statement of this lemma recall that $\hat{\beta}_\varepsilon(t) = \beta_\varepsilon(t) + c_\varepsilon$ and let us consider that we use the optimal window *i.e.* $h = \lambda\varepsilon^{\frac{1}{5}}$. By using again Lemma 4 with $d = 1$ and $q = 2$, we obtain

$$|\mathbb{E}|\beta_\varepsilon(t) + c_\varepsilon(t)|^p - \mathbb{E}\,|B_\varepsilon(t) + c_\varepsilon(t)|^p| \le c\left(\frac{\varepsilon}{h}\right)^{\frac{1}{2}}. \tag{25}$$

In order to obtain the result this computation leads to consider the following difference

$$\frac{1}{\sqrt{h}}\left|\int_0^1 \left(\mathbb{E}\,|B_\varepsilon(t) + c_\varepsilon(t)|^p - \mathbb{E}\,|\Sigma(t)Z + c(t)|^p\right) \mathrm{d}t\right|.$$

By using the inequality

$$\left||\beta + c|^p - |\beta|^p\right| \le p|c|\left(|\beta|^{p-1} + |c|^{p-1}\right),$$

it is enough to show that

$$|\mathbb{E}\,|B_\varepsilon(t) + c_\varepsilon(t)| - \mathbb{E}\,|\Sigma(t)Z + c(t)|\,| = o\left(\frac{1}{\sqrt{h}}\right),$$

which is a consequence of the definition of $\mathbb{E}(\beta_\varepsilon(t))^2$ and a Taylor development of order two.

5.5. **Proof of Lemma 6**

Using Lemma 1 with Mehler's formula yields

$$\mathrm{Cov}\;(\beta_\varepsilon(s), \beta_\varepsilon(t)) \sim \frac{h}{\varepsilon}\int\!\!\int K(u)K(v)\sum_{n=1}^{\infty} a_{2n}(s)a_{2n}(t)(2n)!\,A_\varepsilon^{2n}(s,t,u,v)\,\mathrm{d}u\,\mathrm{d}v,$$

with

$$A_\varepsilon(s,t,u,v) = \frac{1}{\Sigma(s-uh)\Sigma(t-vh)}\int \varphi(x)\varphi\left(x + \frac{t-s}{\varepsilon} + h\frac{u-v}{\varepsilon}\right)\sigma^2(t-vh)\,\mathrm{d}x.$$

The change of variable $v \mapsto z = \frac{t-s}{\varepsilon} + h\frac{u-v}{\varepsilon}$ yields the result using Lebesgue dominated theorem (the corresponding integrals are uniformly convergent).

5.6. **Proof of Lemma 7**

Using again Mehler's formula, we get

$$\mathrm{Var}\;\|B_\varepsilon\|_p^p = \int\!\!\int \sum_{k=1}^{\infty} b_{2k}\,(2k)!\,\mathrm{Cov}^{2k}(B_\varepsilon(s), B_\varepsilon(t))\Sigma^p(t)\Sigma^p(s)\,\mathrm{d}s\,\mathrm{d}t.$$

Now the change of variable $s \mapsto w = \frac{t-s}{h}$ and a systematic use of Lebesgue convergence theorem yield the result.

References

[1] J. Beirlant and D.M. Mason, On the asymptotic normality of the L^p-norm of empirical functional. *Math. Methods Statist.* **4** (1995) 1–19.

[2] C. Berzin-Joseph, J.R. León and J. Ortega, Non-linear functionals of the Brownian bridge and some applications. *Stoch. Proc. Appl.* **92** (2001) 11–30.

[3] P. Brugière, Théorème de limite centrale pour un estimateur non paramétrique de la variance d'un processus de diffusion multidimensionnelle. *Ann. Inst. Henri Poincaré, Probab. Stat.* **29** (1993) 357–389.

[4] P.D. Ditlevsen, S. Ditlevsen and K.K. Andersen, The fast climate fluctuations during the stadial and interstadial climate states. *Ann. Glaciology* **35** (2002).

[5] P. Doukhan, J.R. León and F. Portal, Calcul de la vitesse de convergence dans le théorème central limite vis-à-vis des distances de Prohorov, Dudley et Lévy dans le cas de v. a. dépendantes. *Probab. Math. Statist.* **6** (1985) 19–27.

[6] V. Genon-Catalot, C. Laredo and D. Picard, Non-parametric estimation of the diffusion coefficient by wavelets methods. *Scand. J. Statist.* **19** (1992) 317–335.

[7] I.J. Gihman and A.V. Skorohov, *Stochastic differential equations.* Springer-Verlag, Berlin, New York (1972).

[8] E. Giné, D. Mason and Yu. Zaitsev, *The L^1-norm density estimator process.* To appear in *Ann. Prob.*

[9] A. Gloter, Parameter estimation for a discrete sampling of an integrated Ornstein-Uhlenbeck process. *Statistics* **35** (2000) 225–243.

[10] J. Jacod, *On continuous conditional martingales and stable convergence in law*, sémin. Probab. XXXI, LNM 1655, Springer (1997) 232–246.

[11] P. Major, *Multiple Wiener-Itô integrals.* Springer-Verlag, New York, *Lect. Notes Math.* **849** (1981).
[12] G. Perera and M. Wschebor, Crossings and occupation measures for a class of semimartingales. *Ann. Probab.* **26** (1998) 253–266.
[13] G. Perera and M. Wschebor, Inference on the variance and smoothing of the paths of diffusions. *Ann. Inst. Henri Poincaré, Probab. Stat.* **38** (2002) 1009–1022.
[14] E. Rio, About the Lindeberg method for strongly mixing sequences. *ESAIM: PS* **1** (1995) 35–61.
[15] H.P. Rosenthal, On the subspaces of L^p, $(p > 2)$ spanned by sequences of independent random variables. *Israël Jour. Math.* **8** (1970) 273–303.
[16] V.V. Shergin, On the convergence rate in the central limit theorem for m-dependent random variables. *Theor. Proba. Appl.* **24** (1979) 782–796.
[17] P. Soulier, Non-parametric estimation of the diffusion coefficient of a diffusion process. *Stoch. Anal. Appl.* **16** (1998) 185–200.
[18] G. Terdik, *Bilinear Stochastic Models and Related problems of Nonlinear Time Series.* Springer-Verlag, New York, *Lect. Notes Statist.* **142** (1999).