LAMN property for elliptic diffusion: a Malliavin calculus approach.

Emmanuel GOBET*

February 16, 2000

Abstract

In this paper, 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 linear parameter θ: the sample $(X_{k/n})_{0\leq k\leq n}$ corresponds to an observation of X at equidistant times of the interval $[0,1]$. We prove that LAMN property holds true for the likelihoods, 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.

KEY WORDS: conditional expectation, convergence of sums of random variables, diffusion process, LAMN property, log-likelihood ratios, Malliavin calculus, parametric estimation.

MSC CLASSIFICATION: 60Fxx, 60Hxx, 62Fxx, 62Mxx.

1 Introduction

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

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

for $t \in [0,1]$, where B is a d-dimensional Brownian motion, x is fixed, b and S are known smooth functions of (θ, t, x). θ is a linear 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 when 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 [4] explicit 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. An other interesting issue is to know if these estimators are asymptotically efficient: in some way, this is related to the Local Asymptotic Mixed Normality (LAMN) property, which we now recall (see e.g. Le Cam and Yang [9] chapter 5). If $\mathcal{F}_n = \sigma(X_{t_k} : 0 \leq k \leq n)$, we denote the

*Université Paris VI - UMR 7599, Laboratoire de Probabilités et Modèles Aléatoires, Boîte courrier 188, 4 Place Jussieu, 75252 Paris Cedex 05 - FRANCE, e-mail : gobet@ccr.jussieu.fr

restriction of $\mathbb{P}^\theta$ to $\mathcal{F}_n$ by $\mathbb{P}_n^\theta$, and the log-likelihood ratio of $\mathbb{P}_n^\theta$ w.r.t. $\mathbb{P}_n^{\theta_0}$ by $Z_n(\theta_0, \theta)$. The sequence $((\mathbb{R}^d)^n, \mathcal{F}_n, (\mathbb{P}_n^\theta)_{\theta \in \Theta})$ of statistical models has the LAMN property for the likelihoods, at θ_0, at rate $\sqrt{n}$ and conditional variance $\Gamma(\theta_0) > 0$ if

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

with $R_n \xrightarrow{\mathbb{P}^{\theta_0}} 0$ and $(\Delta_n, \Gamma(\theta_0)) \xrightarrow{\mathcal{L}(\mathbb{P}^{\theta_0})} (\Delta, \Gamma(\theta_0))$, where $\Delta \sim \mathcal{N}(0,1)$ is a Gaussian variable, independent of $\Gamma(\theta_0)$. When the LAMN property holds true for the likelihoods, one can apply minimax theorems (see Jeganathan [6] [7]) 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 likelihoods 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 [4] satisfy the LAMN condition.
Actually, the 1-dimensional case has been considered by Donhal [2]: its proof relies on a good expansion of the transition probability density p^θ of X. Unfortunately, in higher dimension (except for some specific cases, see Genon-Catalot and Jacod [3]), the well-known expansions of p^θ (see e.g. Azencott [1]) are not sufficient to adapt Donhal's proof.

To get the expected 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 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 analyze 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 to suit well to the study of the likelihoods: the case of degenerate coefficient diffusion (with hypo-ellipticity conditions) may be treated in the same way. Furthermore, presumably, this approach may also enable to handle non Markovian situations, s.t. hidden diffusions or stochastic differential equations with memory: this will concern forthcoming papers.

The paper is organized as follows: in section 3, we briefly introduce the material necessary to 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 these kind of expectations (Corollary 4.1). An intermediate technical result is proved in section 5.

2 Assumptions and results

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

Assumption (R): the functions $b(\theta, t, x)$ and $S(\theta, t, x)$ are of class $C^{1+\alpha}$ w.r.t. θ ($\alpha > 0$). The functions b, S, $\dot{b}$, $\dot{S}$, b' and S' are of class $C^{1,2}$ w.r.t. (t, x). Moreover, all these functions (and their derivatives) are uniformly bounded on $\Theta \times [0,1] \times \mathbb{R}^d$.

Assumption (E): the matrix S is symmetric, positive and satisfies an uniform ellipticity condition

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

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

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

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

In the sequel, the indices θ, s, x in $\mathbb{E}_{s,x}^\theta$ stand in reference for the expectation under the law of the diffusion X^θ starting at x at time s.

Remark: our study is restricted to a linear parameter: actually, there is no additional technical difficulties to deal with multidimensional parameters, it is simply more cumbersome to write down. 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 $(X_{t_k})_{0\leq k\leq n}$ 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

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

The main result of the paper is the following Theorem.

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

$$Z_n(\theta_0, \theta_0 + u/\sqrt{n}) \overset{\mathcal{L}(\mathbb{P}^{\theta_0})}{\longrightarrow} u\ \sqrt{\Gamma(\theta_0)}\ \Delta - \frac{u^2}{2}\Gamma(\theta_0)$$

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

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 necessary to these computations.

3 Some basic results on the Malliavin calculus

The reader may refer to Nualart [10] for a detailed exposition of this section.

Fix a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ and let $(W_t)_{t\geq 0}$ be a d-dimensional Brownian motion. Fix $T \in (0,1]$. For $h(.) \in H = L_2([0,T], \mathbb{R}^d)$, $W(h)$ is the Wiener stochastic integral $\int_0^T h(t) \, . \, dW_t$. Let $\mathcal{S}$ denote the class of random variables of the form $F = f(W(h_1), \dots, W(h_N))$ where $f \in C_p^\infty(\mathbb{R}^N)$, $(h_1, \dots, h_N) \in H^N$ and $N \geq 1$. 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

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

The operator $\mathcal{D}$ is closable as an operator from $L_p(\Omega)$ to $L_p(\Omega, H)$, for any $p \geq 1$. Its domain is denoted by $\mathbb{D}^{1,p}$ w.r.t. the norm $\|F\|_{1,p} = \left[\mathbb{E}|F|^p + \mathbb{E}(\|\mathcal{D}F\|_H^p)\right]^{1/p}$. 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 $L_p(\Omega; H^{\otimes k})$, $p \geq 1$. If we define the norm $\|F\|_{k,p} = [\mathbb{E}|F|^p + \sum_{j=1}^k \mathbb{E}(\|\mathcal{D}^j F\|_{H^{\otimes j}}^p)]^{1/p}$, we denote its domain by $\mathbb{D}^{k,p}$.
One has the chain rule property:

Proposition 3.1. *Fix $p \geq 1$. For $f \in C_b^1(\mathbb{R}^d, \mathbb{R})$ and $F = (F_1, \cdots, F_d)$ a random vector whose components belong to $\mathbb{D}^{1,p}$, $f(F) \in \mathbb{D}^{1,p}$ and for $t \geq 0$, one has*

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

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:*

1. *the domain of δ (denoted by* $\mathrm{Dom}(\delta)$*) is the set of processes $u \in L_2([0,T] \times \Omega, \mathbb{R}^d)$ s.t.*

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

2. *if u belongs to* $\mathrm{Dom}(\delta)$*, 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*

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

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

Proposition 3.2.

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

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

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

3. *If F belongs to $\mathbb{D}^{1,2}$, then for any $u\in\mathrm{Dom}(\delta)$ s.t. $\mathbb{E}(F^2\int_0^T u_t^2\,dt)<+\infty$, one has*

$$\delta(F\,u)=F\,\delta(u)-\int_0^T \mathcal{D}_tF\,.\,u_t\,dt, \tag{3.6}$$

whenever the r.h.s. above belongs to $L_2(\Omega)$.

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

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

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

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

At last, we recall the Clark-Ocone's formula.

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

$$F=\mathbb{E}(F)+\int_0^T \mathbb{E}(\mathcal{D}_tF\;/\;\mathcal{F}_t)\,.\,dW_t.$$

4 Proof of the LAMN property

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

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

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

We now derive a new expression for $\frac{\dot{p}^\theta}{p^\theta}(t_k,t_k+T,x,y)$ as a conditional expectation, using Malliavin calculus. For this purpose, let us consider, in all this paragraph, the solution of (2.2) starting at x at time t_k, i.e. the $\mathbb{R}^d$-valued process denoted $(X_{t_k+t}^\theta)_{t\geq 0}$ solving

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

where S_j is the j-th column vector of S, $(W_t)_{t\geq 0}$ a new Brownian motion with its usual filtration $(\mathcal{F}_t)_{t\geq 0}$ (W corresponds to the shift of B at time t_k: even it depends on k, we omit to specify it because there will be no possible confusion in the sequel).
We associate to $(X^\theta_{t_k+t})_{t\geq 0}$ its flow, i.e. the Jacobian matrix $Y^\theta_t := \nabla_x X^\theta_{t_k+t}$, and its derivative w.r.t. θ denoted by $\dot{X}^\theta_t$.

Remark. Our notation with $X^\theta_{t_k+t}$, Y^θ_t and $\dot{X}^\theta_t$ are not homogeneous w.r.t. 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 next computations.

Under **(R)**, it is clear (see Kunita [8]) that Y^θ_t and $\dot{X}^\theta_t$ solve

$$\begin{aligned}
Y^\theta_t &= \mathrm{I}_d + \int_0^t b'(\theta, t_k+s, X^\theta_{t_k+s})\, Y^\theta_s\, ds + \sum_{j=1}^d \int_0^t S'_j(\theta, t_k+s, X^\theta_{t_k+s})\, Y^\theta_s\, dW_{j,s}, \\
\dot{X}^\theta_t &= \int_0^t \left(\dot{b}(\theta, t_k+s, X^\theta_{t_k+s}) + b'(\theta, t_k+s, X^\theta_{t_k+s})\, \dot{X}^\theta_s\right) ds \\
&\quad + \sum_{j=1}^d \int_0^t \left(\dot{S}_j(\theta, t_k+s, X^\theta_{t_k+s}) + S'_j(\theta, t_k+s, X^\theta_{t_k+s})\, \dot{X}^\theta_s\right) dW_{j,s}.
\end{aligned} \tag{4.11}$$

For any $t\geq 0$, the random variables $X^\theta_{t_k+t}$, Y^θ_t and $\dot{X}^\theta_t$ are weakly differentiable: actually, one has $X^\theta_{t_k+t}\in\bigcap_{p\geq 1}\mathbb{D}^{3,p}$, $Y^\theta_t\in\bigcap_{p\geq 1}\mathbb{D}^{2,p}$, $\dot{X}^\theta_t\in\bigcap_{p\geq 1}\mathbb{D}^{2,p}$, with the following estimates

$$\text{for } j=1,2,3, \qquad \sup_{r_1,\cdots,r_j\in[0,T]} \mathbb{E}_{t_k,x}\left(\sup_{r_1\vee\cdots\vee r_j\leq t\leq T} \|\mathcal{D}_{r_1,\cdots,r_j} X^\theta_{t_k+t}\|^p\right) \leq c, \tag{4.12}$$

$$\text{for } j=1,2, \qquad \sup_{r_1,\cdots,r_j\in[0,T]} \mathbb{E}_{t_k,x}\left(\sup_{r_1\vee\cdots\vee r_j\leq t\leq T} \|\mathcal{D}_{r_1,\cdots,r_j} Y^\theta_t\|^p\right) \leq c, \tag{4.13}$$

$$\text{for } j=1,2, \qquad \sup_{r_1,\cdots,r_j\in[0,T]} \mathbb{E}_{t_k,x}\left(\sup_{r_1\vee\cdots\vee r_j\leq t\leq T} \|\mathcal{D}_{r_1,\cdots,r_j} \dot{X}^\theta_t\|^p\right) \leq c, \tag{4.14}$$

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

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

Proposition 4.1. *Assume* **(R)** *and* **(E)**. *Set $T>0$. For $1\leq i\leq d$, let us define $u^k_i = (u^k_{i,s})_{0\leq s\leq T}$ the $\mathbb{R}^d$-valued process whose j-th component is equal to $u^k_{ij,s} = \left(S^{-1}(\theta, t_k+s, X^\theta_{t_k+s})\, Y^\theta_s\, (Y^\theta_T)^{-1}\right)_{i,j}$. Then, one has*

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

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

$$\begin{aligned}\int_\Theta d\theta\, g(\theta)\, \mathbb{E}^\theta_{t_k,x}\left[\nabla f(X^\theta_{t_k+T}) \,.\, \dot{X}^\theta_T\right] &= -\int_\Theta d\theta\, g'(\theta)\, \mathbb{E}^\theta_{t_k,x}\left[f(X^\theta_{t_k+T})\right] \\ &= -\int_\Theta d\theta\, g'(\theta) \int_{\mathbb{R}^d} dy\, p^\theta(t_k, t_k+T, x, y)\, f(y) \\ &= \int_\Theta d\theta\, g(\theta) \int_{\mathbb{R}^d} dy\, \dot{p}^\theta(t_k, t_k+T, x, y)\, f(y),\end{aligned}$$

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

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

Indeed, the proof of Proposition 4.1 now can be easily completed by comparing both expressions obtained for $-\int_\Theta d\theta\, g'(\theta)\, \mathbb{E}^\theta_{t_k,x}[f(X^\theta_{t_k+T})]$.
The derivation of the formula (4.17) for $\mathbb{E}^\theta_{t_k,x}[\nabla f(X^\theta_{t_k+T}) \,.\, \dot{X}^\theta_T]$ 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 $s \le T$, $\mathcal{D}_s X^\theta_{t_k+T} = Y^\theta_T\, (Y^\theta_s)^{-1} S(\theta, t_k{+}s, X^\theta_{t_k+s})$ is invertible, so that $\partial_{x_i} f(X^\theta_{t_k+T}) = \mathcal{D}_s(f(X^\theta_{t_k+T})) \,.\, u^k_{i,s}$. Then, it follows that the l.h.s. of (4.17) is equal to

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

by the relation (3.4). This completes the proof of Proposition 4.1.

□

4.2 About the convergence of a sum of conditional expectations

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

Proposition 4.2. *Assume* **(R)** *and* **(E)**. *Fix $T > 0$. Let us consider H, a $\mathcal{F}_T$-measurable random variable. For any $\theta \in \Theta$ and any $\alpha > \mu_{max}/\mu_{min}$, one has*

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

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

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

The next result is our basic tool to analyze the convergence of the sum of conditional expectations.

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

$$\mathbb{E}^\theta_{t_k,x}\left[H_{t_k}\right] = O(n^{-2}) \qquad \text{and} \qquad \left(\mathbb{E}^\theta_{t_k,x}|H_{t_k}|^{2\,\alpha}\right)^{1/2\alpha} = O(n^{-3/2})$$

uniformly in x*,* k *and* θ*. Then, under* **(R)** *and* **(E)***, one has*

$$\sum_{k=0}^{n-1}\int_{\theta_0}^{\theta_0+u/\sqrt{n}} n\ \mathbb{E}^\theta_{t_k,X_{t_k}}\left[H_{t_k}\ /\ X^\theta_{t_{k+1}} = X_{t_{k+1}}\right]\ d\theta \overset{\mathbb{P}^{\theta_0}}{\longrightarrow} 0.$$

Proof. Set $\xi^n_k = \int_{\theta_0}^{\theta_0+u/\sqrt{n}} n\ \mathbb{E}^\theta_{t_k,X_{t_k}}\left[H_{t_k}\ /\ X^\theta_{t_{k+1}} = X_{t_{k+1}}\right]\ d\theta$: these are $\mathcal{G}_{t_{k+1}}$-measurable random variables. Using Proposition 4.2 and the conditions of the statement of Corollary 4.1, it is easy to check that $\mathbb{E}^{\theta_0}\left[\xi^n_k\ /\ \mathcal{G}_{t_k}\right] = O(n^{-3/2})$ and $\mathbb{E}^{\theta_0}\left[(\xi^n_k)^2\ /\ \mathcal{G}_{t_k}\right] = O(n^{-2})$, uniformly in k. We complete the proof, applying the following classical convergence result about triangular arrays of random variables.

Lemma 4.1. *(Genon-Catalot and Jacod [4], Lemma 9). Let* ξ^n_k*,* U *be random variables, with* ξ^n_k *being* $\mathcal{G}_{t_{k+1}}$*-measurable. The two following conditions imply* $\sum_{k=0}^{n-1}\xi^n_k \overset{\mathbb{P}}{\to} U$*:*

$$\sum_{k=0}^{n-1}\mathbb{E}[\xi^n_k\ /\ \mathcal{G}_{t_k}] \overset{\mathbb{P}}{\to} U \qquad \text{and} \qquad \sum_{k=0}^{n-1}\mathbb{E}\left[(\xi^n_k)^2\ /\ \mathcal{G}_{t_k}\right] \overset{\mathbb{P}}{\to} 0.$$

□

4.3 Convergence of $Z_n(\theta_0, \theta_0 + u/\sqrt{n})$ under $\mathbb{P}^{\theta_0}$

From the equality (4.9) and Proposition 4.1, one deduces that

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

The remainder of the proof consists in expanding $\delta(\dot{X}^\theta_{i,T}\ u^k_i)$ into few random variables $M^{(l)}_{i,t_k}$ (corresponding to the main term) and $H^{(l)}_{i,t_k}$, these ones satisfying the two conditions $\mathbb{E}^\theta_{t_k,x}\left[H^{(l)}_{i,t_k}\right] = O(n^{-2})$ and $\left(\mathbb{E}^\theta_{t_k,x}|H^{(l)}_{i,t_k}|^\alpha\right)^{1/\alpha} = O(n^{-3/2})$ for all $\alpha > 1$, uniformly in x, k, θ. Thus, by Corollary 4.1, we conclude that their contributions converge to 0 in $\mathbb{P}^{\theta_0}$-probability:

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

Set $\hat{u}^k_i = (\hat{u}^k_{i1,t},\cdots,\hat{u}^k_{id,t})^*_{0\le t\le 1/n}$ 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 Ito integral (see (3.5)), i.e. $\sum_j\int_0^{1/n}\hat{u}^k_{ij,t}dW_{j,t}$. Using (3.6), we deduce that

$$
\begin{aligned}
\delta(\dot{X}^\theta_{i,1/n}\, u^k_i) &= \dot{X}^\theta_{i,1/n}\,\delta(u^k_i) - \int_0^{1/n} \mathcal{D}_t\dot{X}^\theta_{i,1/n}\,.\,u^k_{i,t}\,dt \\
&= \left(\sum_{j=1}^d \int_0^{1/n} \dot{S}_{i,j}(\theta, t+t_k, X^\theta_{t_k+t})dW_{j,t}\right)\,\delta(\hat{u}^k_i) - \int_0^{1/n} \mathcal{D}_t\dot{X}^\theta_{i,t}\,.\,\hat{u}^k_{i,t}\,dt \\
&\quad + \left(\dot{X}^\theta_{i,1/n} - \sum_{j=1}^d \int_0^{1/n} \dot{S}_{i,j}(\theta, t+t_k, X^\theta_{t_k+t})dW_{j,t}\right)\,\delta(\hat{u}^k_i) \\
&\quad + \dot{X}^\theta_{i,1/n}\,\delta(u^k_i - \hat{u}^k_i) - \int_0^{1/n} (\mathcal{D}_t\dot{X}^\theta_{i,1/n}\,.\,u^k_{i,t} - \mathcal{D}_t\dot{X}^\theta_{i,t}\,.\,\hat{u}^k_{i,t})\,dt \\
&:= M^{(1)}_{i,t_k} - M^{(2)}_{i,t_k} + H^{(1)}_{i,t_k} + H^{(2)}_{i,t_k} - H^{(3)}_{i,t_k}. \qquad (4.21)
\end{aligned}
$$

4.3.1 Main contributions

Let us denote $\Delta X^\theta_k := X^\theta_{t_{k+1}} - X^\theta_{t_k}$ and $\Delta X_k := X_{t_{k+1}} - X_{t_k}$.

a) Terms $M^{(1)}_{i,t_k}$: since S is invertible, it readily follows

$$
\begin{aligned}
dW_t &= S^{-1}(\theta, t_k+t, X^\theta_{t_k+t})\,dX^\theta_{t_k+t} - S^{-1}(\theta, t_k+t, X^\theta_{t_k+t})\,b(\theta, t_k+t, X^\theta_{t_k+t})\,dt \\
&= S^{-1}(\theta, t_k, X^\theta_{t_k})\,dX^\theta_{t_k+t} \\
&\quad + (\mathrm{I}_d - S^{-1}(\theta, t_k, X^\theta_{t_k})S(\theta, t_k+t, X^\theta_{t_k+t}))\,dW_t - S^{-1}(\theta, t_k, X^\theta_{t_k})\,b(\theta, t_k+t, X^\theta_{t_k+t})\,dt.
\end{aligned}
$$

Thus, easy computations using standard Ito calculus techniques yield

$$
\begin{aligned}
M^{(1)}_{i,t_k} &= \left(\sum_{j,m=1}^d \dot{S}_{i,j}(\theta, t_k, X^\theta_{t_k}) \int_0^{1/n} (S^{-1})_{j,m}(\theta, t_k, X^\theta_{t_k})dX^\theta_{m,t_k+t}\right) \\
&\quad \left(\sum_{j,m=1}^d (S^{-1})_{i,j}(\theta, t_k, X^\theta_{t_k}) \int_0^{1/n} (S^{-1})_{j,m}(\theta, t_k, X^\theta_{t_k})dX^\theta_{m,t_k+t}\right) + H^{(4)}_{i,t_k} \\
&:= \left(\sum_{m=1}^d (\dot{S}S^{-1})_{i,m}(\theta, t_k, X^\theta_{t_k})\Delta X^\theta_{m,k}\right)\left(\sum_{m=1}^d (S^{-2})_{i,m}(\theta, t_k, X^\theta_{t_k})\Delta X^\theta_{m,k}\right) + H^{(4)}_{i,t_k} \qquad (4.22)
\end{aligned}
$$

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

b) 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

$$
\begin{aligned}
M^{(2)}_{i,t_k} &= \sum_{j=1}^d \int_0^{1/n} \left(\dot{S}_{i,j}(\theta, t_k+t, X^\theta_{t_k+t}) + (S'_i(\theta, t_k+t, X^\theta_{t_k+t})\,\dot{X}^\theta_t)_j\right)(S^{-1})_{i,j}(\theta, t_k+t, X^\theta_{t_k+t})\,dt \\
&= \frac{1}{n}\sum_{j=1}^d \dot{S}_{i,j}(\theta, t_k, X^\theta_{t_k})(S^{-1})_{i,j}(\theta, t_k, X^\theta_{t_k}) + H^{(5)}_{i,t_k} := \frac{1}{n}\,(\dot{S}S^{-1})_{i,i}(\theta, t_k, X^\theta_{t_k}) + H^{(5)}_{i,t_k}, \qquad (4.23)
\end{aligned}
$$

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 $Z_n(\theta_0, \theta_0 + u/\sqrt{n})$.

4.3.2 Negligible contributions

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

a) Terms $H^{(1)}_{i,t_k}$: remind that $\delta(\hat{u}^k_i)$ is an Ito integral, so that standard Ito's calculus enables to prove the required estimates.

b) 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 the inequality (3.8) combined with the estimates (4.12), (4.13).

To get 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 Ito integral owing Clark-Ocone's formula (Proposition 3.3), taking into account the relation (3.7), and then, use Ito's calculus combined with the estimates (3.8), (4.12) and (4.13).

c) Terms $H^{(3)}_{i,t_k}$: it is enough to prove that $\mathcal{D}_t\dot{X}^\theta_{i,1/n} \,.\, u^k_{i,t} - \mathcal{D}_t\dot{X}^\theta_{i,t} \,.\, \hat{u}^k_{i,t} = \int_t^{1/n}(\cdots)\,ds + \int_t^{1/n}(\cdots)\,dW_s$, 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 $\hat{Y}$ (resp. $\hat{S}$) is the flow of $\hat{X}$ (resp. its diffusion coefficient): thus, $\mathcal{D}_t\dot{X}^\theta_{i,1/n} \,.\, u^k_{i,t} - \mathcal{D}_t\dot{X}^\theta_{i,t} \,.\, \hat{u}^k_{i,t}$ can be decomposed using Ito's formula between t and $1/n$.

4.3.3 End of the proof of the LAMN property

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

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

with $\xi_k = \int_{\theta_0}^{\theta_0+u/\sqrt{n}}\left[n\,(\dot{S}S^{-1})(\theta, t_k, X_{t_k})\Delta X_k.(S^{-2})(\theta, t_k, X_{t_k})\Delta X_k - \mathrm{Tr}(\dot{S}S^{-1})(\theta, t_k, X_{t_k})\right]\,d\theta$ and $R_n \xrightarrow{\mathbb{P}^{\theta_0}} 0$. Now, if one sets $\Gamma(\theta_0) = 2\,\int_0^1 \mathrm{Tr}(\dot{S}S^{-1})^2(\theta_0, t, X^{\theta_0}_t)\,dt$, it is easy to check that

$$\sum_{k=0}^{n-1}\mathbb{E}^{\theta_0}[\xi_k \,/\, \mathcal{G}_{t_k}] \overset{\mathbb{P}^{\theta_0}}{\to} -u^2\,\Gamma(\theta_0)/2, \qquad \sum_{k=0}^{n-1}\mathbb{E}^{\theta_0}[\xi_k^2 \,/\, \mathcal{G}_{t_k}] - (\mathbb{E}^{\theta_0}[\xi_k \,/\, \mathcal{G}_{t_k}])^2 \overset{\mathbb{P}^{\theta_0}}{\to} u^2\,\Gamma(\theta_0),$$
$$\sum_{k=0}^{n-1}\mathbb{E}^{\theta_0}[\xi_k^4 \,/\, \mathcal{G}_{t_k}] \overset{\mathbb{P}^{\theta_0}}{\to} 0, \qquad \sum_{k=0}^{n-1}\mathbb{E}^{\theta_0}[\xi_k\,\Delta W_{j,k} \,/\, \mathcal{G}_{t_k}] \overset{\mathbb{P}^{\theta_0}}{\to} 0$$

for $j = 1, \cdots, d$. Hence, we complete the proof of the result using a central limit theorem for triangular arrays of random variables (see Jacod [5] Theorem 3-2).

□

5 Proof of Proposition 4.2

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

Proposition 5.1. *Under* **(R)** *and* **(E)**, *for any μ_1 and μ_2 s.t. $\mu_1 < \mu_{min} \leq \mu_{max} < \mu_2$, there exists $c > 0$ s.t.*

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

$$\left|\dot{p}^\theta(s,t,x,y)\right| \quad \leq c\ G_{\mu_1}(t-s,x,y), \tag{5.25}$$

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

Proof. These estimates are classical: they can be found e.g. in Azencott [1] p.478. Note that Azencott [1] assumes in his context more smoothness w.r.t. θ than us, but being a little careful, we see that **(R)** and **(E)** are sufficient assumptions for our purpose.
An other way to derive (5.25) consists in expressing $\dot{p}^\theta$ as the expectation of some random variable, using similar Malliavin arguments as in Proposition 4.1, and to apply standard estimates.

□

We now come back to the proof of the estimates (4.18) and (4.19). It is easy to see that one has

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

Using Hölder inequality (with α and β conjugate) and the estimates (5.24), it follows that the r.h.s. of (5.26) is bounded by

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

since the integral w.r.t. y is finite as soon as $\beta\ \mu_1 + (1-\beta)\ \mu_2 > 0 \Longleftrightarrow \alpha > \mu_1/\mu_2$: 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 get the estimate (4.19), one deduces from (5.26) that

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

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}.

□

Acknowledgement: the author wishes to express his thanks to Professor Dominique Picard for many stimulating conversations, during the preparation of this work.

References

[1] R. Azencott: Densité des diffusions en temps petit: développement asymptotiques - Séminaire de probabilités XVIII - Lecture Notes in Mathematics 1059 - Springer Verlag - 1982-83.

[2] G. Donhal: On estimating the diffusion coefficient - J. Appl. Prob. - Vol.24 - 1987 - pp.105-114.

[3] V. Genon-Catalot and J. Jacod: On the estimation of the diffusion coefficient for diffusion processes: random sampling - Preprint ♯100, Lab. Probabilités, Univ. Paris VI - 1992.

[4] V. Genon-Catalot and J. Jacod: On the estimation of the diffusion coefficient for multidimensional diffusion processes - Ann. Inst. H. Poincaré (Probab. Statist.) - Vol.29 - 1993 - pp.119-151.

[5] J. Jacod: On continuous conditional Gaussian martingales and stable convergence in law - Séminaire de probabiliés XXXI - Lecture Notes in Mathematics 1655 - Springer Verlag - 1997.

[6] P. Jeganathan: On the asymptotic theory of estimation when the limit of the loglikelihhod is mixed normal - Sankya, Series A - Vol.44 - 1982 - pp. 173-212.

[7] P. Jeganathan: Some asymptotic properties of risk functions when the limit of the experiment is mixed normal - Sankya, Series A - Vol.45 - 1983 - pp. 66-87.

[8] H. Kunita: Stochastic differential equations and stochastic flows of diffeomorphisms - Ecole d'Eté de Probabilités de St-Flour XII, 1982 - Lecture Notes in Math. 1097 - Springer Verlag - 1984 - pp. 144-305.

[9] L. Le Cam and G.L. Yang: Asymptotics in Statistics, Springer Verlag - 1990.

[10] D. Nualart: Malliavin calculus and related topics, Springer Verlag - 1995.