Available online at www.sciencedirect.com

SciVerse ScienceDirect

Stochastic Processes and their Applications 123 (2013) 2500–2521

stochastic processes and their applications

www.elsevier.com/locate/spa

An infinite dimensional convolution theorem with applications to the efficient estimation of the integrated volatility

Emmanuelle Clément[a,*], Sylvain Delattre[b], Arnaud Gloter[c]

[a] *Université Paris-Est, Laboratoire d'Analyse et de Mathématiques Appliquées, UMR 8050, 5 Bld Descartes, Champs-sur-Marne, 77454 Marne-la-Vallée Cedex 2, France*

[b] *Laboratoire de Probabilités et Modèles Aléatoires, UMR 7599, Université Paris Diderot, 77251 Paris Cedex 05, France*

[c] *Université d'Évry Val d'Essonne, Département de Mathématiques, 91025 Évry Cedex, France*

Available online 4 April 2013

Abstract

This paper proposes a general approach to obtain asymptotic lower bounds for the estimation of random functionals. The main result is an abstract convolution theorem in a non parametric setting, based on an associated LAMN property. This result is then applied to the estimation of the integrated volatility, or related quantities, of a diffusion process, when the diffusion coefficient depends on an independent Brownian motion.

MSC: primary 62G20; secondary 60F05; 60H05

Keywords: LAMN property; Convolution theorem; Diffusion process

1. Introduction

A fundamental concept in the parametric estimation theory is the notion of *Locally Asymptotically Normal* (LAN) families of distributions, introduced by Le Cam (see Le Cam and Yang [17], Van der Vaart [21]). In particular, this notion permits to establish asymptotic lower bounds for the distribution of any 'regular' estimator of a parameter θ. More precisely, a classical result,

* Corresponding author.
E-mail address: Emmanuelle.Clement@univ-mlv.fr (E. Clément).

0304-4149/$ - see front matter
http://dx.doi.org/10.1016/j.spa.2013.04.004

known as the Hajek convolution theorem, states that the asymptotic distribution of any 'regular' estimator is necessarily a convolution between a Gaussian law and some other law. An advantage of this result is to give a natural way to introduce the notion of efficiency, in the case where the asymptotic distribution reduces to the Gaussian part just mentioned above. In a lot of situations, the LAN property is not satisfied but a more general condition, called *Locally Asymptotically Mixed Normality* (LAMN), can be established. In this latter case, the Hajek convolution theorem can be extended (see Jeganathan [15,16]) and the asymptotic distribution of any 'regular' estimator can be conditionally decomposed as a convolution.

In the LAN situation, some extensions have been done in a non parametric setting by Millar [18], Ibragimov and Kha'sminskii [7] and Golubev [5], but it seems that, up to now, similar results are still unknown in the LAMN case. The aim of this paper is to propose a Hajek type convolution theorem, for the estimation of a random functional, in a LAMN setting. For a random variable F with value in a space B, we consider the general estimation problem of $\Phi(F)$, based on the observation of a random variable with law P_n on a measurable space $(E_n, \mathcal{B}_n)$. The main assumption is that the probability P_n can be decomposed as $P_n(A) = \int_B P_n^f(A) dP^F$, where P^F is the law of the random variable F and $\{P_n^f, f \in B\}$ a statistical experiment, depending on an infinite dimensional parameter f and satisfying the LAMN property. In this Bayesian framework with prior P^F, we establish a convolution theorem which does not require any regularity assumption on the estimator, but requires some regularity on the prior P^F. This convolution theorem permits to define in a rigorous way the notion of asymptotic efficiency for the estimation of random functionals. Moreover, we give some extensions to the estimation of a quantity depending both on the observations and the prior P^F. Such situations occur frequently in practice.

In a second part, we apply our infinite dimensional convolution theorem to the estimation of some functionals of a diffusion process discretely observed. We assume that we observe at times $(t_i^n)_i$ the process X, the solution of

$$X(t) = x_0 + \int_0^t b(X(s))ds + \int_0^t a(X(s), \sigma(s))dW(s),$$

where $(\sigma(t))_t$ is an Itô process, independent of W. This problem can be connected to the preceding abstract framework since we observe a process depending on a random unknown infinite dimensional parameter σ. Our applications concern among others the estimation of quantities which appears in stochastic finance, such as the integrated volatility $\int_0^1 a^2(X(s), \sigma(s))ds$ or some stochastic integrals $\int_0^1 \chi(t, X(t))dX(t)$ related to hedging problems. From our convolution results, we derive explicit lower bounds for the estimation of these quantities based on a discrete sampling of X. Another application deals with the efficiency of discretization schemes such as the Euler scheme.

The paper is organized as follows. In Section 2, we derive an infinite dimensional convolution theorem based on the LAMN property. Section 3 is devoted to the applications to a discretely observed diffusion process. The technical proofs are postponed to the Appendix.

2. Infinite dimensional convolution theorem

2.1. Definitions and notations

Throughout this paper we will consider a real separable Hilbert space H, equipped with the inner product $\langle ., . \rangle$ and the associated norm $\|.\|$, and a subset $B \subset H$. Let I be a linear bounded positive self-adjoint operator on H, then I admits a square-root $I^{1/2}$, such that $I^{1/2} I^{1/2} = I$, which is also positive and self-adjoint. A Gaussian process on H, defined on a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, with covariance operator I, is a centered Gaussian family $N = \{N(h); h \in H\}$ such that for all h and k in H, $\mathbb{E}N(h)N(k) = \langle h, Ik\rangle$. We can observe that this implies that the map $h \mapsto N(h)$ is linear and continuous from H into $L^2(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$. A natural way to construct N, given the operator I and a sequence of independent standard Gaussian variables $(\xi_i)_{i\in\mathbb{N}}$, is to set $N(h) = \sum_i \xi_i \langle h, I^{1/2} e_i\rangle$, for $(e_i)_i$ a complete orthonormal system in H.

Now consider a family of probabilities $\{P_n^f, f \in H\}$ defined on a measurable Polish space $(E_n, \mathcal{E}_n)$. For f_1 and f_2 in H, we will denote by $dP_n^{f_1}/dP_n^{f_2}$ the derivative of the absolute continuous part of the probability $P_n^{f_1}$ with respect to the probability $P_n^{f_2}$, and we are interested in the asymptotic situation as n goes to infinity. We first define the Locally Asymptotically Mixed Normality property (LAMN property) in the direction H_0, for H_0 a linear subspace of H.

Definition 1. The family $\{P_n^f\}$ satisfies the LAMN property at $f \in B$, in the direction H_0, if there exists a sequence of random linear positive bounded operators $(I_n^f)_n$ on H, and random linear real valued functions N_n^f on H, such that $\forall h \in H_0$:

(i) $N_n^f(h)$ and $\langle h, I_n^f h\rangle$ are $\mathcal{E}_n$-measurable,

(ii) we have the decomposition

$$Z_n^f(h) := \log \frac{dP_n^{f+h/\sqrt{n}}}{dP_n^f} = N_n^f(h) - \frac{1}{2}\langle h, I_n^f h\rangle + o_{P_n^f}(1),$$

(iii) $\forall p \geq 1,\ \forall (h_1, \ldots, h_p) \in H_0^p$, we have the convergence in law under P_n^f

$$\begin{pmatrix} (N_n^f(h_i))_{1\leq i\leq p} \\ (\langle h_i, I_n^f h_j\rangle)_{1\leq i,j\leq p} \end{pmatrix} \underset{P_n^f}{\Longrightarrow} \begin{pmatrix} (N^f(h_i))_{1\leq i\leq p} \\ (\langle h_i, I^f h_j\rangle)_{1\leq i,j\leq p} \end{pmatrix}$$

where I^f is a random linear bounded positive self-adjoint operator on H, defined on a probability space $(\Omega, \mathcal{F}, P)$, and conditionally on I^f, N^f is a Gaussian process on H with covariance operator I^f.

In this definition, we could replace the rate of convergence $\sqrt{n}$ by any sequence (u_n) going to infinity.

2.2. *Convolution theorem*

We are interested in estimating the quantity $\Phi(F) = (\Phi^k(F))_{1\leq k\leq d}$, based on the observation of a random variable with law P_n on the measurable space $(E_n, \mathcal{E}_n)$, where $\Phi : H \mapsto \mathbb{R}^d$ is a known function and F a random variable with values in $B \subset H$. We note P^F the law of F. Since H is a separable Banach space, P^F is a Radon measure. This statistical context can be related to a Bayesian framework with prior P^F. We assume that the function Φ is $\mathcal{B}$-measurable, where $\mathcal{B}$ denotes the Borel sigma-field on H. In that follows, we still note $\mathcal{B}$ its trace sigma-field on B.

We make the following hypotheses on the probabilities P_n and P^F and on the function Φ.

H0. Regularity of P^F. For $h \in H_0$, we note $P^{F+h/\sqrt{n}}$ the law of $F + h/\sqrt{n}$ (that is for $A \in \mathcal{B}$, $P^{F+h/\sqrt{n}}(A) = P^F(A - h/\sqrt{n})$). We assume that $\forall h \in H_0$, $\lim_n \left\| P^{F+h/\sqrt{n}} - P^F \right\|_{Var} = 0$, where $\|.\|_{Var}$ denotes the total variation norm.

In the finite dimensional case ($H = \mathbb{R}^p$), H0 is satisfied if P^F admits a density with respect to the Lebesgue measure.

H1. Relation between P_n and P^F. We assume that there exists a family of probabilities $\{(P_n^f)_n; f \in H\}$ on $(E_n, \mathcal{E}_n)$ such that, for all n and for all $A \in \mathcal{E}_n$, the map $f \mapsto P_n^f(A)$ is measurable on $(H, \mathcal{B})$, and such that:

$$\forall A_n \in \mathcal{E}_n, \quad P_n(A_n) = \int_B P_n^f(A_n) dP^F(f).$$

We equip the measurable space $(B \times E_n, \mathcal{B} \otimes \mathcal{E}_n)$ with the probability $\mathbb{P}_n$ defined by

$$\forall A \in \mathcal{B}, \quad \forall A_n \in \mathcal{E}_n, \quad \mathbb{P}_n(A \times A_n) = \int_B P_n^f(A_n) 1_A(f) dP^F(f).$$

H2. LAMN property.

(a) We assume that for a linear subset H_0 such that $\overline{H_0} = H$, (P_n^f) satisfies the LAMN property for all $f \in B$, in the direction H_0 and that the space $(\Omega, \mathcal{F}, P)$, appearing in the LAMN definition, does not depend on f.

(b) We assume moreover that $\forall h, h_1 \in H_0$, $N_n^f(h)$ and $\langle h, I_n^f h_1 \rangle$ are measurable on $(B \times E_n, \mathcal{B} \otimes \mathcal{E}_n)$ and that $(\omega, f) \mapsto I^f(\omega)$ is measurable on $(\Omega \times B, \mathcal{F} \otimes \mathcal{B})$.

(c) We assume that $\forall f \in B$ and $\forall h, h_1, h_2 \in H_0$, $\langle h_1, I_n^{f+h/\sqrt{n}} h_2 \rangle - \langle h_1, I_n^f h_2 \rangle$ goes to zero in P_n^f-probability.

H3. Regularity of Φ.

(a) We assume that $\Phi : H \mapsto \mathbb{R}^d$ is Fréchet differentiable. For $1 \le k \le d$, we note $\dot{\Phi}^k(f)$ the unique vector in H such that $\forall h \in H, \quad \Phi^k(f+h) - \Phi^k(f) = \langle \dot{\Phi}^k(f), h \rangle + o(\|h\|)$. We will use the notation $\langle \dot{\Phi}(f), h \rangle = (\langle \dot{\Phi}^k(f), h \rangle)_{1 \le k \le d}$.

(b) We assume that, for $1 \le k \le d$, $P \otimes P^F(d\omega, df)$ almost surely, $\dot{\Phi}^k(f) \in (I^f(\omega))^{1/2}(H)$.

From the orthogonal decomposition $H = \overline{(I^f)^{1/2}(H)} \oplus \mathrm{K}er(I^f)^{1/2}$, we note $h_f = (h_f^k)_{1 \le k \le d}$ the unique vector in H^d such that for $1 \le k \le d$

$$(I^f)^{1/2} h_f^k = \dot{\Phi}^k(f) \text{ and } h_f^k \in \overline{(I^f)^{1/2}(H)}. \tag{1}$$

Before stating our main result, we recall that a sequence $(\widehat{\Phi}_n)_n$ is an estimator of $\Phi(F)$ if $\forall n$, $\widehat{\Phi}_n$ is $\mathcal{E}_n$-measurable.

In all that follows, we will denote the convergence in law under a probability P by '$\underset{P}{\Longrightarrow}$'.

Theorem 1. *Let $(\widehat{\Phi}_n)_n$ be any estimator of $\Phi(F)$, such that*

$$\sqrt{n}(\widehat{\Phi}_n - \Phi(F)) \underset{\mathbb{P}_n}{\Longrightarrow} Z. \tag{2}$$

Then assuming H0, H1, H2, H3*, the law of Z is a convolution:*

$$Z \underset{law}{=} \Sigma_F^{1/2} G + R, \quad \text{with } \Sigma_F = (\langle (I^F)^{-1/2} \dot{\Phi}^k(F), (I^F)^{-1/2} \dot{\Phi}^l(F) \rangle)_{1 \le k,l \le d}, \tag{3}$$

where conditionally on (F, I^F), R is a random variable independent of G, G is a standard Gaussian vector in $\mathbb{R}^d$, and $(I^F)^{-1/2} \dot{\Phi}^k(F) = h_F^k$ is defined by (1).

Remark 1. We will say that an estimator $\widehat{\Phi}_n$ satisfying (2) is efficient if

$$Z \underset{\text{law}}{=} \Sigma_F^{1/2} G.$$

This means that the dispersion of the conditional asymptotic distribution of $\widehat{\Phi}_n$ is minimal.

Remark 2. We can remark that the lack of regularity assumption for the estimator can be related to Jeganathan results [14], where some almost everywhere convolution theorems are established for a finite dimensional parameter (see Van der Vaart [21] p.115 in the LAN case).

Proof. Let $(h_i)_{i\geq 1}$ be a countable dense sequence in H such that $\forall i, h_i \in H_0$. Such a family exists since $\overline{H_0} = H$. For $p \geq 1$, we note V_p the linear subspace of H_0 generated by $(h_1, \ldots, h_p)$.

We first remark that from H1 and H2(a) (b), we have the convergence in law under $\mathbb{P}_n$

$$\begin{pmatrix} F \\ (N_n^F(h_i))_{1\leq i\leq p} \\ (\langle h_i, I_n^F h_j\rangle)_{1\leq i,j\leq p} \end{pmatrix} \underset{\mathbb{P}_n}{\Longrightarrow} \begin{pmatrix} F \\ (N^F(h_i))_{1\leq i\leq p} \\ (\langle h_i, I^F h_j\rangle)_{1\leq i,j\leq p} \end{pmatrix}. \tag{4}$$

Now since $\sqrt{n}(\widehat{\Phi}_n - \Phi(F))$ converges in law and P^F is a Radon measure, the vector $(\sqrt{n}(\widehat{\Phi}_n - \Phi(F)), F, (N_n^F(h_i))_{1\leq i\leq p}, (\langle h_i, I_n^F h_j\rangle)_{1\leq i,j\leq p})$ is tight and we deduce the convergence in law for a subsequence (n) (that we still note n)

$$\begin{pmatrix} \sqrt{n}(\widehat{\Phi}_n - \Phi(F)) \\ F \\ (N_n^F(h_i))_{1\leq i\leq p} \\ (\langle h_i, I_n^F h_j\rangle)_{1\leq i,j\leq p} \end{pmatrix} \underset{\mathbb{P}_n}{\Longrightarrow} \begin{pmatrix} Z \\ F \\ (N^F(h_i))_{1\leq i\leq p} \\ (\langle h_i, I^F h_j\rangle)_{1\leq i,j\leq p} \end{pmatrix}. \tag{5}$$

Our aim is to describe the law of Z, given (F, I^F). Remarking that the law of (F, I^F) is characterized by the distributions $(F, (\langle h_i, I^F h_j\rangle)_{1\leq i,j\leq p})$ with $p \geq 1$, it is sufficient to compute

$$\mathbb{E} e^{iu^*Z}\varphi(F)\psi((\langle h_i, I^F h_j\rangle)_{1\leq i,j\leq p})$$

for some continuous bounded functions $\varphi : B \mapsto \mathbb{R}$ and $\psi : \mathbb{R}^{p\times p} \mapsto \mathbb{R}$ and where u^* denotes the transpose of the vector $u = (u^k)_{1\leq k\leq d}$.

First we have immediately from (5) and using the notation $\psi_p(I^F) = \psi((\langle h_i, I^F h_j\rangle)_{1\leq i,j\leq p})$ and $\psi_p(I_n^F) = \psi((\langle h_i, I_n^F h_j\rangle)_{1\leq i,j\leq p})$

$$\mathbb{E} e^{iu^*Z}\varphi(F)\psi_p(I^F) = \lim_n \mathbb{E}_{\mathbb{P}_n} e^{iu^*\sqrt{n}(\widehat{\Phi}_n - \Phi(F))}\varphi(F)\psi_p(I_n^F). \tag{6}$$

On the other hand, using H1, we have

$$\mathbb{E}_{\mathbb{P}_n} e^{iu^*\sqrt{n}(\widehat{\Phi}_n - \Phi(F))}\varphi(F)\psi_p(I_n^F) = \int_B \left(\mathbb{E}_{P_n^f} e^{iu^*\sqrt{n}(\widehat{\Phi}_n - \Phi(f))}\varphi(f)\psi_p(I_n^f)\right) dP^F(f).$$

Now we fix $p_0 \leq p$ and $h \in V_{p_0}$ and we change P^F into $P^{F+h/\sqrt{n}}$. We deduce

$$\begin{aligned} &\mathbb{E}_{\mathbb{P}_n} e^{iu^*\sqrt{n}(\widehat{\Phi}_n - \Phi(F))}\varphi(F)\psi_p(I_n^F) \\ &\quad = \int_B \left(\mathbb{E}_{P_n^f} e^{iu^*\sqrt{n}(\widehat{\Phi}_n - \Phi(f))}\varphi(f)\psi_p(I_n^f)\right) dP^{F+h/\sqrt{n}}(f) \\ &\qquad + O\left(\left\| P^{F+h/\sqrt{n}} - P^F \right\|_{Var}\right), \end{aligned}$$

and from H0, we obtain

$$\begin{aligned} &\lim_n \mathbb{E}_{\mathbb{P}_n} e^{iu^*\sqrt{n}(\widehat{\Phi}_n - \Phi(F))}\varphi(F)\psi_p(I_n^F) \\ &\quad = \lim_n \int_B \left(\mathbb{E}_{P_n^f} e^{iu^*\sqrt{n}(\widehat{\Phi}_n - \Phi(f))}\varphi(f)\psi_p(I_n^f)\right) dP^{F+h/\sqrt{n}}(f), \end{aligned} \tag{7}$$

where the right hand side term of (7) is equal to

$$\int_B \left(\mathbb{E}_{P_n^{f+h/\sqrt{n}}} e^{iu^* \sqrt{n}(\widehat{\Phi}_n - \Phi(f+h/\sqrt{n}))} \varphi(f+h/\sqrt{n}) \psi_p(I_n^{f+h/\sqrt{n}}) \right) dP^F(f).$$

Now from H2(a), we have

$$\begin{aligned} &\lim_n \mathbb{E}_{P_n^{f+h/\sqrt{n}}} e^{iu^* \sqrt{n}(\widehat{\Phi}_n - \Phi(f+h/\sqrt{n}))} \varphi(f+h/\sqrt{n}) \psi_p(I_n^{f+h/\sqrt{n}}) \\ &\quad = \lim_n \mathbb{E}_{P_n^f} e^{iu^* \sqrt{n}(\widehat{\Phi}_n - \Phi(f+h/\sqrt{n}))} \varphi(f+h/\sqrt{n}) \psi_p(I_n^{f+h/\sqrt{n}}) e^{Z_n^f(h)}. \end{aligned}$$

It follows from H2(c), H3 and the uniform integrability of $e^{Z_n^f(h)}$ that equation (7) can be rewritten as

$$\begin{aligned} &\lim_n \mathbb{E}_{\mathbb{P}_n} e^{iu^* \sqrt{n}(\widehat{\Phi}_n - \Phi(F))} \varphi(F) \psi_p(I_n^F) \\ &\quad = \lim_n \int_B \left(\mathbb{E}_{P_n^f} e^{iu^* \sqrt{n}(\widehat{\Phi}_n - \Phi(f))} \varphi(f) \psi_p(I_n^f) e^{N_n^f(h) - \frac{1}{2}\langle h, I_n^f h\rangle} e^{-iu^* \langle \dot{\Phi}(f), h\rangle} \right) dP^F(f) \\ &\quad = \lim_n \mathbb{E}_{\mathbb{P}_n} e^{iu^* \sqrt{n}(\widehat{\Phi}_n - \Phi(F))} \varphi(F) \psi_p(I_n^F) e^{N_n^F(h) - \frac{1}{2}\langle h, I_n^F h\rangle} e^{-iu^* \langle \dot{\Phi}(F), h\rangle}. \end{aligned} \tag{8}$$

From the convergence in law (5), we finally deduce, for all $h \in V_{p_0}$:

$$\mathbb{E} e^{iu^* Z} \varphi(F) \psi_p(I^F) = \mathbb{E} e^{iu^* Z} \varphi(F) \psi_p(I^F) e^{N^F(h) - \frac{1}{2}\langle h, I^F h\rangle} e^{-iu^* \langle \dot{\Phi}(F), h\rangle}.$$

Since p, ψ_p and φ are arbitrary, it implies that for all $h \in V_{p_0}$:

$$\mathbb{E}(e^{iu^* Z} | F, I^F) = e^{-iu^* \langle \dot{\Phi}(F), h\rangle - \frac{1}{2}\langle h, I^F h\rangle} \mathbb{E}(e^{iu^* Z} e^{N^F(h)} | F, I^F), \quad \text{almost surely.} \tag{9}$$

Replacing h by zh, with $z \in \mathbb{Q}$, we obtain, for all $h \in V_{p_0}$:

$$\begin{aligned} &\mathbb{E}(e^{iu^* Z} | F, I^F) = e^{-izu^* \langle \dot{\Phi}(F), h\rangle - \frac{1}{2} z^2 \langle h, I^F h\rangle} \mathbb{E}(e^{iu^* Z} e^{z N^F(h)} | F, I^F), \\ &\quad \text{for all } z \in \mathbb{Q}, \text{ almost surely.} \end{aligned} \tag{10}$$

Considering a regular version of the law of $(Z, N^F(h))$ conditional on (F, I^F), we can show that the map $z \mapsto \mathbb{E}(e^{iu^* Z} e^{z N^F(h)} | F, I^F)$ is analytic. Hence, we can deduce that almost-surely, (10) holds for all $z \in \mathbb{C}$. Finally choosing $z = -i$, we deduce that $\forall h \in V_{p_0}$, we have almost surely

$$\mathbb{E}(e^{iu^* Z} | F, I^F) = e^{-\frac{1}{2}(2u^* \langle \dot{\Phi}(F), h\rangle - \langle h, I^F h\rangle)} \mathbb{E}(e^{iu^* Z} e^{-i N^F(h)} | F, I^F). \tag{11}$$

It follows that relation (11) holds, almost surely, for any h in the countable set $\{h_i, i \geq 1\}$. One can consider, for any $h \in H$, the random variable $e^{\frac{1}{2}(2u^* \langle \dot{\Phi}(F), h\rangle - \langle h, I^F h\rangle)} \mathbb{E}(e^{iu^* Z} | F, I^F)$, which is a version of the conditional expectation $\mathbb{E}(e^{iu^* Z} e^{-i N^F(h)} | F, I^F)$ for $h \in \{h_i, i \geq 1\}$. Using the continuity of $h \mapsto 2u^* \langle \dot{\Phi}(F), h\rangle - \langle h, I^F h\rangle$, the L^2-continuity of $h \mapsto N^F(h)$ conditional on (F, I^F), and $\overline{\{h_i, i \geq 1\}} = H$, one can easily check that, this is a version of $\mathbb{E}(e^{iu^* Z} e^{-i N^F(h)} | F, I^F)$ for any $h \in H$. With this choice of conditional expectation, we get that (11) holds true almost surely, with any $h \in H$.

Now, our aim is to maximize with respect to h the quantity $(2u^* \langle \dot{\Phi}(F), h\rangle - \langle h, I^F h\rangle)$.

To finish the proof, we first observe that from H3(b):

$$2u^* \langle \dot{\Phi}(F), h\rangle - \langle h, I^F h\rangle = 2 \sum_{k=1}^{d} u^k \langle I_F^{1/2} h_F^k, h\rangle - \langle h, I^F h\rangle,$$

and since $(I^F)^{1/2}$ is self adjoint we deduce:

$$2u^*\langle \dot{\Phi}(F), h\rangle - \langle h, I^F h\rangle = 2\sum_{k=1}^{d} u^k \langle h_F^k, I_F^{1/2} h\rangle - \langle I_F^{1/2} h, I_F^{1/2} h\rangle. \tag{12}$$

Consequently, the optimization problem can be solved:

$$\begin{aligned}\sup_{h\in H}(2u^*\langle \dot{\Phi}(F), h\rangle - \langle h, I^F h\rangle) &= \sup_{g\in\overline{(I^F)^{1/2}(H)}}\left(2\left\langle \sum_{k=1}^{d} u^k h_F^k, g\right\rangle - \langle g, g\rangle\right),\\ &= \max_{g\in\overline{(I^F)^{1/2}(H)}}\left(2\left\langle \sum_{k=1}^{d} u^k h_F^k, g\right\rangle - \langle g, g\rangle\right) = u^*\Sigma_F u\end{aligned}$$

with $\Sigma_F = (\langle h_F^k, h_F^l\rangle)_{1\le k,l\le d}$, and the maximum is reached for $\sum_{k=1}^{d} u^k h_F^k \in \overline{(I^F)^{1/2}(H)}$.

Now, considering a sequence $(h_m^k)_m$ in H such that $h_F^k = \lim_m (I^F)^{1/2}(h_m^k)$, we have, using (12)

$$u^*\Sigma_F u = \lim_m (2\langle \dot{\Phi}(F), h_m\rangle - \langle h_m, h_m\rangle),$$

with $h_m = \sum_{k=1}^{d} u^k h_m^k$. Turning back to (11) with $h = h_m$, and letting m go to infinity, it follows that

$$\mathbb{E}(e^{iu^*Z}|F, I^F) = e^{-\frac{1}{2}u^*\Sigma_F u}\lim_m \mathbb{E}\left(e^{iu^*Z}e^{-i\sum_k u^k N^F(h_m^k)}\middle| F, I^F\right). \tag{13}$$

Remarking that the function $u \mapsto \lim_m \mathbb{E}(e^{iu^*Z}e^{-i\sum_k u^k N^F(h_m^k)}|F, I^F) = e^{\frac{1}{2}u^*\Sigma_F u}\mathbb{E}(e^{iu^*Z}|F, I^F)$ is continuous at zero, this is the Fourier transform of a probability measure and Theorem 1 is proved. □

Remark 3. We can remark that if H3(b) fails, $\sup_{h\in H}(2u^*\langle \dot{\Phi}(F), h\rangle - \langle h, I^F h\rangle)$ can be infinite. This is the case if there exists k such that $\dot{\Phi}^k(f) \notin \overline{(I^f)^{1/2}(H)}$. This means that the rate of estimation of $\Phi(F)$ is slower than $\sqrt{n}$.

Now, we will extend Theorem 1 in a situation where $\Phi(F)$ is replaced by a sequence of functions $\Phi_n(F)$ defined on $H \times E_n$. This needs the following modification of the hypothesis H3.

H3(n). Regularity of $(\Phi_n)_n$.

(a) We assume that for all n, the restriction of Φ_n on $B \times E_n$ is $\mathcal{B}\otimes\mathcal{E}_n$-measurable and that $f \mapsto \Phi_n(f)$ is differentiable on H, in the following sense : for $1 \le k \le d$, there exists $(\dot{\Phi}_n^k(f))_n$ a sequence of vectors in H such that $\forall h \in H$,

$$\Phi_n^k(f + h/\sqrt{n}) - \Phi_n^k(f) = \frac{1}{\sqrt{n}}\langle \dot{\Phi}_n^k(f), h\rangle + \frac{1}{\sqrt{n}} o_{P_n^f}(1).$$

We assume moreover that for all $h \in H_0$, and for all $1 \le k \le d$, $\dot{\Phi}_n^k(f + h/\sqrt{n}) - \dot{\Phi}_n^k(f)$ goes to zero in P_n^f-probability.

(b) We assume that $\forall f \in B$ and, $\forall p \ge 1$, $\forall h_1, \ldots, h_p \in H_0$, the following convergence in law holds

$$\begin{pmatrix} (\dot{\Phi}_n^k(f))_{1\le k\le d} \\ (N_n^f(h_i))_{1\le i\le p} \\ (\langle h_i, I_n^f h_j\rangle)_{1\le i,j\le p}\end{pmatrix} \underset{P_n^f}{\Longrightarrow} \begin{pmatrix} (\dot{\Phi}^k(f))_{1\le k\le d} \\ (N^f(h_i))_{1\le i\le p} \\ (\langle h_i, I^f h_j\rangle)_{1\le i,j\le p}\end{pmatrix}.$$

We assume that $\dot{\Phi} = (\dot{\Phi}^k)_{1\le k\le d}$ is $\mathcal{B}\otimes\mathcal{F}$ measurable and that for $1 \le k \le d$, $P^F \otimes P$ almost surely, $\dot{\Phi}^k(f) \in (I^f)^{1/2}(H)$.

With these assumptions, we can state an extension of Theorem 1.

Theorem 2. *Let $(\widehat{\Phi}_n)_n$ be any estimator, such that*

$$\sqrt{n}(\widehat{\Phi}_n - \Phi_n(F)) \underset{\mathbb{P}_n}{\Longrightarrow} Z. \tag{14}$$

Then assuming H0, H1, H2, H3*(n), the law of Z is a convolution:*

$$Z \underset{law}{=} \Sigma_F^{1/2} G + R, \quad \text{with } \Sigma_F = (\langle (I^F)^{-1/2}\dot{\Phi}^k(F), (I^F)^{-1/2}\dot{\Phi}^l(F)\rangle)_{1\le k,l\le d}, \tag{15}$$

where conditionally on $(F, I^F, \dot{\Phi}(F))$, R is a random variable independent of G, G is a standard Gaussian vector in $\mathbb{R}^d$.

Proof. The proof is similar to the proof of Theorem 1 and just consists to add $\dot{\Phi}(F) = (\dot{\Phi}^k(F))_{1\le k\le d}$ in the conditioning. We first remark that from H1 and H3(n) we have the convergence in law under $\mathbb{P}_n$:

$$\begin{pmatrix} F \\ (\dot{\Phi}_n^k(F))_{1\le k\le d} \\ (N_n^F(h_i))_{1\le i\le p} \\ (\langle h_i, I_n^F h_j\rangle)_{1\le i,j\le p} \end{pmatrix} \underset{\mathbb{P}_n}{\Longrightarrow} \begin{pmatrix} F \\ (\dot{\Phi}^k(F))_{1\le k\le d} \\ (N^f(h_i))_{1\le i\le p} \\ (\langle h_i, I^F h_j\rangle)_{1\le i,j\le p} \end{pmatrix}. \tag{16}$$

This leads to the modification of (6), adding $\dot{\Phi}(F)$:

$$\mathbb{E}e^{iu^*Z}\varphi(F)\psi_p(I^F)\chi(\dot{\Phi}(F)) = \lim_n \mathbb{E}_{\mathbb{P}_n} e^{iu^*\sqrt{n}(\widehat{\Phi}_n - \Phi_n(F))}\varphi(F)\psi_p(I_n^F)\chi(\dot{\Phi}_n(F)), \tag{17}$$

for a continuous bounded function $\chi : H^d \mapsto \mathbb{R}$, where the limit is taken along a subsequence. We conclude following the same steps as in the proof of Theorem 1. □

3. Applications

In this section, we discuss various applications of the preceding abstract result, based on the observation of a discretized diffusion process. More precisely, we consider the process $(X(t))$ on the time interval [0, 1], the solution of

$$X(t) = x_0 + \int_0^t b(X(s))ds + \int_0^t a(X(s), \sigma(s))dW(s), \tag{18}$$

where $W = (W^i)_{1\le i\le q}$ is a q-dimensional standard Brownian motion, b and a two functions such that $b : \mathbb{R}^q \mapsto \mathbb{R}^q$ and $a : \mathbb{R}^q \times \mathbb{R}^{q'} \mapsto \mathbb{R}^{q\times q}$. We will note a^* the transpose of the matrix a and in the sequel we use the notation $S = aa^*$.

We assume that $\sigma(t)$ is an Itô process, of dimension q', the solution of

$$d\sigma(t) = \beta(t)dt + \gamma(t)dB(t) \tag{19}$$

where B is a q'-dimensional Brownian motion independent of W, and $(\beta(t))$ and $(\gamma(t))$ are progressively measurable, square integrable processes.

We assume that we observe the process X at discrete time $(t_i^n)_{0\le i\le n}$ with $t_0^n = 0$ and $t_n^n = 1$ and we are interested to apply our convolution theorem to the estimation of $\Phi(X, \sigma)$

$= \int_0^1 \phi(X(s), \sigma(s))ds$ for $\phi : \mathbb{R}^q \times \mathbb{R}^{q'} \mapsto \mathbb{R}^d$, from the observations $(X(t_i^n))_i$. For example, if $\phi = a^2$ then $\Phi(X, \sigma) = \int_0^1 a^2(X(s), \sigma(s))ds$ is the integrated volatility.

This statistical problem can easily be related to the abstract framework of Section 2. In fact, the statistical experiment is $(E_n, \mathcal{E}_n, P_n)$, where $E_n = (\mathbb{R}^q)^n$, $\mathcal{E}_n$ is its Borel sigma-field and P_n is the law of $(X(t_i^n))_{1\le i\le n}$. The random parameter $F = \sigma$ takes value in the space $B = \mathcal{C}([0, 1], \mathbb{R}^{q'})$. The Hilbert space H is $L^2([0, 1], \mathbb{R}^{q'})$ with inner product $\langle f, g\rangle = \int_0^1 f^*(t)g(t)dt$ and H_0 the Cameron Martin space. We will note P^σ the law of σ on $\mathcal{C}([0, 1], \mathbb{R}^{q'})$ and we assume that the matrix γ is non degenerated.

A0. $\exists \underline{\gamma} > 0$, such that, almost surely, $\forall t \in [0, 1]$, $(\gamma\gamma^*)(t) \ge \underline{\gamma} Id$.

From the Girsanov theorem, it is easy to check that, assuming A0, the regularity assumption H0 on P^σ is satisfied. Now, if we note P_n^f the law of $(X(t_i^n))_{1\le i\le n}$ conditionally on $\sigma = f$, then H1 is verified. We can observe that P_n^f is the law of $(X^f(t_i^n))_{1\le i\le n}$, where the process $(X^f(t))$ is the solution of

$$X^f(t) = x_0 + \int_0^t b(X^f(s))ds + \int_0^t a(X^f(s), f(s))dW(s). \tag{20}$$

The first step to apply the results of Section 2 is to prove the LAMN property (hypothesis H2) for the family (P_n^f).

3.1. LAMN property

The proof of the LAMN property requires some regularity assumptions on the coefficients b and a of Eq. (18) and on the discretization times $(t_i^n)_i$.

A1. Discretization times. (a) We assume that the measure $\mu_n = \frac{1}{n}\sum_{i=0}^{n-1} \delta_{t_i^n}$ converges weakly, as n goes to infinity, to a measure μ, and that $\sup_i |t_{i+1}^n - t_i^n| \to 0$.

(b) We assume moreover that μ admits a density μ_0 with respect to the Lebesgue measure such that $\forall x$, $\mu_0(x) \ge m > 0$.

A2. Regularity of the coefficients. (a) The functions a and b are $\mathcal{C}^3$ with bounded derivatives. We note $\dot{a}^{(l)}(x, y) = \frac{\partial a}{\partial y^l}(x, y)$ the derivative of a with respect to the coordinate l of the second variable, for $1 \le l \le q'$.

(b) There exist two constants $\underline{a}$ and $\overline{a}$ such that $\forall x \in \mathbb{R}^q$, $\forall y \in \mathbb{R}^{q'}$, $0 < \underline{a} Id \le (aa^*)(x, y) \le \overline{a} Id$.

To simplify the presentation, we give first the LAMN property in the case $q' = 1$ and then extend it to the general case.

3.1.1. Case $q' = 1$

In this subsection, a is defined on $\mathbb{R}^q \times \mathbb{R}$ with value in $\mathbb{R}^{q\times q}$ and we can simplify the notation given in A2(a): we note $\dot{a}(x, y) = \frac{\partial a}{\partial y}(x, y)$ the derivative of a with respect to the second variable.

Proposition 1. *We assume* H1 *and* H2, *then, for all* $f \in \mathcal{C}([0, 1], \mathbb{R})$, *the family* (P_n^f) *satisfies the LAMN property in the direction* H_0 *as defined in Definition* 1, *with* $N_n^f(h)$ *given by:*

$$N_n^f(h) = \sum_{i=0}^{n-1} \frac{h(t_i^n)}{\sqrt{n}(t_{i+1}^n - t_i^n)} \left(\Delta W_{t_i^n}^* (a^{-1}\dot{a})_{t_i^n}^* \Delta W_{t_i^n} - (t_{i+1}^n - t_i^n) Tr(a^{-1}\dot{a})_{t_i^n}^* \right), \tag{21}$$

where $\Delta W_{t_i^n} = W(t_{i+1}^n) - W(t_i^n)$ *and* $(a^{-1}\dot{a})_t^* = (a^{-1}\dot{a})^*(X^f(t), f(t))$. *The operator* I_n^f *is the multiplication operator defined by:*

$$(I_n^f h)(t) := I_n^f(t)h(t) = Tr\left((a^{-1}\dot{a})_{t_i^n}^*(a^{-1}\dot{a})_{t_i^n}^* + (a^{-1}\dot{a})_{t_i^n}(a^{-1}\dot{a})_{t_i^n}^*\right)\mu_0(t)h(t),$$
$$t_i^n \le t < t_{i+1}^n, \tag{22}$$

and $I^f h$ *is given by:*

$$(I^f h)(t) := I^f(t)h(t) = Tr\left((a^{-1}\dot{a})_t^*(a^{-1}\dot{a})_t^* + (a^{-1}\dot{a})_t(a^{-1}\dot{a})_t^*\right)\mu_0(t)h(t). \tag{23}$$

Moreover we have the convergence in law:

$$\begin{pmatrix} W \\ (N_n^f(h_i))_{1\le i\le p} \\ (\langle h_i, I_n^f h_j\rangle)_{1\le i,j\le p} \end{pmatrix} \Longrightarrow \begin{pmatrix} W \\ \left(\int_0^1 h_i(s)\sqrt{I^f(s)}d\tilde{W}(s)\right)_{1\le i\le p} \\ (\langle h_i, I^f h_j\rangle)_{1\le i,j\le p} \end{pmatrix}$$

where $\tilde{W}$ *is a Brownian motion independent of* W.

We can remark that the convergence in law in Proposition 1 is stronger than the one of Definition 1. This result is a straightforward consequence of the LAMN property given by Gobet [4] and we postpone the sketch of the proof to the Appendix.

Remark 4. Using the notation $S_t = (aa^*)(X^f(t), f(t))$, we can observe that

$$Tr\left((a^{-1}\dot{a})_t^*(a^{-1}\dot{a})_t^* + (a^{-1}\dot{a})_t(a^{-1}\dot{a})_t^*\right) = \frac{1}{2}Tr\left((\dot{S}S^{-1}\dot{S}S^{-1})_t\right),$$

and consequently

$$I^f(t) = \frac{1}{2}Tr\left((\dot{S}S^{-1}\dot{S}S^{-1})_t\right)\mu_0(t).$$

3.1.2. General case

The LAMN property remains true for the family (P_n^f), with $f \in \mathcal{C}([0,1], \mathbb{R}^{q'})$. For $h = (h^l)_{1\le l\le q'} \in H_0$, we have:

$$N_n^f(h) = \sum_{i=0}^{n-1} \frac{1}{\sqrt{n}(t_{i+1}^n - t_i^n)} \sum_{l=1}^{q'} h^l(t_i^n)$$
$$\times \left(\Delta W_{t_i^n}^*(a^{-1}\dot{a}^{(l)})_{t_i^n}^* \Delta W_{t_i^n} - (t_{i+1}^n - t_i^n)Tr(a^{-1}\dot{a}^{(l)})_{t_i^n}^*\right), \tag{24}$$

where $(a^{-1}\dot{a}^{(l)})_t^* = (a^{-1}\dot{a}^{(l)})^*(X^f(t), f(t))$. Moreover, the operators I_n^f and I^f are the multiplication operators $(I_n^f h)(t) = I_n^f(t)h(t)$, $(I^f h)(t) = I^f(t)h(t)$, where the matrices $I_n^f(t)$ and $I^f(t)$ (of dimension $q' \times q'$) are respectively given by:

$$I_n^f(t)_{l,l'} = Tr\left((a^{-1}\dot{a}^{(l)})_{t_i^n}^*(a^{-1}\dot{a}^{(l')})_{t_i^n}^* + (a^{-1}\dot{a}^{(l)})_{t_i^n}(a^{-1}\dot{a}^{(l')})_{t_i^n}^*\right)\mu_0(t),$$
$$t_i^n \le t < t_{i+1}^n, \tag{25}$$

$$I^f(t)_{l,l'} = Tr\left((a^{-1}\dot{a}^{(l)})_t^*(a^{-1}\dot{a}^{(l')})_t^* + (a^{-1}\dot{a}^{(l)})_t(a^{-1}\dot{a}^{(l')})_t^*\right)\mu_0(t). \tag{26}$$

As previously, we can observe that

$$I^f(t)_{l,l'} = \frac{1}{2} Tr \left((\dot{S}^{(l)} S^{-1} \dot{S}^{(l')} S^{-1})_t \right) \mu_0(t). \tag{27}$$

3.2. Asymptotic lower bound for the estimation of $\int_0^1 \phi(X(s), \sigma(s))ds$

We state, in this section, the convolution theorem for the estimation of $\Phi(X, \sigma) = \int_0^1 \phi(X(s), \sigma(s))ds$ for $\phi : \mathbb{R}^q \times \mathbb{R}^{q'} \mapsto \mathbb{R}^d$, from the observations $(X(t_i^n))_i$. We make the following assumption on ϕ.

A3. Regularity of ϕ. (a) We assume that $\phi = (\phi^k)_{1 \le k \le d}$ is $\mathcal{C}^1$ (with respect to both variables), and that ϕ and its derivatives are bounded. We note $\dot{\phi}^k$ the vector, in $\mathbb{R}^{q'}$, of partial derivatives with respect to the second variable: $\dot{\phi}^k(x, y) = (\frac{\partial \phi^k}{\partial y^l}(x, y))_{1 \le l \le q'}$.

(b) We assume that $\forall k, \forall f \in \mathcal{C}([0, 1], \mathbb{R}^{q'})$, there exists $h_f^k \in (\mathrm{K}er(I^f)^{1/2})^{\perp}$ such that

$$\dot{\phi}^k(X(t), f(t)) = (I^f(t))^{1/2} h_f^k(t),$$

where the matrix $I^f(t)$ is defined in equation (27).

Remark that A3(b) is satisfied as soon as $I^f(t)$ is invertible for all t.

With these assumptions, we deduce from Theorem 2 the following convolution theorem.

Theorem 3. *Let $\widehat{\Phi}_n$ be any estimator of $\Phi(X, \sigma)$ such that*

$$\sqrt{n}(\widehat{\Phi}_n - \Phi(X, \sigma)) \underset{\mathbb{P}_n}{\Longrightarrow} Z. \tag{28}$$

We assume A0, A1, A2, A3, *and that* $\lim_n \sqrt{n} \sup_i |t_{i+1}^n - t_i^n| = 0$. *Then the law of Z is a convolution:*

$$Z \underset{law}{=} \Sigma_\sigma^{1/2} G + R, \tag{29}$$

with

$$\Sigma_\sigma = \left(\int_0^1 (I^\sigma(t)^{-1/2} \dot{\phi}^k(X(t), \sigma(t)))^* I^\sigma(t)^{-1/2} \dot{\phi}^l(X(t), \sigma(t)) dt \right)_{1 \le k,l \le d}, \tag{30}$$

and where conditionally on $(\sigma, I^\sigma, (\dot{\phi}^k(X, \sigma))_k)$, R is a random variable independent of G, G is a standard Gaussian vector in $\mathbb{R}^d$. Moreover, R is independent of G conditionally on (σ, X).

Remark 5. The result of Theorem 3 shows in particular that, in our context, the optimal rate of convergence in estimating $\Phi(X, \sigma)$ is $\sqrt{n}$. This is no longer the case if we add a jump part to the process $(X(t))$; see Jacod and Reiß [10] for a discussion on the optimal rate of convergence of the integrated volatility in the presence of jumps.

Proof. We already proved that assuming A0, A1 and A2, the assumptions H0, H1 and H2 of Section 2.2 are satisfied.

Now, let $\Phi_n(X, \sigma) = \int_0^1 \phi(X(\varphi_n(t)), \sigma(t))dt$, where $\varphi_n(t) = t_i^n$ if $t_i^n \le t < t_{i+1}^n$. Assuming A3(a) and $\lim_n \sqrt{n} \sup_i |t_{i+1}^n - t_i^n| = 0$, then, by standard arguments, $\sqrt{n}(\Phi(X, \sigma) - \Phi_n(X, \sigma))$ goes to zero in $\mathbb{P}_n$-probability and consequently

$$\sqrt{n}(\widehat{\Phi}_n - \Phi_n(X, \sigma)) \underset{\mathbb{P}_n}{\Longrightarrow} Z.$$

So to apply Theorem 2, we just have to check H3(n). For $f \in B$, using the notations of Section 2.2, we have $\dot{\Phi}_n^k(f) = \dot{\phi}^k(X(\varphi_n(.)), f(.))$ and $\dot{\Phi}^k(f) = \dot{\phi}^k(X(.), f(.))$. We check easily that $\dot{\phi}^k(X(\varphi_n(.)), f(.))$ converges to $\dot{\Phi}^k(f)$ in P_n^f-probability and assuming A3, we deduce H3(n). This gives the first part of Theorem 3.

To obtain the independence of R and G conditionally on (σ, X), we turn back to the proof of Theorem 1 with the following modification. Let $\psi_r : \mathbb{R}^r \mapsto \mathbb{R}$ be a continuous bounded function. We consider the instants $0 < t_1 < \cdots < t_r < 1$, $r \geq 1$. Since we have the LAMN property with a stable convergence in law, we deduce that we have the convergence in law (stronger than (4))

$$\begin{pmatrix} (\sigma(t))_t \\ \psi_r(X(\varphi_n(t_1)), \ldots, X(\varphi_n(t_r))) \\ (N_n^\sigma(h_i))_{1\leq i\leq p} \\ (\langle h_i, I_n^\sigma h_j\rangle)_{1\leq i,j\leq p} \end{pmatrix} \underset{\mathbb{P}_n}{\Longrightarrow} \begin{pmatrix} (\sigma(t))_t \\ \psi_r(X(t_1), \ldots, X(t_r)) \\ \left(\int_0^1 h_i(s)\sqrt{I^\sigma(s)}d\tilde{W}(s)\right)_{1\leq i\leq p} \\ (\langle h_i, I^\sigma h_j\rangle)_{1\leq i,j\leq p} \end{pmatrix}.$$

So by the same arguments as those given in the proof of Theorem 1 and replacing (6) by

$$\begin{aligned} &\mathbb{E}e^{iu^*Z}\varphi(\sigma)\psi_r(X(t_1), \ldots, X(t_r)) \\ &= \lim_n \mathbb{E}_{\mathbb{P}_n} e^{iu^*\sqrt{n}(\widehat{\Phi}_n - \Phi(F))}\varphi(\sigma)\psi_r(X(\varphi_n(t_1)), \ldots, X(\varphi_n(t_r))), \end{aligned} \tag{31}$$

we deduce the decomposition of the law of Z conditionally on (σ, X). □

3.3. Discussion on the efficiency in the p-variation estimation

3.3.1. X and σ of dimension 1

As an illustration of Theorem 3, we consider the estimation of $\int_0^1 a^p(X(t), \sigma(t))dt$ in the simple case $d = q = q' = 1$, for $p \geq 2$. We have $\phi(x, y) = a^p(x, y)$ ($\phi : \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$). Then $I^\sigma(t) = 2\mu_0(t)\frac{\dot{a}^2(X(t),\sigma(t))}{a^2(X(t),\sigma(t))}$ and $\dot{\phi}(x, y) = p(a^{p-1}\dot{a})(x, y)$. We remark that A3 is true if a is $\mathcal{C}^1$. Consequently, from Theorem 3, we deduce the following proposition.

Proposition 2. *We assume* A0, A1, A2*, and* $\lim_n \sqrt{n}\sup_i |t_{i+1}^n - t_i^n| = 0$*, then any estimator of* $\int_0^1 a^p(X(t), \sigma(t))dt$*, with rate of convergence* $\sqrt{n}$*, has an asymptotic conditional variance, on* (σ, X)*, greater than*

$$\Sigma_\sigma = \frac{p^2}{2}\int_0^1 a^{2p}(X(t), \sigma(t))1_{\{\dot{a}(X(t),\sigma(t))\neq 0\}}\frac{1}{\mu_0(t)}dt. \tag{32}$$

We can remark that assuming $\dot{a}(x, y) \neq 0$, $\forall x, y$, the asymptotic minimal variance is

$$\Sigma_\sigma = \frac{p^2}{2}\int_0^1 a^{2p}(X(t), \sigma(t))\frac{1}{\mu_0(t)}dt,$$

and in this case we can discuss the efficiency of classical power variation estimators defined by

$$V_n(p) = \frac{1}{m_p}\sum_{i=0}^{n-1}(t_{i+1}^n - t_i^n)^{1-p/2}|X(t_{i+1}^n) - X(t_i^n)|^p,$$

where m_p denotes the pth absolute moment of a standard normal law. We refer the reader to Jacod [9] and Hayashi–Jacod–Yoshida [6] for the asymptotic properties of these estimators. In our simple case, where the discretization times are deterministic, one can easily see (assuming A1(a) and A2) that $V_n(p)$ converges in probability to $\int_0^1 a^p(X(t),\sigma(t))dt$.

Now if we consider the uniform discretization scheme $t_i^n = i/n$, then we have the convergence in law

$$\sqrt{n}(V_n(p) - \int_0^1 a^p(X(t),\sigma(t))dt) \Longrightarrow \frac{\sqrt{m_{2p}-m_p^2}}{m_p}\int_0^1 a^p(X(t),\sigma(t))d\tilde{W}(t).$$

In this case, A1 is verified with $\mu_0 = 1$, and from Proposition 2 we deduce that $V_n(2)$ is efficient. For $p = 4$, a simple calculation gives $\frac{m_8-m_4^2}{m_4^2} = \frac{96}{9} > 8$, and consequently $V_n(4)$ is not efficient (see Jacod and Rosenbaum [11] for the construction of the efficient estimator of $\int_0^1 a^p(X(s),\sigma(s))ds$ in a more general context).

The situation is more complicated for general discretization schemes, even in the deterministic case, and we restrict ourself to the study of $V_n(2)$. If we make the additional assumption on the discretization scheme (see [6])

$$n\sum_{i=0}^{N_t^n}(t_{i+1}^n - t_i^n)^2 \to \int_0^t a_2(s)ds, \tag{33}$$

where $N_t^n = \sup_i\{i; t_n^i \le t\}$ then we have (see Theorem 3.2 of [6])

$$\sqrt{n}\left(V_n(p) - \int_0^1 a^p(X(t),\sigma(t))dt\right) \Longrightarrow \frac{\sqrt{m_{2p}-m_p^2}}{m_p}\int_0^1 a^p(X(t),\sigma(t))\sqrt{a_2(s)}d\tilde{W}(t).$$

The comparison between $\mu_0(s)$ and $a_2(s)$ is not straightforward in general. However if $t_i^n = g(i/n)$ for a smooth, strictly increasing, function g, mapping $[0,1]$ to $[0,1]$, then $a_2(s) = g'(g^{-1}(s)) = 1/\mu_0(s)$ and we can conclude that $V_n(2)$ is efficient.

3.3.2. *Efficiency for higher dimensions*

Assume that the dimension of the process X is $q \ge 1$ and that one want to estimate the integrated covariance matrix of the process $V = \int_0^1 S(X(s),\sigma(s))ds$, where $S(x,y) = a(x,y)$ $a(x,y)^*$ is the $q\times q$ symmetric local covariance matrix of X. Thus, V is a $d = q(q+1)/2$ dimensional object, and it is known that the multidimensional quadratic variation $V_n = \sum_{i=0}^{n-1}$ $(X(t_{i+1}^n) - X(t_i^n))(X(t_{i+1}^n) - X(t_i^n))^*$ is a consistent estimator of V.

Assume, for simplicity that the sampling is regular $t_i^n = \frac{i}{n}$, then the asymptotic behavior of this estimator can be found for instance in Jacod–Protter [12] (Th.5.4.2 p.162). The error of estimation $\sqrt{n}(V_n - V)$ converges to a conditionally Gaussian variable, with explicit conditional covariance matrix. The asymptotic conditional covariance between the error of estimation of V_{i_1,j_1} and V_{i_2,j_2} where $1 \le i_1 \le j_1 \le q$ and $1 \le i_2 \le j_2 \le q$ is given by

$$\int_0^1 \left[S(X(s),\sigma(s))_{i_1,i_2}S(X(s),\sigma(s))_{j_1,j_2} + S(X(s),\sigma(s))_{i_1,j_2}S(X(s),\sigma(s))_{j_1,i_2}\right]ds. \tag{34}$$

Let us remark, that if the dimension of σ is too small, the quadratic variation might not be efficient. For instance, choose X with dimension 2 and σ one dimensional with the choice $a(x,y) = yI_2$ where I_2 is the unit matrix of size 2. Then, clearly the two components of X are redundant for the estimation of $\int_0^1 \sigma^2(t)dt = \int_0^1 S(X(t), \sigma(t))_{1,1}dt = \int_0^1 S(X(t), \sigma(t))_{2,2}dt$. As a consequence, $\overline{V}_n = \frac{1}{2}\sum_{i=0}^{n-1}[(X_1(t_{i+1}^n) - X_1(t_i^n))^2 + (X_2(t_{i+1}^n) - X_2(t_i^n))^2]$ is clearly an estimator of $V_{1,1}$ with a conditional variance smaller than (34). Moreover, the application of Theorem 3 shows that $\overline{V}_n$ is efficient.

The following proposition states that if q' is large enough, then the quadratic variation is an efficient estimator of the covariance matrix.

Proposition 3. *We assume that* A0 *and* A2 *hold. We denote by* $\mathbb{S}^q$ *the set of symmetric positive definite matrices of size* q *and let* $q' = q(q+1)/2$. *We assume that for all* x, *the function* $y \in \mathbb{R}^{q'} \mapsto S(x, y) \in \mathbb{S}^q$ *is differentiable, and its Jacobian denoted by* DS *is invertible for all* x *and* y.

Assume that $\widehat{V}_n$ *is an estimator of* V *such that* $\sqrt{n}(\widehat{V}_n - V) \underset{\mathbb{P}_n}{\Longrightarrow} Z$. *Then,* Z *is the sum of a conditionally centered Gaussian variable, whose conditional variance is described by* (34), *and some conditionally independent variable.*

Proof. Note $\mathcal{I} = \{(i,j) \mid 1 \le i \le j \le q\}$ and define for $(i,j) \in \mathcal{I}$, the symmetric matrix $E_{i,j}$ as the matrix with all entries equal to zero except the one with index (i,j) or (j,i) where the entry is one. The family $(E_{i,j})_{(i,j)\in\mathcal{I}}$ defines a canonical basis of $\mathbb{S}^q$.

We now apply Theorem 3 with $\phi(X(s), \sigma(s)) = S(X(s), \sigma(s)) = a(X(s), \sigma(s))a(X(s), \sigma(s))^* \in \mathbb{S}^q$, where we consider elements of $\mathbb{S}^q$ as vectors of dimension $d = q(q+1)/2$ indexed by the set $\mathcal{I} = \{(i,j) \mid 1 \le i \le j \le q\}$. With the notation of Theorem 3, we have $\dot{\phi}^{i,j} = (\frac{\partial}{\partial \sigma_l} S(X(s), \sigma(s))_{i,j})_{l=1,\ldots,q'}$ for $(i,j) \in \mathcal{I}$. The Jacobian matrix of $y \in \mathbb{R}^{q'} \in\mapsto S(x,y) \in \mathbb{S}^q$ can be expanded using the canonical basis of $\mathbb{R}^{q'}$ and $\mathbb{S}^q$ as $DS = [\frac{\partial}{\partial \sigma_l} S(X(s), \sigma(s))_{i,j}]_{(i,j)\in\mathcal{I}, l\in\{1,\ldots,q'\}}$.

Recall from (27) that the information matrix in the LAMN property with a multidimensional parameter is:

$$I^\sigma(s)_{l,l'} = \frac{1}{2} Tr\left(\frac{\partial}{\partial \sigma_l} S(X(s), \sigma(s)) S^{-1}(X(s), \sigma(s)) \frac{\partial}{\partial \sigma_{l'}} S(X(s), \sigma(s)) S^{-1}(X(s), \sigma(s))\right) \quad (35)$$

for $(l, l') \in \{1, \ldots, q'\}^2$.

With this setting, by comparison of the expressions (30) and (34), the proof of the proposition consists in showing the relation

$$(\dot{\phi}^{i_1,j_1})^* (I^\sigma(s))^{-1} \dot{\phi}^{i_2,j_2} = S(X(s), \sigma(s))_{i_1,i_2} S(X(s), \sigma(s))_{j_1,j_2} + S(X(s), \sigma(s))_{i_1,j_2} S(X(s), \sigma(s))_{j_1,i_2}. \quad (36)$$

For the sake of shortness we denote the matrix $S(X(s), \sigma(s))$ as S in the rest of the proof. Using that $(E_{i,j})_{1\le i\le j\le q}$ is the canonical basis of $\mathbb{S}^q$, we can rewrite (35) as

$$I^\sigma(s)_{l,l'} = \frac{1}{2} \sum_{1\le i_1\le j_1\le q} \sum_{1\le i_2\le j_2\le q} \frac{\partial S_{i_1,j_1}}{\partial \sigma_l} \frac{\partial S_{i_2,j_2}}{\partial \sigma_{l'}} Tr(E_{i_1,j_1} S^{-1} E_{i_2,j_2} S^{-1}).$$

This relation can be restated as $I^\sigma(s) = (DS)^* J(S)(DS)$ where $J(S)$ is a matrix of size $d \times d$, with components indexed by $\mathcal{I}$ and with entries given by $J(S)_{(i_1,j_1),(i_2,j_2)} = \frac{1}{2} Tr(E_{i_1,j_1} S^{-1} E_{i_2,j_2} S^{-1})$, and DS is the Jacobian matrix of S.

We now write for (i_1, j_1) and (i_2, j_2) in $\mathcal{I}$:

$$(\dot\phi^{i_1,j_1})^*(I^\sigma(s))^{-1}\dot\phi^{i_2,j_2} = (DS(I^\sigma(s))^{-1}DS^*)_{(i_1,j_1),(i_2,j_2)}$$
$$= (DS(DS)^{-1}J(S)^{-1}(DS^*)^{-1}DS^*)_{(i_1,j_1),(i_2,j_2)} = (J(S)^{-1})_{(i_1,j_1),(i_2,j_2)}.$$

Using Lemma 1 we know that $(J(S)^{-1})_{(i_1,j_1),(i_2,j_2)} = S_{i_1,i_2}S_{j_1,j_2} + S_{j_1,i_2}S_{i_1,j_2}$, and equation (36) follows. □

Lemma 1. *Let $S \in \mathbb{S}^q$ and define the matrix of size $d \times d$ indexed by $\mathcal{I}$ as $J(S)_{(i_1,j_1),(i_2,j_2)} = \frac{1}{2}Tr(E_{i_1,j_1}S^{-1}E_{i_2,j_2}S^{-1})$ where $(E_{i,j})_{(i,j)\in\mathcal{I}}$ is the canonical basis of $\mathbb{S}^q$.*

Then the matrix $J(S)$ is invertible and its inverse is $V(S)$ defined by $V(S)_{(i_1,j_1),(i_2,j_2)} = S_{i_1,i_2}S_{j_1,j_2} + S_{j_1,i_2}S_{i_1,j_2}$.

Proof. For (i_1, j_1), (i_2, j_2) in $\mathcal{I}$ we compute,

$$\begin{aligned}(V(S)I(S))_{(i_1,j_1),(i_2,j_2)} &= \sum_{(i_3,j_3)\in\mathcal{I}} V(S)_{(i_1,j_1),(i_3,j_3)}I(S)_{(i_3,j_3),(i_2,j_2)}\\
&= \frac{1}{2}\sum_{(i_3,j_3)\in\mathcal{I}}(S_{i_1,i_3}S_{j_1,j_3} + S_{j_1,i_3}S_{i_1,j_3})Tr(E_{i_3,j_3}S^{-1}E_{i_2,j_2}S^{-1})\\
&= \frac{1}{2}Tr\left(\sum_{(i_3,j_3)\in\mathcal{I}}(S_{i_1,i_3}S_{j_1,j_3} + S_{j_1,i_3}S_{i_1,j_3})E_{i_3,j_3}S^{-1}E_{i_2,j_2}S^{-1}\right).\end{aligned}$$

Denote by M the symmetric matrix of size $q \times q$ whose entries are $M_{i_3,j_3} = S_{i_1,i_3}S_{j_1,j_3} + S_{j_1,i_3}S_{i_1,j_3}$. Then the equations above yield,

$$\begin{aligned}(V(S)I(S))_{(i_1,j_1),(i_2,j_2)} &= \frac{1}{2}Tr(MS^{-1}E_{i_2,j_2}S^{-1}) = \frac{1}{2}Tr(S^{-1}MS^{-1}E_{i_2,j_2})\\
&= \begin{cases}\frac{1}{2}(S^{-1}MS^{-1})_{i_2,j_2} & \text{if } i_2 = j_2\\ (S^{-1}MS^{-1})_{i_2,j_2} & \text{if } i_2 \neq j_2.\end{cases}\end{aligned}$$

With a few algebra, we show that $(S^{-1}MS^{-1})_{i_2,j_2} = (S^{-1}S)_{i_2,i_1}(SS^{-1})_{j_1,i_2} + (S^{-1}S)_{i_2,j_1}(SS^{-1})_{i_1,j_2}$. Then, it is straightforward to deduce that

$$(V(S)I(S))_{(i_1,j_1),(i_2,j_2)} = 1 \quad \text{if } (i_1, j_1) = (i_2, j_2) \quad \text{and}$$
$$(V(S)I(S))_{(i_1,j_1),(i_2,j_2)} = 0 \quad \text{if } (i_1, j_1) \neq (i_2, j_2).$$

This proves that $I(S)^{-1} = V(S)$. □

3.4. Efficient scheme of approximation

In this section, we apply our convolution theorems to prove that some schemes of approximation are efficient.

3.4.1. Approximation of the stochastic integral

Assume that we are in the one dimensional case $q = q' = 1$. Let χ be some $\mathcal{C}^2$ function from $\mathbb{R}^2$ to $\mathbb{R}$ and set $\Psi = \int_0^1 \chi(s, X(s))dX(s)$. The problem of approximating such stochastic integral from $(X(t_i^n))_i$ has been the subject of many works [3,1,20]. This problem is related to the hedging of financial assets.

For simplicity, assume that the sampling is regular and consider $\Psi_n = \sum_{i=0}^{n-1} \chi(\frac{i}{n}, X(\frac{i}{n})) (X(\frac{i+1}{n}) - X(\frac{i}{n}))$ the associated Riemann sum. Then it can be shown (see [20,8])

$$\sqrt{n}(\Psi_n - \Psi) \underset{\mathbb{P}_n}{\Longrightarrow} \frac{1}{\sqrt{2}} \int_0^1 \chi_x'(s, X(s)) a(X(s), \sigma(s))^2 d\tilde{W}(s), \tag{37}$$

where $\tilde{W}$ is some independent Brownian motion and χ_x' the derivative of χ with respect to the second variable.

The following proposition shows that the Riemann sum cannot be improved for reconstructing Ψ.

Proposition 4. *Assume* A0, A2 *and let* $(\widehat{\Psi}_n)_n$ *be any sequence of measurable functions of* $(X(\frac{i}{n}))_{i=0,\dots,n}$, *such that* $\sqrt{n}(\widehat{\Psi}_n - \Psi) \underset{\mathbb{P}_n}{\Longrightarrow} Z$. *Then, the variable* Z *admits the decomposition*

$$Z \underset{law}{=} \frac{1}{\sqrt{2}} \int_0^1 \chi_x'(s, X(s)) a(X(s), \sigma(s))^2 d\tilde{W}(s) + R,$$

where R *is independent of* $\tilde{W}$, *conditionally on* (σ, X).

Proof. We set $\Phi(X, \sigma) = \frac{1}{2} \int_0^1 \chi_x'(s, X(s)) a(X(s), \sigma(s))^2 ds$ and $F(t, x) = \int_0^x \chi(t, u) du$. From Ito's formula we have, $\Phi(X, \sigma) = F(1, X(1)) - \Psi - \int_0^1 F_t'(s, X(s)) ds$. We define $\widehat{\Phi}_n = F(1, X(1)) - \widehat{\Psi}_n - \frac{1}{n} \sum_{i=0}^{n-1} F_t'(\frac{i}{n}, X(\frac{i}{n}))$, then we have $\sqrt{n}(\widehat{\Phi}_n - \Phi) = \sqrt{n}(\Psi - \widehat{\Psi}_n) + o_{\mathbb{P}_n}(1)$ and consequently $\sqrt{n}(\widehat{\Phi}_n - \Phi) \underset{\mathbb{P}_n}{\Longrightarrow} -Z$.

Now, the proposition follows, by a straightforward extension of Theorem 3 to the estimation of $\Phi(X, \sigma) = \int_0^1 \phi(s, X(s), \sigma(s)) ds$ with $\phi(s, x, y) = \frac{1}{2} \chi_x'(s, x) a^2(x, y)$ and recalling that $I^\sigma(t) = 2 \frac{\dot{a}^2(X(t), \sigma(t))}{a^2(X(t), \sigma(t))}$. □

3.4.2. Approximation of solutions of stochastic differential equations

Assume again that X, σ are solutions of (18)–(19) with $q = q' = 1$. Let g and k be smooth real functions, with at most linear growth, and with $1/g$ bounded. We consider the stochastic differential equation driven by X,

$$dY(t) = g(Y(t)) dX(t) + k(Y(t)) dt, \quad Y(0) = y_0 \in \mathbb{R}. \tag{38}$$

As an illustration of our convolution result, we can discuss about the efficiency of the approximation of $Y(1)$ from a functional of $(X(i/n))_{i \in \{0,\dots,n\}}$. First, we recall results concerning the Euler scheme approximation of (38). Let us denote $\varphi_n(s) = \sup\{i/n \mid i/n \le s\}$ and the Euler scheme equation is

$$dY^n(t) = g(Y^n(\varphi_n(t))) dX(t) + k(Y^n(\varphi_n(t))) dt, \quad Y^n(0) = y_0. \tag{39}$$

From the results in [13], the error of the Euler scheme is assessed by,

$$\sqrt{n}(Y(1) - Y^n(1)) \underset{\mathbb{P}_n}{\Longrightarrow} U(1),$$

where $(U(t))_t$ is the solution of

$$dU(t) = g'(Y(t))U(t)dX(t) + k'(Y(t))U(t)dt + \frac{1}{\sqrt{2}}g'(Y(t))g(Y(t))a(X(t),\sigma(t))^2 d\tilde{W}(t), \quad U(0) = 0. \tag{40}$$

The equation (40) is linear and can be solved explicitly using the Doléans-Dade exponential. After several computations, one can deduce that $U(1)$ is a mixed normal variable with an explicit conditional variance equal to

$$\frac{g(Y(1))^2}{2}\int_0^1 g'(Y(s))^2 a(X(s),\sigma(s))^4 \exp\left[\int_s^1 \alpha(u)du\right]ds \tag{41}$$

with

$$\alpha(u) = 2\left[k'(Y(u)) - \frac{g'(Y(u))k(Y(u))}{g(Y(u))}\right] - g(Y(u))g''(Y(u))a^2(X(u),\sigma(u)).$$

The following proposition shows that, when the volatility of the driving semi-martingale X is random, it is impossible to find a scheme, based on $(X(i/n))_i$, with an error smaller than the Euler scheme error.

Proposition 5. *Assume that* A0, A2 *hold and let* $(\widehat{\Phi}_n)_n$ *be any sequence of measurable functions of* $(X(\frac{i}{n}))_{i=0,\ldots,n}$*, such that* $\sqrt{n}(\widehat{\Phi}_n - Y(1)) \underset{\mathbb{P}_n}{\Longrightarrow} Z$*. Then, the variable* Z *admits the decomposition*

$$Z \underset{law}{=} U(1) + R,$$

where R *is independent of* $\tilde{W}$*, conditionally on* (σ, X)*.*

Proof. We shall apply our Theorem 2. We need to approach the random variable $Y(1)$ by some sequence $(\Phi_n)_n$ of random variables measurable with respect to $(X(i/n))_i$ and σ.

First, we transform the equation (38) into a simpler equation. To this end, we take H as a primitive function of $1/g$ with $H(y_0) = 0$, and set $V(t) = H(Y(t))$. Then, V is the solution of the stochastic differential equation

$$V(t) = X(t) + \int_0^t \beta(V(s), X(s), \sigma(s))ds, \tag{42}$$

with $\beta(v,x,y) = \frac{k\circ H^{-1}(v)}{g\circ H^{-1}(v)} - \frac{1}{2}a^2(x,y) \times g' \circ H^{-1}(v)$. We note V^f, the solution of Eq. (42), where σ is replaced by f and we have $V = V^\sigma$. We construct an approximation of $V(1)$ based on the sampling $(X(i/n))_i$. More precisely for f a continuous function, we define $(V_n^f(t))_t$, the solution of

$$V_n^f(t) = X(t) + \int_0^t \beta(V_n^f(\varphi_n(s)), X(\varphi_n(s)), f(s))ds. \tag{43}$$

The variable $V_n^\sigma(1)$ is an approximation of $V(1)$, and the difference only involves drift terms of the corresponding equations. Hence, it can be easily shown that the approximation has a rate greater than $\sqrt{n}$:

$$\sqrt{n}(V_n^\sigma(1) - V(1)) \xrightarrow[\text{proba}]{n\to\infty} 0. \tag{44}$$

For f any continuous function we set

$$\Phi_n(f) = H^{-1}(V_n^f(1)). \tag{45}$$

Using (44) it is simple to see that $\sqrt{n}(\Phi_n(\sigma) - Y(1))$ converges to zero in probability and hence $\sqrt{n}(\widehat{\Phi}_n - \Phi_n(\sigma)) = \sqrt{n}(\widehat{\Phi}_n - Y(1)) + \sqrt{n}(Y(1) - \Phi_n(\sigma)) \underset{\mathbb{P}_n}{\Longrightarrow} Z$.

In order to apply Theorem 2, we need to check H3(n) and especially compute $\dot{\Phi}_n$. First, we determine the derivative of $f \mapsto V_n^f(1)$. From standard results about the differentiability of the solution of S.D.E. with respect to parameters, it comes that $\sqrt{n}(V_n^{f+h/\sqrt{n}}(1) - V_n^f(1)) - \mathcal{V}_n^f(1,h) \xrightarrow{n\to\infty} 0$ where $\mathcal{V}_n^f(s,h)$ is the solution of

$$\mathcal{V}_n^f(t,h) = \int_0^t \bigg\{ \dot{\beta}(V_n^f(\varphi_n(s)), X(\varphi_n(s)), f(s))h(s) \\ + \frac{\partial \beta}{\partial v}(V_n^f(\varphi_n(s)), X(\varphi_n(s)), f(s))\mathcal{V}_n^f(\varphi_n(s),h) \bigg\} ds.$$

Solving this linear equation and using (45), we have $\sqrt{n}(\Phi_n(f + h/\sqrt{n}) - \Phi_n(f)) - \langle \dot{\Phi}_n(f), h\rangle \xrightarrow{n\to\infty} 0$ where $\dot{\Phi}_n(f)$ is the element of $L^2([0,1],\mathbb{R})$ given by,

$$\dot{\Phi}_n(f)(s) = g(H^{-1}(V_n^f(1))) \exp\left(\int_s^1 \frac{\partial \beta}{\partial v}(V_n^f(\varphi_n(u)), X(\varphi_n(u)), f(u)) du \right) \\ \times \dot{\beta}(V_n^f(\varphi_n(s)), X(\varphi_n(s)), f(s)),$$

and, by simple computations,

$$\frac{\partial \beta}{\partial v}(v,x,y) = k' \circ H^{-1}(v) - \frac{kg'}{g} \circ H^{-1}(v) - \frac{1}{2}a^2(x,y) \times g \circ H^{-1}(v) \times g'' \circ H^{-1}(v),$$

and

$$\dot{\beta}(v,x,y) = -a(x,y) \times \dot{a}(x,y) \times g' \circ H^{-1}(v).$$

We deduce that assumption H3(n) is satisfied with $\dot{\Phi}(f)$ the element of $L^2([0,1],\mathbb{R})$ given by,

$$\dot{\Phi}(f)(s) = -g(Y^f(1)) \exp\Bigg[\int_s^1 (k'(Y^f(u)) - \frac{kg'}{g}(Y^f(u)) \\ - \frac{1}{2}a^2(X(u), f(u)) g(Y^f(u)) g''(Y^f(u))) du \Bigg] \dot{a}(X(s), f(s)) \times g'(Y^f(s)), \tag{46}$$

where $Y^f = H^{-1}(V^f(1))$.

Now the proposition follows by application of Theorem 2, recalling that for $h \in L^2([0,1],\mathbb{R})$, $I^\sigma h(t) = 2\frac{\dot{a}^2(X(t),\sigma(t))}{a^2(X(t),\sigma(t))} h(t)$, and remarking that by (46), the quantity $\langle (I^\sigma)^{-1/2}\dot{\Phi}(\sigma), (I^\sigma)^{-1/2} \dot{\Phi}(\sigma)\rangle$ is equal to (41). □

Acknowledgment

The third author's research was benefited by the support of the 'Chaire Risque de crédit', Fédération Bancaire Française.

Appendix

A.1. Sketch of the proof of Proposition 1

The proof is given in Gobet [4] for the uniform discretization scheme and can be easily extend to more general deterministic discretization schemes assuming A1. The expansion of $Z_n^f(h)$ is based on Malliavin calculus and we refer the reader to Nualart [19] for the definitions and notations. We recall that q is the dimension of X and we assume that f takes its values in $\mathbb{R}$. We fix $f \in B$ and $h \in H_0$ and we consider the process $(X^\theta(t))_t$, the solution of

$$X^\theta(t) = x_0 + \int_0^t b(X^\theta(s))ds + \int_0^t a(X^\theta(s), f(s) + \theta h(s)/\sqrt{n})dW(s), \tag{47}$$

where $\theta \in \mathbb{R}$. We omit the dependence in f and we can remark that for $\theta = 0$, the equation (47) is the equation (20) defining X^f.

We denote by $p_n^\theta(t_i^n, t_{i+1}^n, x, y)$ the density of the law of $X^\theta(t_{i+1}^n)$ conditionally on $X^\theta(t_i^n) = x$ and $\dot{p}_n^\theta$ its derivative with respect to θ. We first remark that $Z_n^f(h) = \log \frac{dP_n^{f+h/\sqrt{n}}}{dP_n^f}$ is given by

$$Z_n^f(h) = \sum_{i=0}^{n-1} \int_0^1 \frac{\dot{p}_n^\theta}{p_n^\theta}(t_i^n, t_{i+1}^n, X^f(t_i^n), X^f(t_{i+1}^n))d\theta.$$

Following [4], we have a representation of $\frac{\dot{p}_n^\theta}{p_n^\theta}$ as a conditional expectation of some Malliavin operators that we will explicit in that follows. This representation is based on Malliavin calculus on the time interval $[t_i^n, t_{i+1}^n]$, conditionally on $(W(t))_{t \le t_i^n}$. We first observe that the process $(X^\theta(t))$ admits a derivative with respect to θ that we will denote by $(\dot{X}^\theta(t))$. Moreover $(X^\theta(t))$ and $(\dot{X}^\theta(t))$ belong respectively to the Malliavin spaces $\mathbb{D}^{2,p}$ and $\mathbb{D}^{1,p}$, $\forall p \ge 1$. Now, let φ be a smooth function with compact support, we have from the Lebesgue derivative theorem:

$$\frac{\partial}{\partial\theta}\mathbb{E}_{x,i}\varphi(X^\theta(t_{i+1}^n)) = \int \varphi(y)\dot{p}_n^\theta(t_i^n, t_{i+1}^n, x, y)dy,$$

where $\mathbb{E}_{x,i}$ is the expectation conditionally on $X^\theta(t_i^n) = x$. On the other hand, we have:

$$\frac{\partial}{\partial\theta}\mathbb{E}_{x,i}\varphi(X^\theta(t_{i+1}^n)) = \sum_{j=1}^q \mathbb{E}_{x,i}\varphi'_j(X^\theta(t_{i+1}^n))\dot{X}_j^\theta(t_{i+1}^n).$$

Using the integration by parts formula (see Nualart [19]), we can write

$$\mathbb{E}_{x,i}\varphi'_j(X^\theta(t_{i+1}^n))\dot{X}_j^\theta(t_{i+1}^n) = \mathbb{E}_{x,i}\varphi(X^\theta(t_{i+1}^n))\delta\left(\dot{X}_j^\theta(t_{i+1}^n)\sum_{r=1}^q(\gamma^{-1}_{X^\theta(t_{i+1}^n)})_{j,r}DX_r^\theta(t_{i+1}^n)\right).$$

The operator δ is the divergence operator, $DX^\theta(t_{i+1}^n) = (D^{l'}X_l^\theta(t_{i+1}^n))_{1\le l,l'\le q}$ is the Malliavin derivative of the vector $X^\theta(t_{i+1}^n)$ and $\gamma_{X^\theta(t_{i+1}^n)}$ is the Malliavin variance–covariance matrix of $X^\theta(t_{i+1}^n)$.

This leads to the representation:

$$\frac{\dot{p}_n^\theta}{p_n^\theta}(t_i^n, t_{i+1}^n, x, y) = \mathbb{E}_{x,i}\left(\delta\left(\dot{X}^\theta(t_{i+1}^n)^*\gamma^{-1}_{X^\theta(t_{i+1}^n)}DX^\theta(t_{i+1}^n)\right)|X^\theta(t_{i+1}^n) = y\right). \tag{48}$$

Moreover, we have some explicit expressions for $\dot{X}^\theta(t_{i+1}^n)$, $DX^\theta(t_{i+1}^n)$ and $\gamma_{X^\theta(t_{i+1}^n)}$ (see for example [19]), that we give here for the sake of completeness. We note $Y(t)$ the derivative of the flow of X^θ, $Y(t)$ is a matrix $q \times q$ solution of:

$$Y(t) = Id + \int_0^t b'(X^\theta(s))Y(s)ds + \sum_{j=1}^q \int_0^t a'_j(X^\theta(s), f(s) + \theta h(s)/\sqrt{n})Y(s)dW^j(s), \tag{49}$$

where $b' = (\frac{\partial b_i}{\partial x_j})_{i,j}$ is a $q \times q$ matrix, a_j is the jth column of the matrix a and $a'_j = (\frac{\partial a_{i,j}}{\partial x_k})_{i,k}$ its derivative with respect to x. It is well known that assuming A2, Y is invertible and the pth moments of Y and Y^{-1} are uniformly bounded. Moreover we have for $s \in [t_i^n, t_{i+1}^n]$:

$$D_s X^\theta(t_{i+1}^n) = Y(t_{i+1}^n)Y^{-1}(s)a(X^\theta(s), f(s) + \theta h(s)/\sqrt{n}),$$

$$\gamma_{X^\theta(t_{i+1}^n)} = \int_{t_i^n}^{t_{i+1}^n} D_s X^\theta(t_{i+1}^n) D_s X^\theta(t_{i+1}^n)^* ds,$$

$$\begin{aligned}\dot{X}^\theta(t_{i+1}^n) \;=\; & Y(t_{i+1}^n)Y^{-1}(t_i^n) \int_{t_i^n}^{t_{i+1}^n} (Y(s)Y^{-1}(t_i^n))^{-1} \\ & \times \left[\dot{a}(X^\theta(s), f(s) + \theta h(s)/\sqrt{n}) \frac{h(s)}{\sqrt{n}} dW(s) \right. \\ & - \sum_{j=1}^q \int_{t_i^n}^{t_{i+1}^n} a'_j(X^\theta(s), f(s) + \theta h(s)/\sqrt{n}) \dot{a}_j(X^\theta(s), f(s) \\ & \left. + \theta h(s)/\sqrt{n}) \frac{h(s)}{\sqrt{n}} ds \right].\end{aligned}$$

The main point is that we can approximate the process $\dot{X}^\theta(t_{i+1}^n)^* \gamma^{-1}_{X^\theta(t_{i+1}^n)} D_s X^\theta(t_{i+1}^n)$, for $t_i^n \le s < t_{i+1}^n$ as:

$$\dot{X}^\theta(t_{i+1}^n)^* \gamma^{-1}_{X^\theta(t_{i+1}^n)} D_s X^\theta(t_{i+1}^n) = P_n(s) + U_n(s), \tag{50}$$

where $P_n(s)$ is constant for $s \in [t_i^n, t_{i+1}^n)$

$$P_n(s) = \frac{1}{\sqrt{n}(t_{i+1}^n - t_i^n)} h(t_i^n) \Delta W_{t_i^n}^* (\dot{a}^*(aa^*)^{-1}a)(X^\theta(t_i^n), f(t_i^n) + \theta h(t_i^n)/\sqrt{n}), \tag{51}$$

where $\Delta W_{t_i^n} = W(t_{i+1}^n) - W(t_i^n)$. The process P_n gives the principal contribution and U_n has a negligible contribution. This leads to the decomposition

$$\frac{\dot{p}_n^\theta}{p_n^\theta}(t_i^n, t_{i+1}^n, x, y) = \mathbb{E}_{x,i}\left(\delta(P_n) + \delta(U_n) | X^\theta(t_{i+1}^n) = y\right). \tag{52}$$

We can compute $\delta(P_n)$:

$$\delta(P_n) = \frac{1}{\sqrt{n}(t_{i+1}^n - t_i^n)} h(t_i^n) \left(\Delta W_{t_i^n}^* (a^{-1}\dot{a})^* \Delta W_{t_i^n} - (t_{i+1}^n - t_i^n) Tr(a^{-1}\dot{a})^* \right), \tag{53}$$

where $a^{-1}\dot{a}$ is evaluated at $(X^\theta(t_i^n), f(t_i^n)+\theta h(t_i^n)/\sqrt{n})$. Moreover, we remark that $\mathbb{E}_{x,i}\delta(U_n) = 0$ and that we have the following bounds on U_n

$$\mathbb{E}_{x,i}|\delta(U_n)|^p \leq C\left(\frac{\sup_i \sqrt{t_{i+1}^n - t_i^n} + \sup_i |f(t_{i+1}^n) - f(t_i^n)|}{\sqrt{n}}\right)^p, \quad p \geq 1.$$

We omit the details, but these bounds are sufficient to prove that the contribution of U_n is negligible in the expansion of $Z_n^f(h)$ (see [4] Proposition 4.2).

To compute the conditional expectation of $\delta(P_n)$ on $X^\theta(t_{i+1}^n) = y$, we remark that we can approximate $\Delta W_{t_i^n}$ by $a^{-1}(X^\theta(t_i^n), f(t_i^n) + \theta h(t_i^n)/\sqrt{n})(X^\theta(t_{i+1}^n) - X^\theta(t_i^n))$ and so

$$\mathbb{E}_{x,i}\left(\delta(P_n)|X^\theta(t_{i+1}^n) = y\right) = \frac{1}{\sqrt{n}(t_{i+1}^n - t_i^n)}h(t_i^n)\Big[(y-x)^*(a^{-1})^*(a^{-1}\dot{a})^*a^{-1}(y-x) - (t_{i+1}^n - t_i^n)Tr(a^{-1}\dot{a})^*\Big] + R_{n,i}$$

where now $a^{-1}\dot{a}$ and a^{-1} are evaluated at $(x, f(t_i^n)+\theta h(t_i^n)/\sqrt{n})$, and where $R_{n,i}$ is a remainder term. By a Taylor expansion up to order one with respect to θ, we obtain after some calculus:

$$\begin{aligned}
&\int_0^1 \mathbb{E}_{x,i}\left(\delta(P_n)|X^\theta(t_{i+1}^n) = y\right)d\theta \\
&= \frac{1}{\sqrt{n}(t_{i+1}^n - t_i^n)}h(t_i^n)\Big[(y-x)^*(a^{-1})^*(a^{-1}\dot{a})^*a^{-1}(y-x) - (t_{i+1}^n - t_i^n)Tr(a^{-1}\dot{a})^*\Big] \\
&\quad - \frac{1}{2n}\frac{h(t_i^n)^2}{(t_{i+1}^n - t_i^n)}(y-x)^*(a^{-1})^*\left((a^{-1}\dot{a})^*(a^{-1}\dot{a})^* + (a^{-1}\dot{a})(a^{-1}\dot{a})^*\right)a^{-1}(y-x) \\
&\quad + \tilde{R}_{n,i},
\end{aligned}$$

where now $a^{-1}\dot{a}$ and a^{-1} are evaluated at $(x, f(t_i^n))$.

Finally, replacing (x, y) by $(X^f(t_i^n), X^f(t_{i+1}^n))$, and using the approximation $X^f(t_{i+1}^n) - X^f(t_i^n) = a_{t_i^n}\Delta W_{t_i^n}$, we deduce:

$$\begin{aligned}
Z_n^f(h) &= \frac{1}{\sqrt{n}}\sum_{i=0}^{n-1}\frac{h(t_i^n)}{(t_{i+1}^n - t_i^n)}\left(\Delta W_{t_i^n}^*(a^{-1}\dot{a})_{t_i^n}^*\Delta W_{t_i^n} - (t_{i+1}^n - t_i^n)Tr(a^{-1}\dot{a})_{t_i^n}^*\right) \\
&\quad - \frac{1}{2n}\sum_{i=0}^{n-1}\frac{h(t_i^n)^2}{(t_{i+1}^n - t_i^n)}\Delta W_{t_i^n}^*\left((a^{-1}\dot{a})_{t_i^n}^*(a^{-1}\dot{a})_{t_i^n}^* + (a^{-1}\dot{a})_{t_i^n}(a^{-1}\dot{a})_{t_i^n}^*\right)\Delta W_{t_i^n} \\
&\quad + o_{P_n^f}(1),
\end{aligned} \tag{54}$$

where $(a^{-1}\dot{a})_{t_i^n} = (a^{-1}\dot{a})(X^f(t_i^n), f(t_i^n))$. The second term in (54) converges in P_n^f-probability to

$$-\frac{1}{2}\langle h, I^f h\rangle = -\frac{1}{2}\int_0^1 h^2(s)Tr((a^{-1}\dot{a})_s^*(a^{-1}\dot{a})_s^* + (a^{-1}\dot{a})_s(a^{-1}\dot{a})_s^*)\mu_0(s)ds.$$

Moreover, we have

$$\langle h, I^f h\rangle = \langle h, I_n^f h\rangle + o_{P_n^f}(1)$$

where $I_n^f h$ is defined by (22). This leads to

$$Z_n^f(h) = N_n^f(h) - \frac{1}{2}\langle h, I_n^f h\rangle + o_{p_n^f}(1), \tag{55}$$

with $N_n^f(h)$ given by (21). We conclude by establishing the stable convergence in law of $N_n^f(h)$ using a central limit theorem for triangular arrays of random variables (see Jacod [8], Genon-Catalot and Jacod [2]).

References

[1] S. Geiss, A. Toivola, Weak convergence of error processes in discretizations of stochastic integrals and besov spaces, Bernoulli 15 (2009) 925–954.

[2] Valentine Genon-Catalot, Jean Jacod, On the estimation of the diffusion coefficient for multi-dimensional diffusion processes, Ann. Inst. H. Poincaré Probab. Stat. 29 (1) (1993) 119–151.

[3] E. Gobet, E. Temam, Discrete time hedging errors for options with irregular pay-offs, Finance Stoch. 5 (3) (2001) 357–367.

[4] Emmanuel Gobet, Local asymptotic mixed normality property for elliptic diffusion: a Malliavin calculus approach, Bernoulli 7 (6) (2001) 899–912.

[5] G.K. Golubev, Local asymptotic normality in problems of non parametric estimation of functions, and lower bounds for quadratic risks, Theory Probab. Appl. 1 (1992) 152–157.

[6] Takaki Hayashi, Jean Jacod, Nakahiro Yoshida, Irregular sampling and central limit theorems for power variations: the continuous case, Ann. Inst. H. Poincaré Probab. Stat. 47 (4) (2011) 1197–1218.

[7] I.A. Ibragimov, R.Z. Khas'minskiĭ, Asymptotically normal families of distributions and efficient estimation, Ann. Statist. 19 (4) (1991) 1681–1724.

[8] Jean Jacod, On continuous conditional Gaussian martingales and stable convergence in law, in: Séminaire de Probabilités, XXXI, in: Lecture Notes in Math., vol. 1655, Springer, Berlin, 1997, pp. 232–246.

[9] Jean Jacod, Asymptotic properties of realized power variations and related functionals of semimartingales, Stochastic Process. Appl. 118 (4) (2008) 517–559.

[10] Jean Jacod, Markus Reiß, A remark on the rates of convergence for integrated volatility estimation in the presence of jumps. arXiv:1209.4173v1 [math.ST], 2012.

[11] Jean Jacod, Mathieu Rosenbaum, Quarticity and other functionals of volatility: efficient estimation. arXiv:1207.3757v2 [math.ST], 2012.

[12] Jean Jacod, Philip Protter, Discretization of Processes, in: Stochastic Modelling and Applied Probability, vol. 67, Springer, Heidelberg, 2012.

[13] Jean Jacod, Philip Protter, Asymptotic error distributions for the Euler method for stochastic differential equations, Ann. Probab. 26 (1) (1998) 267–307.

[14] P. Jeganathan, On a decomposition of the limit distribution of a sequence of estimators, Sankhyā A 43 (1) (1981) 26–36.

[15] P. Jeganathan, On the asymptotic theory of estimation when the limit of the log-likelihood ratios is mixed normal, Sankhyā A 44 (2) (1982) 173–212.

[16] P. Jeganathan, Some asymptotic properties of risk functions when the limit of the experiment is mixed normal, Sankhyā A 45 (1) (1983) 66–87.

[17] Lucien Le Cam, Grace Lo Yang, Asymptotics in Statistics, in: Springer Series in Statistics, Springer-Verlag, New York, 1990, Some basic concepts.

[18] P.W. Millar, Nonparametric applications of an infinite-dimensional convolution theorem, Z. Wahrsch. Verw. Gebiete 68 (4) (1985) 545–556.

[19] David Nualart, The Malliavin calculus and related topics, in: Probability and its Applications (New York), second ed., Springer-Verlag, Berlin, 2006.

[20] H. Rootzen, Limit distributions for the error in approximations of stochastic integral, Ann. Probab. 80 (2) (1980) 241–251.

[21] A.W. van der Vaart, Asymptotic Statistics, in: Cambridge Series in Statistical and Probabilistic Mathematics, vol. 3, Cambridge University Press, Cambridge, 1998.