Bernoulli **7**(6), 2001, 899–912

Local asymptotic mixed normality property for elliptic diffusion: a Malliavin calculus approach

EMMANUEL GOBET

Centre de Mathématiques Appliquées, École Polytechnique, 91128 Palaiseau Cedex, France. E-mail: emmanuel.gobet@polytechnique.fr

We address the problem of the validity of the local asymptotic mixed normality (LAMN) property when the model is a multidimensional diffusion process X whose coefficients depend on a scalar parameter θ: the sample $(X_{k/n})_{0\leqslant k\leqslant n}$ corresponds to an observation of X at equidistant times in the interval [0, 1]. We prove that the LAMN property holds true for the likelihood under an ellipticity condition and some suitable smoothness assumptions on the coefficients of the stochastic differential equation. Our method is based on Malliavin calculus techniques: in particular, we derive for the log-likelihood ratio a tractable representation involving conditional expectations.

Keywords: conditional expectation; convergence of sums of random variables; diffusion process; local asymptotic mixed normality property; log-likelihood ratios; Malliavin calculus; parametric estimation

1. Introduction

Let P^θ be the law of the $\mathbb{R}^d$-valued diffusion process

$$X_t^\theta = x + \int_0^t b(\theta, s, X_s^\theta)\mathrm{d}s + \int_0^t S(\theta, s, X_s^\theta)\mathrm{d}B_s \tag{1.1}$$

for $t \in [0, 1]$, where B is a d-dimensional Brownian motion, x is fixed, and b and S are known smooth functions of (θ, t, x). The parameter θ is a scalar parameter which belongs to Θ, an open interval of $\mathbb{R}$. In this paper, we focus on the case where S is non-degenerate. We are interested in an estimation problem where we observe X at n regularly spaced times $t_k = k/n$ on the time interval [0, 1]: asymptotics are taken when n goes to $+\infty$. In this setting, exhibiting suitable contrasts, Genon-Catalot and Jacod (1993) construct consistent estimators $\hat{\theta}_n$ of θ_0. Furthermore, they prove the weak convergence at rate $\sqrt{n}$ of their renormalized error $\sqrt{n}(\hat{\theta}_n - \theta_0)$ to a mixed Gaussian variable. Another interesting issue is whether these estimators are asymptotically efficient: in some way, this is related to the local asymptotic mixed normality (LAMN) property, which we now recall (see, for example, Le Cam and Lo Yang 1990, Chapter 5). If $\mathcal{F}_n = \sigma(X_{t_k}: 0 \leqslant k \leqslant n)$, we denote the restriction of P^θ to $\mathcal{F}_n$ by P_n^θ, and the log-likelihood ratio of P_n^θ with respect to $P_n^{\theta_0}$ by $Z_n(\theta_0, \theta)$. The sequence $((\mathbb{R}^d)^n, \mathcal{F}_n, (P_n^\theta)_{\theta\in\Theta})$ of statistical models has the LAMN property for the likelihood, at θ_0, at rate $\sqrt{n}$ and conditional variance $\Gamma(\theta_0) > 0$ if

1350–7265

$$Z_n\left(\theta_0, \theta_0 + \frac{u}{\sqrt{n}}\right) = u\Delta_n\sqrt{\Gamma(\theta_0)} - \frac{u^2}{2}\Gamma(\theta_0) + R_n$$

with $R_n \overset{P^{\theta_0}}{\to} 0$ and $(\Delta_n, \Gamma(\theta_0)) \overset{\mathcal{L}(P^{\theta_0})}{\to} (\Delta, \Gamma(\theta_0))$, where $\Delta \sim N(0, 1)$ is a Gaussian variable independent of $\Gamma(\theta_0)$. When the LAMN property holds true for the likelihood, one can apply minimax theorems (see Jeganathan 1982; 1983) and derive, in particular, lower bounds for the variance of estimators.

In this paper, we intend to prove the validity of the LAMN property for the likelihood at rate $\sqrt{n}$, for the model (1.1), if the diffusion coefficient S is non-degenerate, under suitable smoothness assumptions on b and S. The result is new in a multidimensional setting, even if it is not surprising, since the estimators exhibited by Genon-Catalot and Jacod (1993) satisfy the LAMN condition. We observe that the one-dimensional case has been considered by Donhal (1987): its proof relies on an expansion of the transition probability density p^θ of X. Unfortunately, in higher dimensions (except for some specific cases – see Genon-Catalot and Jacod 1994), the well-known expansion of p^θ (see Azencott 1984) is not sufficient to adapt Donhal's proof.

To obtain our result we adopt a new strategy. The first step consists in transforming the log-likelihood ratio $Z_n(\theta_0, \theta_0 + u/\sqrt{n})$ using a Malliavin calculus integration-by-parts formula: we derive a representation of Z_n as a sum of conditional expectations (see Proposition 4.1). The second step is to get an appropriate convergence result to analyse the weak convergence of this kind of sum: in Corollary 4.1, we give simple conditions to achieve this purpose. Finally, simple expansions of the conditioned random variables yield the result.

The Malliavin calculus approach we develop here is quite general and seems well suited to the study of the likelihood: the case of degenerate coefficient diffusion (with one hypoellipticity condition) may be treated in the same way. Furthermore, presumably, this approach may also enable non-Markovian situations, such as hidden diffusions or stochastic differential equations with memory, to be tackled: consideration of these is deferred to forthcoming papers.

The present paper is organized as follows. In Section 3, we briefly introduce the material required for our Malliavin calculus computations. Section 4 is devoted to the proof of the result: the main steps are the representation of $Z_n(\theta_0, \theta_0 + u/\sqrt{n})$ involving conditional expectations (Proposition 4.1) and the weak convergence of expectations of this kind (Corollary 4.1); a technical result is proved in Section 5.

2. Assumptions and results

Let Θ be an open interval of $\mathbb{R}$. We consider a map b (a map S) from $\Theta \times [0, 1] \times \mathbb{R}^d$ into $\mathbb{R}^d$ (into $\mathbb{R}^d \otimes \mathbb{R}^d$). As usual, derivation with respect to θ (with respect to space variables) is denoted by a dot (by a prime). We assume that the following two hypotheses are fulfilled.

Assumption R. *The functions $b(\theta, t, x)$ and $S(\theta, t, x)$ are of class $C^{1+\alpha}$ with respect to θ*

($\alpha > 0$). *The functions* b, S, $\dot{b}$, $\dot{S}$, b' *and* S' *are of class* $C^{1,2}$ *with respect to* (t, x). *Moreover,* b' *and* S' *are uniformly bounded on* $\Theta \times [0, 1] \times \mathbb{R}^d$.

Assumption E. *The matrix* S *is symmetric and positive definite:*

$$\forall(\theta, t, x) \in \Theta \times [0, 1] \times \mathbb{R}^d: \qquad \inf_{\xi \in \mathbb{R}^d : \|\xi\|=1} \xi.S(\theta, t, x)\xi > 0.$$

Let $(B_t)_{t \geqslant 0}$ be a standard Brownian motion in $\mathbb{R}^d$ (with $(\mathcal{G}_t)_{t\geqslant 0}$ its usual filtration), and $(X_t^\theta)_{t\geqslant 0}$ be the inhomogeneous diffusion process which solves

$$X_t^\theta = x + \int_0^t b(\theta, s, X_s^\theta)\mathrm{d}s + \int_0^t S(\theta, s, X_s^\theta)\mathrm{d}B_s. \tag{2.2}$$

In the following, $\mathrm{E}_{s,x}^\theta$ denotes the expectation under the law of the diffusion X^θ starting at x at time s.

Remark. Our study is restricted to a scalar parameter; however, there are no additional technical difficulties in dealing with multidimensional parameters, beyond a more cumbersome notation. Here, the true difficulty comes from the fact that the process takes its values in $\mathbb{R}^d$.

Fix $\theta_0 \in \Theta$. For $n \in \mathbb{N}^*$, we now consider the sample $(X_{t_k})_{0 \leqslant k \leqslant n}$ of the diffusion X observed at equidistant discretization times $t_k = k/n$ on the interval [0, 1]. For $u \in \mathbb{R}$, we introduce the log-likelihood ratio

$$Z_n\left(\theta_0, \theta_0 + \frac{u}{\sqrt{n}}\right) := \log\left(\frac{dP_n^{\theta_0 + u/\sqrt{n}}}{dP_n^{\theta_0}}\right)(X_0, X_{1/n}, \ldots, X_1). \tag{2.3}$$

The main result of the paper is the following theorem.

Theorem 2.1. *Under Assumptions R and E, the LAMN property holds for the likelihood in* θ_0, *that is, there is an extra Gaussian variable* $\Delta \sim N(0, 1)$ *independent of* $\mathcal{G}_1$ *such that*

$$Z_n\left(\theta_0, \theta_0 + \frac{u}{\sqrt{n}}\right) \overset{\mathcal{L}(P^{\theta_0})}{\rightarrow} u\sqrt{\Gamma(\theta_0)}\Delta - \frac{u^2}{2}\Gamma(\theta_0),$$

where $\Gamma(\theta_0) = 2\int_0^1 \mathrm{tr}(\dot{S}S^{-1})^2(\theta_0, t, X_t^{\theta_0})\mathrm{d}t$.

The remainder of the paper is devoted to its proof: the first step of our approach consists in transforming the log-likelihood ratio using Malliavin calculus techniques, to obtain a simple and tractable representation of the ratio $Z_n(\theta_0, \theta_0 + u/\sqrt{n})$, as a conditional expectation. We first introduce the material required for these computations.

3. Some basic results on the Malliavin calculus

The reader may refer to Nualart (1995) for a detailed exposition of this section.

Fix a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ and let $(W_t)_{t\geqslant 0}$ be a d-dimensional Brownian motion. Fix $T \in (0, 1]$. For $h(\cdot) \in H = L_2([0, T], \mathbb{R}^d)$, $W(h)$ is the Wiener stochastic integral $\int_0^T h(t)\cdot \mathrm{d}W_t$. Let $\mathcal{S}$ denote the class of random variables of the form $F = f(W(h_1), \ldots, W(h_N))$ where $f \in C_p^\infty(\mathbb{R}^N)$, $(h_1, \ldots, h_N) \in H^N$ and $N \geqslant 1$. For $F \in \mathcal{S}$, we define its derivative $\mathcal{D}F = (\mathcal{D}_t F)_{t\in[0,T]}$ as the H-valued random variable given by

$$\mathcal{D}_t F = \sum_{i=1}^N \partial_{x_i} f(W(h_1), \ldots, W(h_N))h_i(t).$$

The operator $\mathcal{D}$ is closable as an operator from $L_p(\Omega)$ to $L_p(\Omega, H)$, for any $p \geqslant 1$. Its domain is denoted by $\mathbb{D}^{1,p}$ with respect to the norm $\|F\|_{1,p} = [\mathrm{E}|F|^p + \mathrm{E}(\|\mathcal{D}F\|_H^p)]^{1/p}$. We can define the iteration of the operator $\mathcal{D}$ in such a way that for a smooth random variable F, the derivative $\mathcal{D}^k F$ is a random variable with values on $H^{\otimes k}$. As in the case $k = 1$, the operator $\mathcal{D}^k$ is closable from $S \subset L_p(\Omega)$ into $L_p(\Omega; H^{\otimes k})$, $p \geqslant 1$. If we define the norm $\|F\|_{k,p} = [\mathrm{E}|F|^p + \sum_{j=1}^k \mathrm{E}(\|\mathcal{D}^j F\|_{H^{\otimes j}}^p)]^{1/p}$, we denote its domain by $\mathbb{D}^{k,p}$. We will require the chain rule property:

Proposition 3.1. *Fix* $p \geqslant 1$. *If* $f \in C_b^1(\mathbb{R}^d, \mathbb{R})$ *and* $F = (F_1, \ldots, F_d)$ *a random vector whose components belong to* $\mathbb{D}^{1,p}$, *then* $f(F) \in \mathbb{D}^{1,p}$ *and for* $t \geqslant 0$, *we have*

$$\mathcal{D}_t(f(F)) = \sum_{i=1}^d \partial_{x_i} f(F)\mathcal{D}_t F_i.$$

We now introduce δ, the Skorohod integral, defined as the adjoint operator of $\mathcal{D}$:

Definition 3.1. δ *is a linear operator on* $L_2([0, T] \times \Omega, \mathbb{R}^d)$ *with values in* $L_2(\Omega)$ *such that:*

(i) *the domain of* δ (*denoted by* Dom(δ)) *is the set of processes* $u \in L_2([0, T] \times \Omega, \mathbb{R}^d)$ *such that*

$$\forall F \in \mathbb{D}^{1,2}, \qquad \left|\mathrm{E}\left(\int_0^T \mathcal{D}_t F \cdot u_t \,\mathrm{d}t\right)\right| \leqslant c(u)\|F\|_2.$$

(ii) *if* u *belongs to* Dom(δ), *then* $\delta(u) = \int_0^T u_t \delta W_t$ *is the element of* $L_2(\Omega)$ *characterized by the integration by parts formula*

$$\forall F \in \mathbb{D}^{1,2}, \qquad \mathrm{E}(F\delta(u)) = \mathrm{E}\left(\int_0^T \mathcal{D}_t F \cdot u_t \,\mathrm{d}t\right). \tag{3.4}$$

In the following proposition, we summarize some of the properties of the Skorohod integral:

Proposition 3.2.

(i) *The space of weakly differentiable H-valued variables $\mathbb{D}^{1,2}(H)$ belongs to* $\mathrm{Dom}(\delta)$.

(ii) *If u is an adapted process belonging to $L_2([0, T] \times \Omega, \mathbb{R}^d)$, then the Skorohod integral and the Itô integral coincide:*

$$\delta(u) = \int_0^T u_t \delta W_t = \int_0^T u_t \, \mathrm{d}W_t. \tag{3.5}$$

(iii) *If F belongs to $\mathbb{D}^{1,2}$, then for any $u \in \mathrm{Dom}(\delta)$ such that $\mathrm{E}(F^2 \int_0^T u_t^2 \, \mathrm{d}t) < +\infty$, we have*

$$\delta(Fu) = F\delta(u) - \int_0^T \mathcal{D}_t F \cdot u_t \, \mathrm{d}t, \tag{3.6}$$

whenever the right-hand side belongs to $L_2(\Omega)$.

(iv) *For u belonging to $\mathbb{D}^{2,2}(H)$, $\mathcal{D}$ and δ satisfy the following commutativity relationship:*

$$\mathcal{D}_t(\delta(u)) = u_t + \int_0^T \mathcal{D}_t(u_s)\delta W_s. \tag{3.7}$$

(v) *The operator δ is continuous from $\mathbb{D}^{k,p}(H)$ into $\mathbb{D}^{k-1,p}$ for all $k \geqslant 1$ and $p > 1$. In particular, if $k = 1$, for $p > 1$, we have*

$$\|\delta(u)\|_p \leqslant c_p(\|u\|_{L_p(\Omega,\, H)} + \|\mathcal{D}u\|_{L_p(\Omega, H \otimes H)}). \tag{3.8}$$

Finally, we recall the Clark–Ocone formula.

Proposition 3.3. *Any random variable $F \in \mathbb{D}^{1,2}$ has the integral representation*

$$F = \mathrm{E}(F) + \int_0^T \mathrm{E}(\mathcal{D}_t F | \mathcal{F}_t) \cdot \mathrm{d}W_t.$$

4. Proof of the LAMN property

4.1. Localization of the assumptions

First, let us justify the fact that it is sufficient to prove Theorem 2.1 under the following stronger assumptions.

Assumption R′. *Assumption R holds, and the functions b, S, $\dot{b}$, $\dot{S}$, b' and S' (of class $C^{1,2}$ with respect to (t, x)) and their derivatives are uniformly bounded on $\Theta \times [0, 1] \times \mathbb{R}^d$.*

Assumption E′. *The symmetric matrix S satisfies a uniform ellipticity condition*

$$\forall(\theta,\, t,\, x) \in \Theta \times [0, 1] \times \mathbb{R}^d, \qquad \mu_{\min} \mathrm{I}_d \leqslant S^2(\theta,\, t,\, x) \leqslant \mu_{\max} \mathrm{I}_d$$

for some real numbers $0 < \mu_{\min} \leqslant \mu_{\max} < +\infty$.

Lemma 4.1. *If the result of Theorem* 2.1 holds true under Assumptions R′ and E′, it also holds true under Assumptions R and E.

Proof. This is more or less standard and we only give the main arguments (see Delattre 1997). For q integer, consider two functions $b_q : \Theta \times [0, 1] \times \mathbb{R}^d \mapsto \mathbb{R}^d$ and $\sigma_q : \Theta \times [0, 1] \times \mathbb{R}^d \mapsto \mathbb{R}^d \otimes \mathbb{R}^d$, which, on the one hand, coincide with b and σ on the set $\Theta \times [0, 1] \times \{z : |z| \leqslant q\}$, and, on the other hand, satisfy Assumptions R′ and E′. Denote by P_q^θ the law of the stochastic differential equation with the coefficients b_q and σ_q, and by $P_{q,n}^\theta$ its restriction to $\mathcal{F}_n$. Denote by $\|\mu\|$ the total variation of the signed measure μ and put $\Delta_{q,n}(\theta) = \|P_n^\theta - P_{q,n}^\theta\|$: it is straightforward to prove that for any bounded sequence $(\theta_n)_n$, one has $\lim_q \limsup_n \Delta_{q,n}(\theta_n) = 0$. Moreover, Proposition 4.3.2 in Le Cam (1986) yields

$$\int \left(1 \wedge \left| \frac{dP_n^{\theta_0+u/\sqrt{n}}}{dP_n^{\theta_0}} - \frac{dP_{q,n}^{\theta_0+u/\sqrt{n}}}{dP_{q,n}^{\theta_0}} \right| \right) \mathrm{d}(P_n^{\theta_0} + P_{q,n}^{\theta_0})$$
$$\leqslant \Delta_{q,n}(\theta_0) + 2\Delta_{q,n}\left(\theta_0 + \frac{u}{\sqrt{n}}\right) + 2\sqrt{\Delta_{q,n}(\theta_0)}.$$

The proof of the lemma is now easily completed: we omit the details. □

As a consequence of the above lemma, we can consider for the rest of the proof of Theorem 2.1 Assumptions R′ and E′, instead of R and E.

4.2. Transformation of $Z_n(\theta_0, \theta_0 + u/\sqrt{n})$ using Malliavin calculus

Under Assumptions R′ and E′, the law of X_t^θ conditionally on $X_s^\theta = x$ $(t > s)$ has a strictly positive transition density $p^\theta(s, t, x, y)$, which is smooth with respect to θ (see Proposition 5.1). Thus, the Markov property enables us to write

$$Z_n\left(\theta_0, \theta_0 + \frac{u}{\sqrt{n}}\right) = \sum_{k=0}^{n-1} \int_{\theta_0}^{\theta_0+u/\sqrt{n}} \frac{\dot{p}^\theta}{p^\theta}(t_k, t_{k+1}, X_{t_k}, X_{t_{k+1}})\mathrm{d}\theta. \tag{4.9}$$

We now derive a new expression for the term being integrated in (4.9), as a conditional expectation, using Malliavin calculus. For this purpose, let us consider, throughout this subsection, the solution of (2.2) starting at x at time t_k, that is, the $\mathbb{R}^d$-valued process denoted by $(X_{t_k+t}^\theta)_{t\geqslant 0}$ solving

$$X_{t_k+t}^\theta = x + \int_0^t b(\theta, t_k + s, X_{t_k+s}^\theta)\mathrm{d}s + \sum_{j=1}^d \int_0^t S_j(\theta, t_k + s, X_{t_k+s}^\theta)\mathrm{d}W_{j,s}, \tag{4.10}$$

where S_j is the jth column vector of S, and $(W_t)_{t\geqslant 0}$ a new Brownian motion with its usual filtration $(\mathcal{F}_t)_{t\geqslant 0}$ (W corresponds to the shift of B at time t_k: although it depends on k, we suppress this dependence in the notation because there is no possible cause for confusion).

We associate with $(X^\theta_{t_k+t})_{t\geqslant 0}$ its flow, that is, the Jacobian matrix $Y^\theta_t := \nabla_x X^\theta_{t_k+t}$, and its derivative with respect to θ denoted by $\dot{X}^\theta_t$.

Remark. Our notation with $X^\theta_{t_k+t}$, Y^θ_t and $\dot{X}^\theta_t$ is not homogeneous with respect to the time variable: denoting $X^\theta_{t_k+t}$ by X^θ_t would have been convenient at this stage of the proof, but nevertheless, the notation with $X^\theta_{t_k+t}$ will be clearer for the purpose of the following computations.

Under Assumption R′, it is clear (see Kunita 1984) that Y^θ_t and $\dot{X}^\theta_t$ solve

$$Y^\theta_t = \mathrm{I}_d + \int_0^t b'(\theta,\ t_k+s,\ X^\theta_{t_k+s})Y^\theta_s\,\mathrm{d}s + \sum_{j=1}^d \int_0^t S'_j(\theta,\ t_k+s,\ X^\theta_{t_k+s})Y^\theta_s\,\mathrm{d}W_{j,s},$$

$$\dot{X}^\theta_t = \int_0^t (\dot{b}(\theta,\ t_k+s,\ X^\theta_{t_k+s}) + b'(\theta,\ t_k+s,\ X^\theta_{t_k+s})\dot{X}^\theta_s)\mathrm{d}s$$
$$+ \sum_{j=1}^d \int_0^t (\dot{S}_j(\theta,\ t_k+s,\ X^\theta_{t_k+s}) + S'_j(\theta,\ t_k+s,\ X^\theta_{t_k+s})\dot{X}^\theta_s)\mathrm{d}W_{j,s}. \tag{4.11}$$

For any $t\geqslant 0$, the random variables $X^\theta_{t_k+t}$, Y^θ_t and $\dot{X}^\theta_t$ are weakly differentiable (see Nualart 1995, Section 2.2): actually, we have $X^\theta_{t_k+t}\in\bigcap_{p\geqslant 1}\mathbb{D}^{3,p}$, $Y^\theta_t\in\bigcap_{p\geqslant 1}\mathbb{D}^{2,p}$, $\dot{X}^\theta_t\in\bigcap_{p\geqslant 1}\mathbb{D}^{2,p}$, with the following estimates

$$\text{for } j=1,\ 2,\ 3,\quad \sup_{r_1,\ldots,r_j\in[0,T]} \mathrm{E}_{t_k,x}(\sup_{r_1\vee\ldots\vee r_j\leqslant t\leqslant T} \|\mathcal{D}_{r_1,\ldots,r_j}X^\theta_{t_k+t}\|^p) \leqslant c, \tag{4.12}$$

$$\text{for } j=1,\ 2,\quad \sup_{r_1,\ldots,r_j\in[0,T]} \mathrm{E}_{t_k,x}(\sup_{r_1\vee\ldots\vee r_j\leqslant t\leqslant T} \|\mathcal{D}_{r_1,\ldots,r_j}Y^\theta_t\|^p) \leqslant c, \tag{4.13}$$

$$\text{for } j=1,\ 2,\quad \sup_{r_1,\ldots,r_j\in[0,T]} \mathrm{E}_{t_k,x}(\sup_{r_1\vee\ldots\vee r_j\leqslant t\leqslant T} \|\mathcal{D}_{r_1,\ldots,r_j}\dot{X}^\theta_t\|^p) \leqslant c, \tag{4.14}$$

for some constant c (uniform in x, k, θ and $T\leqslant 1$). Finally, $\mathcal{D}_s X^\theta_{t_k+t}$ is given by:

$$\mathcal{D}_s X^\theta_{t_k+t} = Y^\theta_t (Y^\theta_s)^{-1} S(\theta,\ t_k+s,\ X^\theta_{t_k+s})\mathbf{1}_{s\leqslant t}. \tag{4.15}$$

Proposition 4.1. *Suppose that Assumptions R′ and E′ hold. Set $T>0$. For $1\leqslant i\leqslant d$, let us define $u^k_i = (u^k_{i,s})_{0\leqslant s\leqslant T}$, the $\mathbb{R}^d$-valued process whose jth component is equal to $u^k_{ij,s} = (S^{-1}(\theta,\ t_k+s,\ X^\theta_{t_k+s})Y^\theta_s(Y^\theta_T)^{-1})_{i,j}$. Then*

$$\frac{\dot{p}^\theta}{p^\theta}(t_k,\ t_k+T,\ x,\ y) = \sum_{i=1}^d \frac{1}{T}\mathrm{E}^\theta_{t_k,x}[\delta(\dot{X}^\theta_{i,T}u^k_i)|X^\theta_{t_k+T}=y]. \tag{4.16}$$

Proof. Let f and g be two smooth functions with compact support. Then

$$\int_\Theta d\theta g(\theta)\mathrm{E}^\theta_{t_k,x}[\nabla f(X^\theta_{t_k+T})\cdot \dot{X}^\theta_T] = -\int_\Theta d\theta g'(\theta)\mathrm{E}^\theta_{t_k,x}[f(X^\theta_{t_k+T})]$$
$$= -\int_\Theta d\theta g'(\theta)\int_{\mathbb{R}^d} dy p^\theta(t_k,\ t_k+T,\ x,\ y)f(y)$$
$$= \int_\Theta d\theta g(\theta)\int_{\mathbb{R}^d} dy \dot{p}^\theta(t_k,\ t_k+T,\ x,\ y)f(y),$$

where we have used a simple integration by parts in two different ways. It remains to prove that

$$\mathrm{E}^\theta_{t_k,x}[\nabla f(X^\theta_{t_k+T})\cdot \dot{X}^\theta_T] = \sum_{i=1}^d \frac{1}{T}\mathrm{E}^\theta_{t_k,x}[f(X^\theta_{t_k+T})\delta(\dot{X}^\theta_{i,T}u^k_i)]. \tag{4.17}$$

Indeed, the proof of Proposition 4.1 can now easily be completed by comparing both expressions obtained for $-\int_\Theta d\theta g'(\theta)\mathrm{E}^\theta_{t_k,x}[f(X^\theta_{t_k+T})]$.

The derivation of the formula (4.17) for $\mathrm{E}^\theta_{t_k,x}[\nabla f(X^\theta_{t_k+T})\cdot \dot{X}^\theta_T]$ is based on the duality relationship (3.4) between $\mathcal{D}$ and δ. First, the chain rule (Proposition 3.1) leads to $\mathcal{D}_s(f(X^\theta_{t_k+T})) = \mathcal{D}_s X^\theta_{t_k+T}\nabla f(X^\theta_{t_k+T})$: for $s \leqslant T$, $\mathcal{D}_s X^\theta_{t_k+T} = Y^\theta_T(Y^\theta_s)^{-1}S(\theta,\ t_k+s,\ X^\theta_{t_k+s})$ is invertible, so that $\partial_{x_i} f(X^\theta_{t_k+T}) = \mathcal{D}_s(f(X^\theta_{t_k+T}))\cdot u^k_{i,s}$. Then, it follows that the left-hand side of (4.17) is equal to

$$\sum_{i=1}^d \frac{1}{T}\mathrm{E}^\theta_{t_k,x}\left[\int_0^T \partial_{x_i} f(X^\theta_{t_k+T})\dot{X}^\theta_{i,T}\, ds\right] = \sum_{i=1}^d \frac{1}{T}\mathrm{E}^\theta_{t_k,x}\left[\int_0^T \mathcal{D}_s(f(X^\theta_{t_k+T}))\cdot(\dot{X}^\theta_{i,T}u^k_{i,s})ds\right]$$
$$= \sum_{i=1}^d \frac{1}{T}\mathrm{E}^\theta_{t_k,x}[f(X^\theta_{t_k+T})\delta(\dot{X}^\theta_{i,T}u^k_i)]$$

by (3.4). This completes the proof. □

4.3. On the convergence of a sum of conditional expectations

Owing to Proposition 4.1 and equality (4.9), $Z_n(\theta_0,\ \theta_0 + u/\sqrt{n})$ is represented as a sum of conditional expectations. To analyse its convergence, we need an appropriate convergence result (Corollary 4.1 below). To prove it, we first state an intermediate result, the proof of which is postponed until Section 5.

Proposition 4.2. *Suppose that Assumptions R' and E' hold. Fix $T > 0$. Let us consider H, an $\mathcal{F}_T$-measurable random variable. For any $\theta \in \Theta$ and any $\alpha > \mu_{\max}/\mu_{\min}$,*

$$\mathrm{E}^{\theta_0}_{t_k,x}\mathrm{E}^\theta_{t_k,x}[|H|\,|X^\theta_{t_k+T} = X^{\theta_0}_{t_k+T}] \leqslant c(\mathrm{E}^\theta_{t_k,x}|H|^\alpha)^{1/\alpha}, \tag{4.18}$$

$$|\mathrm{E}^{\theta_0}_{t_k,x}(\mathrm{E}^\theta_{t_k,x}[H|X^\theta_{t_k+T} = X^{\theta_0}_{t_k+T}]) - \mathrm{E}^\theta_{t_k,x}[H]| \leqslant c|\theta-\theta_0|(\mathrm{E}^\theta_{t_k,x}|H|^\alpha)^{1/\alpha}, \tag{4.19}$$

for some constant c uniform in x, k, θ and $T \leqslant 1$.

The next result is our basic tool in the analysis of the convergence of the sum of conditional expectations.

Corollary 4.1. *Let* $(H_{t_k})_{0\leqslant k\leqslant n-1}$ *be* $\mathcal{F}_{1/n}$*-measurable random variables which satisfy, for some* $\alpha > \mu_{\max}/\mu_{\min}$, *the conditions*

$$\mathrm{E}^{\theta}_{t_k,x}[H_{t_k}] = O(n^{-2}) \quad \text{and} \quad (\mathrm{E}^{\theta}_{t_k,x}|H_{t_k}|^{2\alpha})^{1/2\alpha} = O(n^{-3/2})$$

uniformly in x, k and θ. Then, under Assumptions R′ and E′,

$$\sum_{k=0}^{n-1}\int_{\theta_0}^{\theta_0+u/\sqrt{n}} n\mathrm{E}^{\theta}_{t_k,X_{t_k}}[H_{t_k}|X^{\theta}_{t_{k+1}} = X_{t_{k+1}}]\,\mathrm{d}\theta \overset{P^{\theta_0}}{\rightarrow} 0.$$

Proof. Set $\xi^n_k = \int_{\theta_0}^{\theta_0+u/\sqrt{n}} n\mathrm{E}^{\theta}_{t_k,X_{t_k}}[H_{t_k}|X^{\theta}_{t_{k+1}} = X_{t_{k+1}}]\mathrm{d}\theta$: these are $\mathcal{G}_{t_{k+1}}$-measurable random variables. Using Proposition 4.2 with $T = n^{-1}$ and the conditions of the statement of Corollary 4.1, it is easy to check that $\mathrm{E}^{\theta_0}[\xi^n_k|\mathcal{G}_{t_k}] = O(n^{-3/2})$ and $\mathrm{E}^{\theta_0}[(\xi^n_k)^2|\mathcal{G}_{t_k}] = O(n^{-2})$, uniformly in k. Application of the following classical convergence result (Genon-Catalot and Jacod 1993, Lemma 9) on triangular arrays of random variables completes the proof. □

Lemma 4.2. *Let* ξ^n_k, *U be random variables, with* ξ^n_k *being* $\mathcal{G}_{t_{k+1}}$*-measurable. The two following conditions imply* $\sum_{k=0}^{n-1}\xi^n_k \overset{P}{\rightarrow} U$:

$$\sum_{k=0}^{n-1}\mathrm{E}[\xi^n_k|\mathcal{G}_{t_k}] \overset{P}{\rightarrow} U \quad \text{and} \quad \sum_{k=0}^{n-1}\mathrm{E}[(\xi^n_k)^2|\mathcal{G}_{t_k}] \overset{P}{\rightarrow} 0.$$

4.4. Convergence of $Z_n(\theta_0, \theta_0 + u/\sqrt{n})$ under P^{θ_0}

From (4.9) and Proposition 4.1, we deduce that

$$Z_n(\theta_0, \theta_0 + u/\sqrt{n}) = \sum_{i=1}^{d}\sum_{k=0}^{n-1}\int_{\theta_0}^{\theta_0+u/\sqrt{n}} n\mathrm{E}^{\theta}_{t_k,X_{t_k}}[\delta(\dot{X}^{\theta}_{i,1/n}\ u^k_i)|X^{\theta}_{t_{k+1}} = X_{t_{k+1}}]\mathrm{d}\theta. \tag{4.20}$$

The rest of the proof of the LAMN property consists in expanding $\delta(\dot{X}^{\theta}_{i,T}\ u^k_i)$ into several random variables $M^{(l)}_{i,t_k}$ (corresponding to the main term) and $H^{(l)}_{i,t_k}$, the latter satisfying the two conditions $\mathrm{E}^{\theta}_{t_k,x}[H^{(l)}_{i,t_k}] = O(n^{-2})$ and $(\mathrm{E}^{\theta}_{t_k,x}|H^{(l)}_{i,t_k}|^{\alpha})^{1/\alpha} = O(n^{-3/2})$ for all $\alpha > 1$, uniformly in x, k, θ. Thus, by Corollary 4.1, we conclude that their contributions converge to 0 in P^{θ_0}-probability:

$$\sum_{k=0}^{n-1}\int_{\theta_0}^{\theta_0+u/\sqrt{n}} n\mathrm{E}^{\theta}_{t_k,X_{t_k}}[H^{(l)}_{i,t_k}|X^{\theta}_{t_{k+1}} = X_{t_{k+1}}]\mathrm{d}\theta \overset{P^{\theta_0}}{\rightarrow} 0.$$

Set $\hat{u}^k_i = (\hat{u}^k_{i1,t}, \ldots, \hat{u}^k_{id,t})^*_{0\leqslant t\leqslant 1/n}$ with $\hat{u}^k_{ij,t} = (S^{-1})_{i,j}(\theta, t_k + t, X^{\theta}_{t_k+t})$. Since $\hat{u}^k_i$ is adapted, $\delta(\hat{u}^k_i)$ is simply an Itô integral (see (3.5)), that is, $\sum_j \int_0^{1/n} \hat{u}^k_{ij,t}\,\mathrm{d}W_{j,t}$. Using (3.6), we deduce that

$$\delta(\dot{X}^{\theta}_{i,1/n}u^k_i) = \dot{X}^{\theta}_{i,1/n}\delta(u^k_i) - \int_0^{1/n} \mathcal{D}_t\dot{X}^{\theta}_{i,1/n}\cdot u^k_{i,t}\,\mathrm{d}t$$

$$= \left(\sum_{j=1}^{d}\int_0^{1/n} \dot{S}_{i,j}(\theta,\ t+t_k,\ X^{\theta}_{t_k+t})\mathrm{d}W_{j,t}\right)\delta(\hat{u}^k_i) - \int_0^{1/n} \mathcal{D}_t\dot{X}^{\theta}_{i,t}\cdot \hat{u}^k_{i,t}\,\mathrm{d}t$$

$$+ \left(\dot{X}^{\theta}_{i,1/n} - \sum_{j=1}^{d}\int_0^{1/n} \dot{S}_{i,j}(\theta,\ t+t_k,\ X^{\theta}_{t_k+t})\mathrm{d}W_{j,t}\right)\delta(\hat{u}^k_i)$$

$$+ \dot{X}^{\theta}_{i,1/n}\delta(u^k_i - \hat{u}^k_i) - \int_0^{1/n} (\mathcal{D}_t\dot{X}^{\theta}_{i,1/n}\cdot u^k_{i,t} - \mathcal{D}_t\dot{X}^{\theta}_{i,t}\cdot \hat{u}^k_{i,t})\mathrm{d}t$$

$$:= M^{(1)}_{i,t_k} - M^{(2)}_{i,t_k} + H^{(1)}_{i,t_k} + H^{(2)}_{i,t_k} - H^{(3)}_{i,t_k}. \tag{4.21}$$

4.4.1. Main contributions

Let us write $\Delta X^{\theta}_k := X^{\theta}_{t_{k+1}} - X^{\theta}_{t_k}$ and $\Delta X_k := X_{t_{k+1}} - X_{t_k}$. We first consider terms $M^{(1)}_{i,t_k}$. Since S is invertible, it readily follows that

$$\mathrm{d}W_t = S^{-1}(\theta,\ t_k+t,\ X^{\theta}_{t_k+t})\mathrm{d}X^{\theta}_{t_k+t} - S^{-1}(\theta,\ t_k+t,\ X^{\theta}_{t_k+t})b(\theta,\ t_k+t,\ X^{\theta}_{t_k+t})\mathrm{d}t$$

$$= S^{-1}(\theta,\ t_k,\ X^{\theta}_{t_k})\mathrm{d}X^{\theta}_{t_k+t} + (\mathrm{I}_d - S^{-1}(\theta,\ t_k,\ X^{\theta}_{t_k})S(\theta,\ t_k+t,\ X^{\theta}_{t_k+t}))\mathrm{d}W_t$$

$$- S^{-1}(\theta,\ t_k,\ X^{\theta}_{t_k})b(\theta,\ t_k+t,\ X^{\theta}_{t_k+t})\mathrm{d}t.$$

Thus, easy computations using standard Itô calculus techniques yield

$$M^{(1)}_{i,t_k} = \left(\sum_{j,m=1}^{d} \dot{S}_{i,j}(\theta,\ t_k,\ X^{\theta}_{t_k})\int_0^{1/n}(S^{-1})_{j,m}(\theta,\ t_k,\ X^{\theta}_{t_k})\mathrm{d}X^{\theta}_{m,t_k+t}\right)$$

$$\times\left(\sum_{j,m=1}^{d}(S^{-1})_{i,j}(\theta,\ t_k,\ X^{\theta}_{t_k})\int_0^{1/n}(S^{-1})_{j,m}(\theta,\ t_k,\ X^{\theta}_{t_k})\mathrm{d}X^{\theta}_{m,t_k}+t\right) + H^{(4)}_{i,t_k}$$

$$:= \left(\sum_{m=1}^{d}(\dot{S}S^{-1})_{i,m}(\theta,\ t_k,\ X^{\theta}_{t_k})\Delta X^{\theta}_{m,k}\right)\left(\sum_{m=1}^{d}(S^{-2})_{i,m}(\theta,\ t_k,\ X^{\theta}_{t_k})\Delta X^{\theta}_{m,k}\right) + H^{(4)}_{i,t_k} \tag{4.22}$$

with $\mathrm{E}^{\theta}_{t_k,x}[H^{(4)}_{i,t_k}] = O(n^{-2})$, $(\mathrm{E}^{\theta}_{t_k,x}|H^{(4)}_{i,t_k}|^{\alpha})^{1/\alpha} = O(n^{-3/2})$ for all $\alpha > 1$, uniformly in x, k, θ.

Turning to terms $M^{(2)}_{i,t_k}$, we first deduce from (4.11) that $(\mathcal{D}_t\dot{X}^{\theta}_{i,t})_j = \dot{S}_{i,j}(\theta,\ t_k+t,\ X^{\theta}_{t_k+t}) + (S'_i(\theta,\ t_k+t,\ X^{\theta}_{t_k+t})\dot{X}^{\theta}_t)_j$, so that it readily follows that

$$M^{(2)}_{i,t_k} = \sum_{j=1}^{d} \int_0^{1/n} (\dot{S}_{i,j}(\theta,\, t_k+t,\, X^\theta_{t_k+t}) + (S'_i(\theta,\, t_k+t,\, X^\theta_{t_k+t})\dot{X}^\theta_t)_j)(S^{-1})_{i,j}(\theta,\, t_k+t,\, X^\theta_{t_k+t})\mathrm{d}t$$

$$= \frac{1}{n}\sum_{j=1}^{d} \dot{S}_{i,j}(\theta,\, t_k,\, X^\theta_{t_k})(S^{-1})_{i,j}(\theta,\, t_k,\, X^\theta_{t_k}) + H^{(5)}_{i,t_k} := \frac{1}{n}(\dot{S}S^{-1})_{i,i}(\theta,\, t_k,\, X^\theta_{t_k}) + H^{(5)}_{i,t_k}, \tag{4.23}$$

with the required estimates on the mean and the L_α-norms of $H^{(5)}_{i,t_k}$ to ensure that it gives a negligible contribution in $Z_n(\theta_0,\, \theta_0 + u/\sqrt{n})$.

4.4.2. Negligible contributions

We must verify that, for $l = 1, 2, 3$, $\mathrm{E}^\theta_{t_k,x}[H^{(l)}_{i,t_k}] = O(n^{-2})$ and $(\mathrm{E}^\theta_{t_k,x}|H^{(l)}_{i,t_k}|^\alpha)^{1/\alpha} = O(n^{-3/2})$ for all $\alpha > 1$, uniformly in x, k, θ. We only sketch the proof of these estimates.

For terms $H^{(1)}_{i,t_k}$, recall that $\delta(\hat{u}^k_i)$ is an Itô integral, so that standard Itô calculus enables us to prove the required estimates.

Turning to terms $H^{(2)}_{i,t_k}$, the L_α-norms of order $n^{-3/2}$ can be directly obtained using $\|\dot{X}^\theta_{1/n}\|_p = O(n^{-1/2})$ and inequality (3.8) combined with the estimates (4.12), (4.13). To obtain the $O(n^{-2})$-estimate for the mean, first transform the random variable $\delta(u^k_i - \hat{u}^k_i) \in \mathbb{D}^{1,2}$ into an Itô integral using the Clark–Ocone formula (Proposition 3.3), taking into account relation (3.7), and then use Itô's calculus combined with the estimates (3.8), (4.12) and (4.13).

Finally, for terms $H^{(3)}_{i,t_k}$, it is enough to prove that $\mathcal{D}_t\dot{X}^\theta_{i,1/n}\cdot u^k_{i,t} - \mathcal{D}_t\dot{X}^\theta_{i,t}\cdot\hat{u}^k_{i,t} = \int_t^{1/n}(\cdots)\,\mathrm{d}s + \int_t^{1/n}(\cdots)\mathrm{d}W_s$, with adequate L_p controls on the adapted integrands. For this, if we put $\hat{X}_t = (\dot{X}^\theta_t,\, X^\theta_t)$ (this is a $\mathbb{R}^{2d}$-valued diffusion process), note that $\mathcal{D}_t\hat{X}_{1/n} = \hat{Y}_{1/n}(\hat{Y}_t)^{-1}\hat{S}(\theta,\, t_k+t,\, \hat{X}_t)$ (see equality (4.15)), where $\hat{Y}$ (or $\hat{S}$) is the flow of $\hat{X}$ (or its diffusion coefficient): thus, $\mathcal{D}_t\dot{X}^\theta_{i,1/n}\cdot u^k_{i,t} - \mathcal{D}_t\dot{X}^\theta_{i,t}\cdot\hat{u}^k_{i,t}$ can be decomposed using Itô's formula between t and $1/n$.

4.4.3. Conclusion of the proof of the LAMN property

Plugging (4.22) and (4.23) into (4.21) and (4.20), we have proved that $Z_n(\theta_0,\, \theta_0 + u/\sqrt{n})$ is equal to

$$\sum_{i=1}^{d}\sum_{k=0}^{n-1}\int_{\theta_0}^{\theta_0+u/\sqrt{n}}\left[n\left(\sum_{m=1}^{d}(\dot{S}S^{-1})_{i,m}(\theta,\, t_k,\, X_{t_k})\Delta X_{m,k}\right)\left(\sum_{m=1}^{d}(S^{-2})_{i,m}(\theta,\, t_k,\, X_{t_k})\Delta X_{m,k}\right) - (\dot{S}S^{-1})_{i,i}(\theta,\, t_k,\, X_{t_k})\right]\mathrm{d}\theta + R_n := \sum_{k=0}^{n-1}\xi_k + R_n$$

with

$$\xi_k = \int_{\theta_0}^{\theta_0+u/\sqrt{n}} [n(\dot{S}S^{-1})(\theta,\, t_k,\, X_{t_k})\Delta X_k.(S^{-2})(\theta,\, t_k,\, X_{t_k})\Delta X_k - \mathrm{tr}(\dot{S}S^{-1})(\theta,\, t_k,\, X_{t_k})]\mathrm{d}\theta$$

and $R_n \xrightarrow{P^{\theta_0}} 0$. Now, if we set $\Gamma(\theta_0) = 2\int_0^1 \mathrm{tr}(\dot{S}S^{-1})^2(\theta_0,\, t,\, X_t^{\theta_0})\mathrm{d}t$, it is easy to check that

$$\sum_{k=0}^{n-1} \mathrm{E}^{\theta_0}[\xi_k|\mathcal{G}_{t_k}] \xrightarrow{P^{\theta_0}} -u^2\Gamma(\theta_0)/2, \qquad \sum_{k=0}^{n-1} \mathrm{E}^{\theta_0}[\xi_k^2|\mathcal{G}_{t_k}] - (\mathrm{E}^{\theta_0}[\xi_k|\mathcal{G}_{t_k}])^2 \xrightarrow{P^{\theta_0}} u^2\Gamma(\theta_0),$$

$$\sum_{k=0}^{n-1} \mathrm{E}^{\theta_0}[\xi_k^4|\mathcal{G}_{t_k}] \xrightarrow{P^{\theta_0}} 0, \qquad \sum_{k=0}^{n-1} \mathrm{E}^{\theta_0}[\xi_k \Delta W_{j,k}|\mathcal{G}_{t_k}] \xrightarrow{P^{\theta_0}} 0$$

for $j = 1, \ldots, d$. Hence, we complete the proof of the result using a central limit theorem for triangular arrays of random variables (see Jacod 1997, Theorem 3-2). □

5. Proof of Proposition 4.2

We first state some preliminary estimates for the transition density of X^θ. For $\mu > 0$, we denote by $G_\mu(t, x, y)$ the transition density of the scaled Brownian motion $(x + (1/\sqrt{\mu})W_t)_{t\geqslant 0}$, that is, the Gaussian kernel $G_\mu(t, x, y) = (2\pi t)^{-d/2}\mu^{d/2}\exp(-\mu\|y - x\|^2/2t)$. The density $p^\theta(s, t, x, y)$ satisfies the following estimates:

Proposition 5.1. *Under Assumptions R′ and E′, for any μ_1 and μ_2 such that $\mu_1 < \mu_{\min} \leqslant \mu_{\max} < \mu_2$, there exists $c > 0$ such that*

$$\frac{1}{c} G_{\mu_2}(t-s,\, x,\, y) \leqslant p^\theta(s,\, t,\, x,\, y) \leqslant cG_{\mu_1}(t-s,\, x,\, y), \tag{5.24}$$

$$|\dot{p}^\theta(s,\, t,\, x,\, y)| \leqslant cG_{\mu_1}(t-s,\, x,\, y), \tag{5.25}$$

for $0 \leqslant s < t \leqslant 1$ and $(\theta, x, y) \in \Theta \times \mathbb{R}^d \times \mathbb{R}^d$.

Proof. These estimates are classical: they can be found, for example in Azencott (1984, p. 478). Note that Azencott (1984) assumes in his context more smoothness with respect to θ than do we, but, with a little care, we see that Assumptions R′ and E′ are sufficient for our purpose.

Another way to derive (5.25) consists in expressing $\dot{p}^\theta$ as the expectation of some random variable, using similar Malliavin arguments to those in Proposition 4.1, and applying standard estimates. □

We now return to the proof of estimates (4.18) and (4.19). It is easy to see that

$$\mathrm{E}^{\theta_0}_{t_k,x}(\mathrm{E}^{\theta}_{t_k,x}[H|X^\theta_{t_k+T} = X^{\theta_0}_{t_k+T}]) = \mathrm{E}^{\theta}_{t_k,x}\left[H\frac{p^{\theta_0}}{p^\theta}(t_k,\, t_k+T,\, x,\, X^\theta_{t_k+T})\right]. \tag{5.26}$$

Using the Hölder inequality (with α and β conjugate) and (5.24), it follows that the right-hand side of (5.26) is bounded by

$$C(\mathrm{E}^{\theta}_{t_k,x}|H|^{\alpha})^{1/\alpha}\left(\int_{\mathbb{R}^d} G^{\beta}_{\mu_1}(T,\, x,\, y)G^{1-\beta}_{\mu_2}(T,\, x,\, y)\mathrm{d}y\right)^{1/\beta} \leqslant C'(\mathrm{E}^{\theta}_{t_k,x}|H|^{\alpha})^{1/\alpha},$$

since the integral with respect to y is finite as soon as $\beta\mu_1 + (1-\beta)\mu_2 > 0 \Leftrightarrow \alpha > \mu_1/\mu_2$: this condition is satisfied up to modifying μ_1 and μ_2 from the beginning. It completes the proof of the estimate (4.18).

To obtain (4.19), we deduce from (5.26) that

$$\mathrm{E}^{\theta_0}_{t_k,x}(\mathrm{E}^{\theta}_{t_k,x}[H|X^{\theta}_{t_k+T} = X^{\theta_0}_{t_k+T}]) = \mathrm{E}^{\theta}_{x}[H] + \int_{\theta}^{\theta_0} \mathrm{d}\theta' \mathrm{E}^{\theta}_{t_k,x}\left[H\frac{\dot{p}^{\theta'}}{p^{\theta}}(t_k,\, t_k+T,\, x,\, X^{\theta}_{t_k+T})\right].$$

We estimate the last expectation using the same arguments as before, exploiting the upper bound (5.25) for $\dot{p}^{\theta'}$ instead of those for p^{θ_0}. □

Acknowledgements

This work was partially supported by the Université Pierre et Marie Curie Paris VI. The author wishes to express his thanks to Professor Dominique Picard for many stimulating conversations during the preparation of this work.

References

Azencott, R. (1984) Densité des diffusions en temps petit: Développements asymptotiques. In J. Azéma and M. Yor (eds), *Séminaire de Probabilités XVIII*, Lecture Notes in Math. 1059. Berlin: Springer-Verlag.

Delattre, S. (1997) Estimation du coefficient de diffusion d'un processus de diffusion en présence d'erreurs d'arrondis. Doctoral thesis, Université Paris VI.

Donhal, G. (1987) On estimating the diffusion coefficient. *J. Appl. Probab.*, **24**, 105–114.

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

Genon-Catalot, V. and Jacod, J. (1994) Estimation of the diffusion coefficient for diffusion processes: Random sampling. *Scand. J. Statist.*, **21**(3), 193–221.

Jacod, J. (1997) On continuous conditional Gaussian martingales and stable convergence in law. In J. Azéma, M. Émery and M. Yor (eds), *Séminaire de probabiliés XXXI*, Lecture Notes in Math. 1655. Berlin: Springer-Verlag.

Jeganathan, P. (1982) On the asymptotic theory of estimation when the limit of the log-likelihood is mixed normal. *Sankhyā Ser. A*, **44**, 173–212.

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

Kunita, H. (1984) Stochastic differential equations and stochastic flows of diffeomorphisms. In P.L. Hennequin (ed.), *École d'Été de Probabilités de Saint-Flour XII, 1982*, Lecture Notes in Math. 1097, pp. 144–305. Berlin: Springer-Verlag.

Le Cam, L. (1986) *Asymptotic Methods in Statistical Decision Theory.* Berlin: Springer-Verlag.
Le Cam, L. and Lo Yang, G. (1990) *Asymptotics in Statistics*. Berlin: Springer-Verlag.
Nualart, D. (1995) *Malliavin Calculus and Related Topics*. New York: Springer-Verlag.

Received February 2000 and revised March 2001