LAN property for ergodic diffusions with discrete observations.

Emmanuel GOBET*†

September, 2000

ABSTRACT. We consider a multidimensional elliptic diffusion $X^{\alpha,\beta}$, whose drift $b(\alpha, x)$ and diffusion coefficients $S(\beta, x)$ depend on multidimensional parameters α and β. We assume some various hypotheses on b and S, which ensure that $X^{\alpha,\beta}$ is ergodic, and we address the problem of the validity of the Local Asymptotic Normality (LAN in short) property for the likelihoods, when the sample is $(X_{k\Delta_n})_{0\leq k\leq n}$, under the conditions $\Delta_n \to 0$ and $n\Delta_n \to +\infty$. We prove that the LAN property is satisfied, at rate $\sqrt{n\Delta_n}$ for α and $\sqrt{n}$ for β: our approach is based on a Malliavin calculus transformation of the likelihoods.

KEY WORDS: ergodic diffusion process, LAN property, loglikelihood ratio, Malliavin calculus, parametric estimation.

MSC CLASSIFICATION: 60A10 - 60Fxx - 60Gxx - 60H07 - 62E20 - 62F12 - 62Mxx.

RÉSUMÉ. Nous considèrons un processus de diffusion multidimensionnel elliptique $X^{\alpha,\beta}$, dont les coefficients de dérive $b(\alpha, x)$ et de diffusion $S(\beta, x)$ dépendent de paramètres multidimensionnels α et β. Nous formulons plusieurs jeux d'hypothèses sur b and S, assurant l'ergodicité de $X^{\alpha,\beta}$, et nous nous intéressons à la validité de la propriète LAN (Local Asymptotic Normality) pour les vraisemblances, quand l'échantillon observé est $(X_{k\Delta_n})_{0\leq k\leq n}$, sous les conditions $\Delta_n \to 0$ et $n\Delta_n \to +\infty$. Nous démontrons que la propriètè LAN est vérifiée, avec les vitesses $\sqrt{n\Delta_n}$ pour α et $\sqrt{n}$ pour β: notre approche repose sur une réécriture du rapport de vraisemblance à l'aide du calcul de Malliavin.

Introduction

Let $\mathbb{P}^{\alpha,\beta}$ be the law of $(X_t^{\alpha,\beta})_{t\geq 0}$, the $\mathbb{R}^d$-valued process solution of

$$X_t^{\alpha,\beta} = x_0 + \int_0^t b(\alpha, X_s^{\alpha,\beta})\, ds + \int_0^t S(\beta, X_s^{\alpha,\beta})\, dB_s, \tag{0.1}$$

where B is a d-dimensional Brownian motion, x_0 is fixed and known, b and S are known smooth functions. We are concerned with the estimation of the multidimensional parameters (α, β) which belong to Θ, an open subset of $\mathbb{R}^{n_\alpha} \times \mathbb{R}^{n_\beta}$ $(n_\alpha \geq 1, n_\beta \geq 1)$, when the observation is the discretized path

*Ecole Polytechnique - Centre de Mathématiques Appliquées - 91128 Palaiseau Cedex - FRANCE, e-mail : gobet@cmapx.polytechnique.fr

†This work was partially supported by Université Pierre et Marie Curie Paris 6.

$(X_{k\Delta_n})_{0\leq k\leq n}$. The asymptotic are $n\to+\infty$ and we consider the case when $\Delta_n\to 0$ and $n\Delta_n\to+\infty$, assuming that $X^{\alpha,\beta}$ is ergodic.

The purpose of this paper is to prove the Local Asymptotic Normality (LAN in short) property for the likelihoods under appropriate assumptions on b and S. We give a precise formulation of the problem in our setting (for a general account on the subject, see Le Cam and Yang [19] e.g.). If $\mathcal{F}_n=\sigma(X_{k\Delta_n}:0\leq k\leq n)$, we denote the restriction of $\mathbb{P}^{\alpha,\beta}$ to $\mathcal{F}_n$ by $\mathbb{P}_n^{\alpha,\beta}$. The sequence $((\mathbb{R}^d)^n,\mathcal{F}_n,(\mathbb{P}_n^{\alpha,\beta})_{(\alpha,\beta)\in\Theta})$ of statistical models has the LAN property for the likelihoods, at (α^0,β^0), with rates $\sqrt{n\Delta_n}$ for α^0 and $\sqrt{n}$ for β^0, with covariance matrix $\Gamma^{\alpha^0,\beta^0}\in\mathbb{R}^{n_\alpha+n_\beta}\otimes\mathbb{R}^{n_\alpha+n_\beta}$ if for any $u\in\mathbb{R}^{n_\alpha}$ and any $v\in\mathbb{R}^{n_\beta}$, one has

$$(0.2)\quad \log\left(\frac{d\mathbb{P}_n^{\alpha^0+\frac{u}{\sqrt{n\Delta_n}},\beta^0+\frac{v}{\sqrt{n}}}}{d\mathbb{P}_n^{\alpha^0,\beta^0}}\right)((X_{k\Delta_n})_{0\leq k\leq n})=\begin{pmatrix}u\\ v\end{pmatrix}.\mathcal{N}_n^{\alpha^0,\beta^0}-\frac{1}{2}\begin{pmatrix}u\\ v\end{pmatrix}.\Gamma^{\alpha^0,\beta^0}\begin{pmatrix}u\\ v\end{pmatrix}+R_n,$$

where $R_n=R_n(u,v)\overset{\mathbb{P}^{\alpha^0,\beta^0}}{\longrightarrow}0$ and $\mathcal{N}_n^{\alpha^0,\beta^0}\overset{\mathcal{L}(\mathbb{P}^{\alpha^0,\beta^0})}{\longrightarrow}\mathcal{N}^{\alpha^0,\beta^0}$ defined as a centered Gaussian vector with covariance matrix $\Gamma^{\alpha^0,\beta^0}$.
If the LAN property holds true and if $\Gamma^{\alpha^0,\beta^0}$ is non degenerate (this is somehow related to an identification condition on the statistical models), minimax theorems can be applied (see Hajek [10], Le Cam [18], or Le Cam and Yang [19] for a review) and $(\Gamma^{\alpha^0,\beta^0})^{-1}$ gives the lower bound for the asymptotic variance of estimators. This justifies the importance of such a property in parametric estimation problems.

The estimation procedure has been studied by several authors, mainly when $d=1$ (see Prakasa Rao [22], Florens Zmirou [4], Kessler [16]), while Yoshida [23] adopts a multidimensional setting. The estimators they propose are contrast ones: their construction is based either on a discretization of the likelihood associated to the continuous observation (see Yoshida [23] and also Genon-Catalot [6]), either on the use of some approximative schemes (see Florens Zmirou [4], Kessler [16]) (see also Genon-Catalot et al. [7] for general contrast functions in a different asymptotic framework). It is worth noticing that these estimators are asymptotically efficient, since their variance achieve the lower bound given by $(\Gamma^{\alpha^0,\beta^0})^{-1}$ as the reader may see from the statement of the LAN property (see Theorem 4.1 below). Some significant progresses have been recently realized by Kessler [16] concerning the assumptions on the form of the coefficients: in particular, he allows to deal with quite general diffusion coefficients S, whereas the previous works were restricted to the cases when $S(\beta,x)$ did not depend on x or was linear w.r.t the parameters. Moreover, to derive the asymptotic normality of the estimators, Kessler overcomes the restriction $n{\Delta_n}^2\to 0$ (see Prakasa Rao [22]) or $n{\Delta_n}^3\to 0$ (see Florens Zmirou [4] and Yoshida [23]), assuming only $n{\Delta_n}^p\to 0$ for some $p>1$: here, to get the LAN property, we need not require some specific form of the coefficients or some additional assumption on the decreasing rate of Δ_n.
In a Markov setting, the loglikelihood ratio can be naturally expressed as a sum of terms of the form $\log p^{\alpha,\beta}(\Delta_n,X_{k\Delta_n},X_{(k+1)\Delta_n})$, where $p^{\alpha,\beta}(t,x,y)$ is the transition density function of $X_t^{\alpha,\beta}$, and to derive asymptotic properties, one may follow one of the four following strategies.

1. Either, $p^{\alpha,\beta}$ is explicit (since $X^{\alpha,\beta}$ has a Gaussian law e.g.) and the computations can follow a more or less classical routine: in this way, one can prove that the LAN property holds true for the Ornstein-Uhlenbeck processes (see Jacod [13]).

2. Either, one assumes that some specific estimates on $p^{\alpha,\beta}(t,x,y)$, its derivatives and some integrals involving these quantities are satisfied, and asymptotic properties may be deduced (see Genon-Catalot et al. [8] when the observation is restricted to $[0,1]$). But in general, the validity of these estimates turns to be impossible to check. See also Höpfner, Jacod and Ladelli [12] for the case of Markov chains or Markov step processes.

3. Either, one uses an expansion of $p^{\alpha,\beta}(t,x,y)$ w.r.t. t,α,β up to an appropriate order. This strategy has been successfully performed by Dacunha-Castelle et al. [2] in the case of an one-dimensional elliptic diffusion for estimation purposes, by deriving for $p^{\alpha,\beta}(t,x,y)$ a quasi-explicit representation using a Brownian bridge. This approach has also been used by Donhal [3] to prove the LAMN property when $d=1$, in the asymptotic assumption $n\Delta_n = 1$. For our objective, this strategy has some drawbacks: it essentially restricts the study to the one-dimensional case, since the representation of $p^{\alpha,\beta}(t,x,y)$ can not be extended, using the same arguments, to a general multidimensional situation; even for $d=1$ in our setting, we need to impose a condition on the decreasing rate of Δ_n and more smoothness conditions on the coefficients than needed.

4. Either, and this is the approach we are going to adopt, instead of expanding $\log p^{\alpha,\beta}(\Delta_n,x,y)$ when α,β,x,y are fixed, we first transform $\log(p^{\alpha^0+\frac{u}{\sqrt{n\Delta_n}},\beta^0+\frac{v}{\sqrt{n}}}/p^{\alpha^0,\beta^0})(\Delta_n, X_{k\Delta_n}, X_{(k+1)\Delta_n})$ using a Malliavin calculus integration by parts formula, and then, compute a stochastic expansion. We followed this approach in [9] and derived, in a quite straightforward way, the LAMN property when $n\Delta_n = 1$, generalizing the result of Donhal [3] in a multidimensional setting.

The content of the paper is the following. Our purpose is to derive the LAN property defined in (0.2), when the observation is $(X_{k\Delta_n})_{0\le k\le n}$, with the asymptotic $\Delta_n \to 0$ [1] and $n\Delta_n \to +\infty$: this property is known to be true only in the case of Ornstein-Uhlenbeck processes (see Jacod [13]). We consider different sets of hypotheses on b and S, under which $X^{\alpha,\beta}$ is ergodic. The diffusion coefficient S is always assumed to be uniformly strictly elliptic, whereas various hypotheses on b will be made (including the case of unbounded coefficients). A first set of models (which include the Ornstein-Uhlenbeck processes) is defined in section 1, whereas extensions will be briefly exposed in section 5. In section 1, we state preliminary results concerning estimates on the transition density (their proofs are postponed in Appendix A) and we define the notation used in all the paper. To understand the chain of arguments to get the LAN property, we propose a step-by-step proof. It starts in section 2, where we expose Malliavin calculus ideas, which allow to transform the loglikelihood ratio in a tractable way. Section 3 is devoted to the stochastic expansion of this loglikelihood, to exhib the main order contribution: this is the crucial and technical part of the paper. Then, we state the LAN property in section 4 (see Theorem 4.1) and complete easily its proof, using the results of section 3. The validity of LAN property under other assumptions is discussed in section 5.

1 Assumptions, notations and preliminary results

As usual, we denote the i-th coordinate of the vector u by u_i, or $u_{i,t}$ if $u = u_t$ is time dependent. For smooth functions $g(w)$, $\partial_{w_i} g(w)$ stands for the partial derivative of g w.r.t. w_i.

Now, let us consider Θ_α (resp. Θ_β) an open subset of $\mathbb{R}^{n_\alpha}$ (resp. $\mathbb{R}^{n_\beta}$) for some integer $n_\alpha \ge 1$ (resp. $n_\beta \ge 1$): these two sets are used to define the parameterization of the coefficients of the model of SDE's which we are interested in.

[1] in the sequel, we assume, without restriction, that $\Delta_n \le 1$ for all n

Let $b(\alpha, x)$ be a map from $\Theta_\alpha \times \mathbb{R}^d$ into $\mathbb{R}^d$, and $S(\beta, x)$ a map from $\Theta_\beta \times \mathbb{R}^d$ into $\mathbb{R}^d \otimes \mathbb{R}^d$. For fixed α and β, these maps as function of x are supposed to be globally Lipschitz, so that there is an unique strong solution $(X_t^{\alpha,\beta})_{t\geq 0}$ to the homogeneous stochastic differential equation

$$X_t^{\alpha,\beta} = x_0 + \int_0^t b(\alpha, X_s^{\alpha,\beta})\, ds + \int_0^t S(\beta, X_s^{\alpha,\beta})\, dB_s, \tag{1.3}$$

where $(B_t)_{t\geq 0}$ is a standard Brownian motion in $\mathbb{R}^d$ (with $(\mathcal{G}_t)_{t\geq 0}$ its usual filtration) and x_0 is a deterministic initial condition. In the sequel, the indices α, β, x_0 in $\mathbb{E}_{x_0}^{\alpha,\beta}$ stand in reference for the expectation under the law of the diffusion $X^{\alpha,\beta}$ starting at x_0. When no confusion is possible, we may simply write X_t instead of $X_t^{\alpha,\beta}$.

In order to get asymptotic properties on the likelihood ratio, it is necessary to put additional regularity conditions on the coefficients. To include the important case of Ornstein-Uhlenbeck processes, we allow the drift coefficient to be unbounded: this hypothesis will lead to technical difficulties, mainly concerning some estimates on the transition density (see Proposition 1.2 below). The easier case of bounded drift coefficient is discussed in section 5. In the sequel, we assume the following hypotheses.

Assumption (R):

1. the functions $b(\alpha, x)$ and $S(\beta, x)$ are [2] of class $C^{1+\gamma}$ w.r.t. (α, x) or (β, x), for some $\gamma \in (0, 1)$.
2. each partial derivative $\partial_{\alpha_i} b(\alpha, x)$, $\partial_{x_i} b(\alpha, x)$, $\partial_{\beta_i} S(\beta, x)$, $\partial_{x_i} S(\beta, x)$ is of class C^1 w.r.t. x.
3. the following estimates hold:
   (a) $|b(\alpha, x)| \leq c(1 + |x|)$ and $|\partial_{x_i} b(\alpha, x)| + |S(\beta, x)| + |\partial_{x_i} S(\beta, x)| \leq c$
   (b) $|g(., x)| \leq c(1 + |x|^q)$ for $g = \partial_{\alpha_i} b,\ \partial^2_{x_i,x_j} b,\ \partial^2_{x_i,\alpha_j} b,\ \partial_{\beta_i} S,\ \partial^2_{x_i,x_j} S$ or $\partial^2_{x_i,\beta_j} S$
   (c) $\dfrac{|\partial_{\alpha_i} b(\alpha, x) - \partial_{\alpha_i} b(\alpha', x)|}{|\alpha - \alpha'|^\gamma} + \dfrac{|\partial_{\beta_i} S(\beta, x) - \partial_{\beta_i} S(\beta', x)|}{|\beta - \beta'|^\gamma} \leq c(1 + |x|^q)$

   for some positive constants c and q, independent of $(\alpha, \alpha', \beta, \beta', x) \in \Theta_\alpha^2 \times \Theta_\beta^2 \times \mathbb{R}^d$.

To ensure the ergodicity of the process (1.3), we impose two conditions derived from Has'minskii [11]: the drift coefficient b is strongly re-entrant and the matrix S is strongly non degenerate.

Assumption (D): one has $\forall(\alpha, x) \in \Theta_\alpha \times \mathbb{R}^d \quad b(\alpha, x).x \leq -c_0\, |x|^2 + K$ for some constant $c_0 > 0$.

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

$$\forall(\beta, x) \in \Theta_\beta \times \mathbb{R}^d \qquad \frac{1}{c_1}\, \mathrm{I}_d(x) \leq S(\beta, x) \leq c_1\, \mathrm{I}_d(x)$$

for some constant $c_1 \geq 1$.

[2]as usual, 'f is of class $C^{1+\gamma}$' means that f is of class C^1 and its partial derivatives are γ-Hölder continuous

Example 1.1. *Set* $\Theta_\alpha = (\alpha_1^{min}, \alpha_1^{max}) \times K$ *(K is some open bounded subset of $\mathbb{R}$) and* $\Theta_\beta = (\beta^{min}, \beta^{max})$. *Then, the linear Ornstein-Uhlenbeck process* $X_t^{\alpha,\beta} = x_0 + \int_0^t (\alpha_1 X_s^{\alpha,\beta} + \alpha_2)\, ds + \beta\, B_t$ *fulfills the above assumptions when* $\alpha_1^{max} < 0$ *and* $\beta^{min} > 0$.

Under assumptions **(R)**, **(D)** and **(E)**, the process $X^{\alpha,\beta}$ has an unique invariant probability measure: we denote it by $\mu^{\alpha,\beta}$ and we are going to prove that it has squared exponential moments.

Proposition 1.1. *Under* **(R)**, **(D)** *and* **(E)**, *there is a constant* $C_e > 0$ *such that*

1. *for any* $C \in [0, C_e)$ *and for any* $\lambda > 0$, *one has:*

$$\forall\, t \geq 0 \qquad \mathbb{E}_{x_0}^{\alpha,\beta} \exp\left(C|X_t|^2\right) \leq \exp\left(C|x_0|^2\right)\, \exp\left(-\lambda\, t\right) + K, \tag{1.4}$$

   for some constant $K = K(C, \lambda)$

2. *for any* $C < C_e$, *one has*

$$\int_{\mathbb{R}^d} \exp\left(C|x|^2\right)\, \mu^{\alpha,\beta}(dx) < \infty. \tag{1.5}$$

Proof. Set $f(x) = \exp\left(C|x|^2\right)$ and denote by $L^{\alpha,\beta}$ the infinitesimal generator of the diffusion $X^{\alpha,\beta}$. From assumptions **(R)** and **(D)** and putting $C_e = c_0/c_1^2$, one easily deduces:

$$\begin{aligned} L^{\alpha,\beta} f(x) &= 2Cf(x) \sum_i b_i(\alpha, x) x_i + 2C^2 f(x) \sum_{i,j} (S^2)_{i,j}(\beta, x) x_i x_j + Cf(x) \sum_i (S^2)_{i,i}(\beta, x) \\ &\leq 2C\left(-c_0 + Cc_1^2\right)|x|^2 f(x) + Kf(x) = -c_0'|x|^2 f(x) + Kf(x), \end{aligned}$$

using for the last inequality $C \in [0, C_e)$, so that $c_0' > 0$.
Now, it readily follows that $L^{\alpha,\beta} f(x) \leq -\lambda f(x) + K'(\lambda)$ for any $\lambda > 0$; thus, if $g(t) = \mathbb{E}_{x_0}^{\alpha,\beta}(f(X_t))$, one has $g'(t) \leq -\lambda g(t) + K'(\lambda)$. To derive (1.4), compute $(g(t)\exp(\lambda t))'$ and use the previous inequality. We deduce (1.5) from (1.4). Let U be a compact subset of $\mathbb{R}^d$: from the ergodic theorem, one gets

$$\int_{\mathbb{R}^d} \exp\left(C|x|^2\right)\, \mathbf{1}_{x\in U}\, \mu^{\alpha,\beta}(dx) = \lim_{t\to+\infty} \mathbb{E}_{x_0}^{\alpha,\beta}\left(\exp\left(C|X_t|^2\right)\, \mathbf{1}_{X_t\in U}\right) \leq K,$$

for some constant K independent of U. Now, let U increase to $\mathbb{R}^d$ and apply monotone convergence theorem, to complete the proof.

$\square$

Under **(R)** and **(E)**, the law of $X_t^{\alpha,\beta}$ $(t > 0)$ conditionally on $X_0^{\alpha,\beta} = x$ has a strictly positive transition density $p^{\alpha,\beta}(t, x, y)$, which is, in particular, differentiable w.r.t. α and β (see Proposition 2.2 below). Furthermore, $p^{\alpha,\beta}(t, x, y)$ and its derivatives satisfy the following estimates.

Proposition 1.2. *Assume* **(R)** *and* **(E)**. *There exist constants* $c > 1$ *and* $K > 1$ *s.t.*

$$p^{\alpha,\beta}(t, x, y) \leq \frac{K}{t^{d/2}} \exp\left(-\frac{|x-y|^2}{c\, t}\right) \exp\left(c\, t\, |x|^2\right), \tag{1.6}$$

$$p^{\alpha,\beta}(t, x, y) \geq \frac{1}{K\, t^{d/2}} \exp\left(-c\frac{|x-y|^2}{t}\right) \exp\left(-c\, t\, |x|^2\right), \tag{1.7}$$

and for any $\nu > 1$, there exist other constants $c > 1$, $K > 1$, $q > 0$ s.t.

$$\mathbb{E}_x^{\overline{\alpha},\overline{\beta}} \left| \frac{\partial_{\alpha_i} p^{\alpha,\beta}}{p^{\alpha,\beta}}(t,x,X_t) \right|^\nu \leq K\, t^{\nu/2} \exp\left(c\, t\, |x|^2\right) \, (1+|x|)^q, \tag{1.8}$$

$$\mathbb{E}_x^{\overline{\alpha},\overline{\beta}} \left| \frac{\partial_{\beta_j} p^{\alpha,\beta}}{p^{\alpha,\beta}}(t,x,X_t) \right|^\nu \leq K \exp\left(c\, t\, |x|^2\right) \, (1+|x|)^q, \tag{1.9}$$

for $0 < t \leq 1$, $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$, $1 \leq i \leq n_\alpha$, $1 \leq j \leq n_\beta$ and $(\alpha, \overline{\alpha}, \beta, \overline{\beta}) \in \Theta_\alpha \times \Theta_\alpha \times \Theta_\beta \times \Theta_\beta$.

Analogous bounds for $|\partial_{\alpha_i} p^{\alpha,\beta}(t,x,y)|$ and $|\partial_{\beta_j} p^{\alpha,\beta}(t,x,y)|$ are also available, but we will not use them in the sequel. To derive estimates (1.8) and (1.9), we somehow exploit Malliavin calculus representations which we introduce in Section 2 below: so, we admit for a while this proposition, the proof being postponed in Appendix A.

As far as the author knows, these estimates seem to be new in the context of unbounded drift and bounded diffusion coefficients. Actually, when the functions b and S (and some of their derivatives) are bounded, Gaussian type bounds (i.e. of the form $\frac{K}{t^\mu}\exp(-\frac{|x-y|^2}{c\,t}))$ for p and its derivatives are available (see e.g. Theorem 4.5, Friedman [5]), whereas when b and S have a linear growth (think of the geometric Brownian motion e.g.), the upper bounds are not of Gaussian type, but only decreasing faster than any polynomials.
Here, the boundedness of the diffusion coefficient enables to keep Gaussian bounds, up to the factor $\exp\left(\pm c\, t\, |x|^2\right)$. Actually, this latter term is unavoidable. Indeed, consider again the Ornstein-Uhlenbeck process from Example 1.1: one has $p^{\alpha_1,0,1}(t,x,x) \underset{t\to 0}{\sim} \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{1}{2}x^2\alpha_1^2 t\right)$, estimate which should be compared to inequality (1.7).

Notation.
In all the paper, the multi-index of parameters $(\alpha_1, \cdots, \alpha_{n_\alpha}, \beta_1, \cdots, \beta_{n_\beta})$ is going to be simply denoted by (α, β), and $(\alpha^0, \beta^0) = (\alpha_1^0, \cdots, \alpha_{n_\alpha}^0, \beta_1^0, \cdots, \beta_{n_\beta}^0) \in \Theta_\alpha \times \Theta_\beta$ corresponds as usual to the true value of the parameters. Besides, (α, β) might be a row vector as well a column vector: we will not distinguish the notation, since in the further contexts, no confusion will be possible.
To define the local likelihood ratio around the parameter (α^0, β^0), we fix $u \in \mathbb{R}^{n_\alpha}$, $v \in \mathbb{R}^{n_\beta}$ and set

$$(\alpha^+, \beta^+) := \left(\alpha^0 + \frac{u}{\sqrt{n\Delta_n}}, \beta^0 + \frac{v}{\sqrt{n}}\right) = \left(\alpha_1^0 + \frac{u_1}{\sqrt{n\Delta_n}}, \cdots, \alpha_{n_\alpha}^0 + \frac{u_{n_\alpha}}{\sqrt{n\Delta_n}}, \beta_1^0 + \frac{v_1}{\sqrt{n}}, \cdots, \beta_{n_\beta}^0 + \frac{v_{n_\beta}}{\sqrt{n}}\right).$$

Our main issue is to study the weak convergence (under $\mathbb{P}^{\alpha^0,\beta^0}$ and under the assumptions $\Delta_n \to 0$, $n\Delta_n \to \infty$) of the local likelihood ratio

$$Z_n := \frac{d\mathbb{P}_n^{\alpha^+,\beta^+}}{d\mathbb{P}_n^{\alpha^0,\beta^0}}\left((X_{k\Delta_n})_{0\leq k\leq n}\right),$$

or the convergence of its logarithm $z_n = \log(Z_n)$, which can be rewritten using the transition densities as:

$$z_n = \sum_{k=0}^{n-1} \log\left(\frac{p^{\alpha^+,\beta^+}}{p^{\alpha^0,\beta^0}}\right)(\Delta_n, X_{k\Delta_n}, X_{(k+1)\Delta_n}). \tag{1.10}$$

But to deal with some perturbations around (α^0,β^0), we adopt more specific notation:

$$\begin{aligned}
(\alpha_i^+,\beta^+) &= \left(\alpha_1^0,\cdots,\alpha_{i-1}^0,\alpha_i^0+\frac{u_i}{\sqrt{n\Delta_n}},\cdots,\alpha_{n_\alpha}^0+\frac{u_{n_\alpha}}{\sqrt{n\Delta_n}},\beta_1^0+\frac{v_1}{\sqrt{n}},\cdots,\beta_{n_\beta}^0+\frac{v_{n_\beta}}{\sqrt{n}}\right)\\
(\alpha_i(l),\beta^+) &= \left(\alpha_1^0,\cdots,\alpha_{i-1}^0,\alpha_i^0+l\frac{u_i}{\sqrt{n\Delta_n}},\alpha_{i+1}^0+\frac{u_{i+1}}{\sqrt{n\Delta_n}},\cdots,\alpha_{n_\alpha}^0+\frac{u_{n_\alpha}}{\sqrt{n\Delta_n}},\beta_1^0+\frac{v_1}{\sqrt{n}},\cdots,\beta_{n_\beta}^0+\frac{v_{n_\beta}}{\sqrt{n}}\right)\\
(\alpha,\beta_i^+) &= \left(\alpha_1^0,\cdots,\alpha_{n_\alpha}^0,\beta_1^0,\cdots,\beta_{i-1}^0,\beta_i^0+\frac{v_i}{\sqrt{n}},\cdots,\beta_{n_\beta}^0+\frac{v_{n_\beta}}{\sqrt{n}}\right)\\
(\alpha,\beta_i(l)) &= \left(\alpha_1^0,\cdots,\alpha_{n_\alpha}^0,\beta_1^0,\cdots,\beta_{i-1}^0,\beta_i^0+l\,\frac{v_i}{\sqrt{n}},\beta_{i+1}^0+\frac{v_{i+1}}{\sqrt{n}},\cdots,\beta_{n_\beta}^0+\frac{v_{n_\beta}}{\sqrt{n}}\right).
\end{aligned}$$

We also introduce the mean vector and the covariance matrix of $X_{\Delta_n}^{\alpha,\beta}$:

$$\begin{aligned}
m^{\alpha,\beta}(x) &= \left(m_i^{\alpha,\beta}(x)\right)_i = \left(\mathbb{E}_x^{\alpha,\beta}[X_{i,\Delta_n}]\right)_i\\
V^{\alpha,\beta}(x) &= \left(V_{i,j}^{\alpha,\beta}(x)\right)_{i,j} = \left(\mathbb{E}_x^{\alpha,\beta}[X_{i,\Delta_n}-m_i^{\alpha,\beta}(x)][X_{j,\Delta_n}-m_j^{\alpha,\beta}(x)]\right)_{i,j}
\end{aligned}$$

We may write $g(n,x,\alpha,\beta)=R(\epsilon_n,x)$ if the function g satisfies the estimate $|g(n,x,\alpha,\beta)|\le K(1+|x|^q)\epsilon_n$, for some positive constants K and q, independent of x, n, $\alpha\in\Theta_\alpha$ and $\beta\in\Theta_\beta$.
Besides, the notation K will be kept for all finite positive constants, (independent of x, n, α, β and so on), which will appear in proofs.

2 Transformation of the log-likelihood ratio using Malliavin calculus

In this section, we present the methodology to derive the convergence of the local log-likelihood ratio: the main new idea is to use Malliavin calculus techniques to rewrite this ratio in a tractable way. This strategy has already been performed in Gobet [9] and we briefly expose it in this new setting.

Since one may write $\frac{p^{\alpha^+,\beta^+}}{p^{\alpha^0,\beta^0}}=\frac{p^{\alpha_1^+,\beta^+}}{p^{\alpha_2^+,\beta^+}}\cdots\frac{p^{\alpha_i^+,\beta^+}}{p^{\alpha_{i+1}^+,\beta^+}}\cdots\frac{p^{\alpha_{n_\alpha}^+,\beta^+}}{p^{\alpha,\beta^+}}\,\frac{p^{\alpha,\beta_1^+}}{p^{\alpha,\beta_2^+}}\cdots\frac{p^{\alpha,\beta_j^+}}{p^{\alpha,\beta_{j+1}^+}}\cdots\frac{p^{\alpha,\beta_{n_\beta}^+}}{p^{\alpha^0,\beta^0}}$, one easily deduces, using the smoothness property of $p^{\alpha,\theta}$, that equation (1.10) can be transformed as

$$(2.11)\quad z_n=\sum_{k=0}^{n-1}\zeta_k^{\alpha_1}+\cdots+\sum_{k=0}^{n-1}\zeta_k^{\alpha_i}\cdots+\sum_{k=0}^{n-1}\zeta_k^{\alpha_{n_\alpha}}+\sum_{k=0}^{n-1}\zeta_k^{\beta_1}+\cdots+\sum_{k=0}^{n-1}\zeta_k^{\beta_j}\cdots+\sum_{k=0}^{n-1}\zeta_k^{\beta_{n_\beta}},$$

where

$$(2.12)\qquad \zeta_k^{\alpha_i} = \frac{u_i}{\sqrt{n\Delta_n}}\int_0^1 dl\ \frac{\partial_{\alpha_i}p^{\alpha_i(l),\beta^+}}{p^{\alpha_i(l),\beta^+}}(\Delta_n,X_{k\Delta_n},X_{(k+1)\Delta_n}),$$

$$(2.13)\qquad \zeta_k^{\beta_j} = \frac{v_j}{\sqrt{n}}\int_0^1 dl\ \frac{\partial_{\beta_j}p^{\alpha,\beta_j(l)}}{p^{\alpha,\beta_j(l)}}(\Delta_n,X_{k\Delta_n},X_{(k+1)\Delta_n}).$$

The core of the analysis of the weak convergence of z_n is of course based on a good understanding of the stochastic behavior of $\zeta_k^{\alpha_i}$ and $\zeta_k^{\beta_j}$, which is going to be analyzed through some stochastic expansions. To this purpose, the first step of this program is to rewrite $\frac{\partial_{\alpha_i}p^{\alpha,\beta}}{p^{\alpha,\beta}}(T,x,y)$ and $\frac{\partial_{\beta_j}p^{\alpha,\beta}}{p^{\alpha,\beta}}(T,x,y)$ as a conditional expectation, using Malliavin calculus. For this, we need to introduce the material necessary to our computations (for more details, see Nualart [21]).

2.1 Basic facts on Malliavin calculus

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

Definition 2.1. *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* Dom(δ)*) is the set of processes $u \in L_2([0,T]\times\Omega, \mathbb{R}^d)$ such that $\forall F \in \mathbb{D}^{1,2}$, one has $\left|\mathbb{E}\left(\int_0^T \mathcal{D}_t F \,.\, u_t\, dt\right)\right| \leq c(u)\, \|F\|_2$.*
2. *if u belongs to* Dom(δ)*, then $\delta(u)$ is the element of $L_2(\Omega)$ characterized by the integration by parts formula: $\forall F \in \mathbb{D}^{1,2}$, $\quad \mathbb{E}(F\,\delta(u)) = \mathbb{E}\left(\int_0^T \mathcal{D}_t F \,.\, u_t\, dt\right)$.*

We now state some properties of the Skorohod integral, which are going to be useful in the sequel:

Proposition 2.1.

1. *For any $p > 1$, the space of weakly differentiable H-valued variables $\mathbb{D}^{1,p}(H)$ belongs to* Dom(δ) *and one has*

$$\|\delta(u)\|_p \leq c_p\, \left(\|u\|_{L_p(\Omega,H)} + \|\mathcal{D}u\|_{L_p(\Omega,H\otimes H)}\right). \tag{2.14}$$

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

$$\delta(F\,u) = F\,\delta(u) - \int_0^T \mathcal{D}_t F \,.\, u_t\, dt, \tag{2.15}$$

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

2.2 Transformation of $\frac{\partial_{\alpha_i} p^{\alpha,\beta}}{p^{\alpha,\beta}}(T,x,y)$ and $\frac{\partial_{\beta_j} p^{\alpha,\beta}}{p^{\alpha,\beta}}(T,x,y)$

To allow some Malliavin calculus computations on transition densities while avoiding some confusion with the observed process (1.3) generated by the Brownian motion $(B_t)_{t\geq 0}$, we consider an independent

Brownian motion $(W_t)_{t\geq 0}$ (with its usual filtration $(\mathcal{F}_t)_{t\geq 0}$) to which we associate an independent copy of $X^{\alpha,\beta}$ (still denoted by $X^{\alpha,\beta}$), which consequently solves

$$X_t^{\alpha,\beta} = x + \int_0^t b(\alpha, X_s^{\alpha,\beta})\, ds + \sum_{l=1}^{d} \int_0^t S_l(\beta, X_s^{\alpha,\beta})\, dW_{l,s}, \tag{2.16}$$

where S_l is the l-th column vector of S.
Since $b(\alpha, x)$ and $S(\beta, x)$ are assumed to be of class $C^{1+\gamma}$, $X_t^{\alpha,\beta}$ is differentiable as a function of x, α and β (see Kunita [17]), so that we can introduce its flow, i.e. the Jacobian matrix $Y_t^{\alpha,\beta} := \nabla_x X_t^{\alpha,\beta}$, and its derivative w.r.t. α_i (resp. β_j) denoted by $\partial_{\alpha_i} X_t^{\alpha,\beta}$ (resp. $\partial_{\beta_j} X_t^{\alpha,\beta}$). This defines new processes, which solve a system of SDE's:

$$\begin{aligned}
Y_t^{\alpha,\beta} &= \mathrm{I}_d + \int_0^t \nabla_x b(\alpha, X_s^{\alpha,\beta})\, Y_s^{\alpha,\beta}\, ds + \sum_{l=1}^{d} \int_0^t \nabla_x S_l(\beta, X_s^{\alpha,\beta})\, Y_s^{\alpha,\beta}\, dW_{l,s}, \\
\partial_{\alpha_i} X_t^{\alpha,\beta} &= \int_0^t \left(\partial_{\alpha_i} b(\alpha, X_s^{\alpha,\beta}) + \nabla_x b(\alpha, X_s^{\alpha,\beta})\, \partial_{\alpha_i} X_s^{\alpha,\beta} \right) ds \\
&\quad + \sum_{l=1}^{d} \int_0^t \nabla_x S_l(\beta, X_s^{\alpha,\beta})\, \partial_{\alpha_i} X_s^{\alpha,\beta}\, dW_{l,s}, \qquad (2.17) \\
\partial_{\beta_j} X_t^{\alpha,\beta} &= \int_0^t \nabla_x b(\alpha, X_s^{\alpha,\beta})\, \partial_{\beta_j} X_s^{\alpha,\beta}\, ds \\
&\quad + \sum_{l=1}^{d} \int_0^t \left(\partial_{\beta_j} S_l(\beta, X_s^{\alpha,\beta}) + \nabla_x S_l(\beta, X_s^{\alpha,\beta})\, \partial_{\beta_j} X_s^{\alpha,\beta} \right) dW_{l,s}. \qquad (2.18)
\end{aligned}$$

Under **(R)**, for any $t \geq 0$, the random variables $X_t^{\alpha,\beta}$, $Y_t^{\alpha,\beta}$, $(Y_t^{\alpha,\beta})^{-1}$, $(\partial_{\alpha_i} X_t^{\alpha,\beta})_i$ and $(\partial_{\beta_j} X_t^{\alpha,\beta})_j$ belong to $\mathbb{D}^{1,p}$ for any $p \geq 1$ (see Nualart [21], Section 2.2). Besides, the following crude estimates hold true:

$$\mathbb{E}_x^{\alpha,\beta}\left(\sup_{0\leq t\leq 1} \|Z_t\|^p \right) + \sup_{r\in[0,1]} \mathbb{E}_x^{\alpha,\beta}\left(\sup_{r\leq t\leq 1} \|\mathcal{D}_r Z_t\|^p \right) = R(1,x) \tag{2.19}$$

for $Z_t = X_t^{\alpha,\beta}$, $Y_t^{\alpha,\beta}$ or $(Y_t^{\alpha,\beta})^{-1}$. We now state the useful result for the analysis of the loglikelihood.

Proposition 2.2. *(Gobet [9], Proposition 4.1). Assume* **(R)** *and* **(E)** *and set* $T \in (0,1]$. *For* $1 \leq l_1 \leq d$, *let us define* $U_{l_1} = (U_{l_1,t})_{0\leq t\leq T}$ *the* $\mathbb{R}^d$*-valued process whose* l_2*-th component is equal to* $U_{l_1 l_2,t} = \left(S^{-1}(\beta, X_t^{\alpha,\beta})\, Y_t^{\alpha,\beta}\, (Y_T^{\alpha,\beta})^{-1} \right)_{l_1,l_2}$. *Then, one has*

$$\begin{aligned}
\frac{\partial_{\alpha_i} p^{\alpha,\beta}}{p^{\alpha,\beta}}(T,x,y) &= \frac{1}{T}\, \mathbb{E}_x^{\alpha,\beta}\left[\sum_{l_1=1}^{d} \delta(\partial_{\alpha_i} X_{l_1,T}^{\alpha,\beta}\, U_{l_1}) \mid X_T^{\alpha,\beta} = y \right], \\
\frac{\partial_{\beta_j} p^{\alpha,\beta}}{p^{\alpha,\beta}}(T,x,y) &= \frac{1}{T}\, \mathbb{E}_x^{\alpha,\beta}\left[\sum_{l_1=1}^{d} \delta(\partial_{\beta_j} X_{l_1,T}^{\alpha,\beta}\, U_{l_1}) \mid X_T^{\alpha,\beta} = y \right].
\end{aligned}$$

3 Expansion of the local log-likelihood ratio

From Proposition 2.2, each random variable $\zeta_k^{\alpha_i}$ (or $\zeta_k^{\beta_j}$) can be rewritten as

$$\zeta_k^{\alpha_i} = \frac{u_i}{\sqrt{n\Delta_n}} \int_0^1 dl \, \frac{1}{\Delta_n} \mathbb{E}_{X_{k\Delta_n}}^{\alpha_i(l),\beta^+} \left[H_n^{\alpha_i(l),\beta^+} \mid X_{\Delta_n}^{\alpha_i(l),\beta^+} = X_{(k+1)\Delta_n} \right],$$

for some random variable $H_n^{\alpha_i(l),\beta^+}$. To derive the convergence $\sum_{k=0}^{n-1} \zeta_k^{\alpha_i}$ under $\mathbb{P}^{\alpha^0,\beta^0}$, we may apply a classical convergence theorem for triangular arrays of random variables, by checking the convergence of the sum of some conditional moments, e.g. $\sum_{k=0}^{n-1} \mathbb{E}^{\alpha^0,\beta^0}\left[\zeta_k^{\alpha_i} | \mathcal{G}_{k\Delta_n}\right]$: the fact that the expectations outside and inside refer to different probability measures ($\mathbb{P}^{\alpha^0,\beta^0}$ and $\mathbb{P}^{\alpha_i(l),\beta^+}$) is a sizable difficulty.
Our approach to this problem is to perform a stochastic expansion of $H_n^{\alpha_i(l),\beta^+}$ under $\mathbb{P}^{\alpha_i(l),\beta^+}$. The miracle arises from the fact that this random variable is equal at the first order to some function $g(\alpha_i(l), \beta^+, n, X_0^{\alpha_i(l),\beta^+}, X_{\Delta_n}^{\alpha_i(l),\beta^+})$: consequently, its conditional expectation is immediate to compute and thus, the checking of the convergence of the sum of the conditional moments (under $\mathbb{P}^{\alpha^0,\beta^0}$) of $g(\alpha_i(l), \beta^+, n, X_{k\Delta_n}, X_{(k+1)\Delta_n})$ becomes much more easy.
Nevertheless, we have to prove that the remainder terms in these expansions have no contribution in the limit of z_n. For this, it is necessary to obtain some specific results on the convergence in probability of sums of conditional expectations: our crucial tools are Propositions 3.1 and 3.2.

3.1 Some convergence results

The main purpose of this section is to prove the two following Propositions.

Proposition 3.1. *Assume* **(R)**, **(D)** *and* **(E)**. *Set $i \in \{1, \cdots, n_\alpha\}$. Let H be a $\mathcal{F}_{\Delta_n}$-measurable random variable, which satisfies for any $\mu > 1$:*

$$\mathbb{E}_x^{\alpha_i(l),\beta^+}[H] = 0 \qquad \textit{and} \qquad \left(\mathbb{E}_x^{\alpha_i(l),\beta^+}|H|^\mu\right)^{1/\mu} = R({\Delta_n}^{3/2}\epsilon_n, x),$$

for some sequence $\epsilon_n \to 0$. Then, one has

$$\sum_{k=0}^{n-1} \frac{u_i}{\sqrt{n\Delta_n}} \int_0^1 dl \, \frac{1}{\Delta_n} \, \mathbb{E}_{X_{k\Delta_n}}^{\alpha_i(l),\beta^+} \left[H \mid X_{\Delta_n}^{\alpha_i(l),\beta^+} = X_{(k+1)\Delta_n} \right] \overset{\mathbb{P}^{\alpha^0,\beta^0}}{\longrightarrow} 0.$$

Proposition 3.2. *Assume* **(R)**, **(D)** *and* **(E)**. *Set $j \in \{1, \cdots, n_\beta\}$. Let H be a $\mathcal{F}_{\Delta_n}$-measurable random variable, which satisfies for any $\mu > 1$:*

$$\mathbb{E}_x^{\alpha,\beta_j(l)}[H] = 0 \qquad \textit{and} \qquad \left(\mathbb{E}_x^{\alpha,\beta_j(l)}|H|^\mu\right)^{1/\mu} = R(\Delta_n\epsilon_n, x),$$

for some sequence $\epsilon_n \to 0$. Then, one has

$$\sum_{k=0}^{n-1} \frac{v_j}{\sqrt{n}} \int_0^1 dl \, \frac{1}{\Delta_n} \, \mathbb{E}_{X_{k\Delta_n}}^{\alpha,\beta_j(l)} \left[H \mid X_{\Delta_n}^{\alpha,\beta_j(l)} = X_{(k+1)\Delta_n} \right] \overset{\mathbb{P}^{\alpha^0,\beta^0}}{\longrightarrow} 0.$$

Actually, analogous results are proved in Gobet [9] (see Corollary 4.1), but they are inefficient for our purpose. The main difference concerns the assumption on the mean of H, which is taken to be 0 in this paper, whereas in [9], it was dominated by some power of Δ_n. This difference turns out to be crucial, and being a little careful in the proof below, we may note that if the mean of H is only supposed to be of order ${\Delta_n}^\nu$, we can not obtain the result of the Propositions above, unless we impose (as in Kessler [16] and others) some restrictive conditions on the decreasing rate of Δ_n such as $n{\Delta_n}^{\nu'} \to 0$.
In order to prove Propositions 3.1 and 3.2 and further results, we need a classical discrete time ergodic theorem, which following version is adapted from Kessler [16].

Lemma 3.1. *Assume* (**R**), (**D**) *and* (**E**). *There is a constant $C'_e > 0$, such that, if g is a differentiable function satisfying $|g(x)| + |\nabla g(x)| \leq K \exp(C|x|^2)$ with $C < C'_e$, then*

$$\frac{1}{n}\sum_{k=0}^{n-1} g(X_{k\Delta_n}) \overset{\mathbb{P}^{\alpha^0,\beta^0}}{\longrightarrow} \int_{\mathbb{R}^d} g(x)\mu^{\alpha^0,\beta^0}(dx),$$

where the limit above is finite.

Proof. Take $C'_e \leq C_e$ where C_e is defined in Proposition 1.1: the continuous time ergodic theorem ensures that $\frac{1}{n\Delta_n}\int_0^{n\Delta_n} g(X_s)\, ds \overset{\mathbb{P}^{\alpha^0,\beta^0}}{\longrightarrow} \int_{\mathbb{R}^d} g(x)\mu^{\alpha^0,\beta^0}(dx)$. Thus, it is enough to prove that

$$\mathbb{E}_{x_0}^{\alpha^0,\beta^0}\Big|\frac{1}{n\Delta_n}\int_0^{n\Delta_n} g(X_s)\, ds - \frac{1}{n}\sum_{k=0}^{n-1} g(X_{k\Delta_n})\Big| \leq \frac{1}{n\Delta_n}\sum_{k=0}^{n-1}\int_{k\Delta_n}^{(k+1)\Delta_n} \mathbb{E}_{x_0}^{\alpha^0,\beta^0}|g(X_s) - g(X_{k\Delta_n})|\, ds$$

converges to 0. But using standard Itô's calculus, one gets (for some $\lambda > 0$)

$$\mathbb{E}_{x_0}^{\alpha^0,\beta^0}|g(X_s) - g(X_{k\Delta_n})| \leq K\sqrt{\Delta_n}\sqrt{\mathbb{E}_{x_0}^{\alpha^0,\beta^0}\exp(\lambda C|X_{k\Delta_n}|^2)\, \mathbb{E}_{x_0}^{\alpha^0,\beta^0}\exp(\lambda C|X_s|^2)} \leq K\sqrt{\Delta_n},$$

for some new constant K, which is independent of $k\Delta_n$ and s owing the uniform estimates of Proposition 1.1 up to choosing C small enough. The completion of the proof now follows easily.

□

The above Lemma is going to be often combined with the following classical convergence result about triangular arrays of random variables.

Lemma 3.2. *(Genon-Catalot et al. [7], Lemma 9). Let ξ_k^n, U be random variables, with ξ_k^n being $\mathcal{G}_{(k+1)\Delta_n}$-measurable. The two following conditions imply $\sum_{k=0}^{n-1}\xi_k^n \overset{\mathbb{P}}{\to} U$:*

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

Proof of Proposition 3.1. Set $\xi_k^n = \frac{u_i}{\sqrt{n\Delta_n}}\int_0^1 dl\, \frac{1}{\Delta_n}\, \mathbb{E}_{X_{k\Delta_n}}^{\alpha_i(l),\beta^+}\left[H \mid X_{\Delta_n}^{\alpha_i(l),\beta^+} = X_{(k+1)\Delta_n}\right]$: these are $\mathcal{G}_{(k+1)\Delta_n}$-measurable random variables, to which we are going to apply Lemma 3.2.

1-Evaluation of $\mathbb{E}^{\alpha^0,\beta^0}\left[\xi_k^n \mid \mathcal{G}_{k\Delta_n}\right]$. It reduces to evaluate

$$\mathbb{E}^{\alpha^0,\beta^0}\left[\mathbb{E}^{\alpha_i(l),\beta+}_{X_{k\Delta_n}}\left[H \mid X^{\alpha_i(l),\beta+}_{\Delta_n}=X_{(k+1)\Delta_n}\right] \mid \mathcal{G}_{k\Delta_n}\right]=\mathbb{E}^{\alpha_i(l),\beta+}_{X_{k\Delta_n}}\left[H\ \frac{p^{\alpha^0,\beta^0}}{p^{\alpha_i(l),\beta+}}(\Delta_n,X_0,X_{\Delta_n})\right]$$

$$=\mathbb{E}^{\alpha_i(l),\beta+}_{X_{k\Delta_n}}\left[H\ \frac{p^{\alpha_i(l),\beta+}}{p^{\alpha_i(l),\beta+}}(\Delta_n,X_0,X_{\Delta_n})\right. \tag{3.20}$$

$$+H\ \frac{(p^{\alpha^+_{i+1},\beta+}-p^{\alpha_i(l),\beta+})+(p^{\alpha^+_{i+2},\beta+}-p^{\alpha^+_{i+1},\beta+})+\cdots+(p^{\alpha,\beta+}-p^{\alpha^+_{n_\alpha},\beta+})}{p^{\alpha_i(l),\beta+}}(\Delta_n,X_0,X_{\Delta_n}) \tag{3.21}$$

$$\left.+H\ \frac{(p^{\alpha,\beta^+_2}-p^{\alpha,\beta^+_1})+\cdots+(p^{\alpha^0,\beta^0}-p^{\alpha,\beta^+_{n_\beta}})}{p^{\alpha_i(l),\beta+}}(\Delta_n,X_0,X_{\Delta_n})\right] \tag{3.22}$$

The term (3.20) is equal to $\mathbb{E}^{\alpha_i(l),\beta+}_{X_{k\Delta_n}}[H]=0$.
Each difference in (3.21) (strictly speaking, not the first one, but nevertheless, the following arguments also apply to it) is equal to

$$\begin{aligned}&\mathbb{E}^{\alpha_i(l),\beta+}_{X_{k\Delta_n}}\left[H\frac{(p^{\alpha^+_{m+1},\beta+}-p^{\alpha^+_m,\beta+})}{p^{\alpha_i(l),\beta+}}(\Delta_n,X_0,X_{\Delta_n})\right]\\ &=\ -\frac{u_m}{\sqrt{n\Delta_n}}\int_0^1 dl\ \mathbb{E}^{\alpha_i(l),\beta+}_{X_{k\Delta_n}}\left[H\frac{\partial\alpha_m p^{\alpha_m(l),\beta+}}{p^{\alpha_m(l),\beta+}}\frac{p^{\alpha_m(l),\beta+}}{p^{\alpha_i(l),\beta+}}(\Delta_n,X_0,X_{\Delta_n})\right].\end{aligned}$$

Using Hölder's inequality (with ν_1, ν_2 and ν_3 conjugate) and the estimate on $(\mathbb{E}^{\alpha_i(l),\beta+}_{X_{k\Delta_n}}|H|^{\nu_1})^{1/\nu_1}$, the inequality (1.8), upper/lower bounds (1.6) and (1.7), it follows that the r.h.s. of the above equality is bounded by

$$\begin{aligned}&\frac{K}{\sqrt{n\Delta_n}}\times R({\Delta_n}^{3/2}\epsilon_n,X_{k\Delta_n})\times\sqrt{\Delta_n}\ \exp\left(c\ \Delta_n\ |X_{k\Delta_n}|^2\right)\\ \times\ &\left(\int_{\mathbb{R}^d}\frac{1}{{\Delta_n}^{\nu_3 d/2}}\ \mathrm{e}^{-\nu_3\frac{|X_{k\Delta_n}-y|^2}{c\ \Delta_n}+\nu_3 c\ \Delta_n\ |X_{k\Delta_n}|^2}\ \frac{1}{{\Delta_n}^{(1-\nu_3)d/2}}\ \mathrm{e}^{-(1-\nu_3)c\frac{|X_{k\Delta_n}-y|^2}{\Delta_n}-(1-\nu_3)c\ \Delta_n\ |X_{k\Delta_n}|^2)}\ dy\right)^{1/\nu_3}\\ &\le R\left(\frac{{\Delta_n}^{3/2}\epsilon_n}{\sqrt{n}},X_{k\Delta_n}\right)\ \exp\left(c'\ \Delta_n\ |X_{k\Delta_n}|^2\right),\end{aligned}$$

since the integral w.r.t. y is finite as soon as $-\nu_3/c-(1-\nu_3)c<0$: this condition is satisfied up to choosing ν_3 closed to 1, i.e. ν_1 and ν_2 enough large.
Using analogous arguments (and in particular estimate (1.9)), check that each difference in (3.22) satisfies the following inequality

$$\left|\mathbb{E}^{\alpha_i(l),\beta+}_{X_{k\Delta_n}}\left[H\ \frac{(p^{\alpha,\beta^+_{m+1}}-p^{\alpha,\beta^+_m})}{p^{\alpha_i(l),\beta+}}(\Delta_n,X_0,X_{\Delta_n})\right]\right|\le R\left(\frac{{\Delta_n}^{3/2}\epsilon_n}{\sqrt{n}},X_{k\Delta_n}\right)\ \exp\left(c'\ \Delta_n\ |X_{k\Delta_n}|^2\right).$$

Taking into account that $c'\ \Delta_n\le C'_e/2$ for n large enough, one has proved that

$$\left|\mathbb{E}^{\alpha^0,\beta^0}\left[\xi_k^n \mid \mathcal{G}_{k\Delta_n}\right]\right|\le\frac{1}{\sqrt{n\Delta_n}}\frac{1}{\Delta_n}R\left(\frac{{\Delta_n}^{3/2}\epsilon_n}{\sqrt{n}},X_{k\Delta_n}\right)\ \mathrm{e}^{c'\ \Delta_n\ |X_{k\Delta_n}|^2}\le\epsilon_n\times\frac{1}{n}R(1,X_{k\Delta_n})\ \mathrm{e}^{\frac{1}{2}C'_e\ |X_{k\Delta_n}|^2}.$$

Apply Lemma 3.1 to the function $R(1,x)\, e^{\frac{1}{2}C'_e\,|x|^2}$ and conclude that $\sum_{k=0}^{n-1}\mathbb{E}^{\alpha^0,\beta^0}\left[\xi_k^n \mid \mathcal{G}_{k\Delta_n}\right] \overset{\mathbb{P}^{\alpha^0,\beta^0}}{\to} 0$.

2-Evaluation of $\mathbb{E}^{\alpha^0,\beta^0}\left[(\xi_k^n)^2 \mid \mathcal{G}_{k\Delta_n}\right]$. Using repeatedly Jensen's inequality, one has

$$\begin{aligned}\mathbb{E}^{\alpha^0,\beta^0}\left[(\xi_k^n)^2 \mid \mathcal{G}_{k\Delta_n}\right] &\leq \frac{|u_i|^2}{n\Delta_n}\,\frac{1}{{\Delta_n}^2}\int_0^1 dl\; \mathbb{E}^{\alpha_i(l),\beta^+}_{X_{k\Delta_n}}\left[H^2\,\frac{p^{\alpha^0,\beta^0}}{p^{\alpha_i(l),\beta^+}}(\Delta_n, X_0, X_{\Delta_n})\right]\\ &\leq \epsilon_n^2 \times \frac{1}{n}R\left(1, X_{k\Delta_n}\right)\,\exp\left(\frac{1}{2}C'_e\,|X_{k\Delta_n}|^2\right),\end{aligned}$$

where the expectation under $\mathbb{P}^{\alpha_i(l),\beta^+}$ has been evaluated as before, i.e. using Hölder's inequality, the estimate on $(\mathbb{E}^{\alpha_i(l),\beta^+}_{X_{k\Delta_n}}|H|^{2\nu_1})^{1/\nu_1}$ and upper/lower bounds (1.6) and (1.7). Lemma 3.1 completes the proof of $\sum_{k=0}^{n-1}\mathbb{E}^{\alpha^0,\beta^0}\left[(\xi_k^n)^2 \mid \mathcal{G}_{k\Delta_n}\right] \overset{\mathbb{P}^{\alpha^0,\beta^0}}{\to} 0$. Thus, Proposition 3.1 is proved.

The proof of Proposition 3.2 is very similar to the previous one: we omit it.

$\square$

3.2 Stochastic expansion

The objective of this section is to derive some good approximations of the sums $\sum_{k=0}^{n-1}\zeta_k^{\alpha_i}$ and $\sum_{k=0}^{n-1}\zeta_k^{\beta_j}$ from (2.11): as explained before, it consists in performing a stochastic expansion (w.r.t. the small time Δ_n) of the random variables $\sum_{l_1=1}^{d}\delta(\partial_{\alpha_i}X^{\alpha_i(l),\beta^+}_{l_1,\Delta_n}\;U_{l_1})$ and $\sum_{l_1=1}^{d}\delta(\partial_{\beta_j}X^{\alpha,\beta_j(l)}_{l_1,\Delta_n}\;U_{l_1})$ defined in Proposition 2.2. To neglect the contribution of the remainder terms, we apply Propositions 3.1 and 3.2 above. The main difference with what we did in [9] is that we have to keep in mind that these remainder terms have to be centered random variables: this may explain that the next computations are little more intricate than in [9].

3.2.1 Contributions of the drift coefficient

Lemma 3.3. *Assume* **(R)**, **(D)** *and* **(E)**. *Set* $i\in\{1,\cdots,n_\alpha\}$. *If one defines*

$$\hat{\zeta}_k^{\alpha_i} = \frac{1}{\sqrt{n\Delta_n}}\int_0^1 dl\;\partial_{\alpha_i}b(\alpha_i(l),X_{k\Delta_n}).\left[S^{-2}(\beta^+,X_{k\Delta_n})(X_{(k+1)\Delta_n}-m^{\alpha_i(l),\beta^+}(X_{k\Delta_n}))\right],$$

then one has $\displaystyle\sum_{k=0}^{n-1}\zeta_k^{\alpha_i} - u_i\sum_{k=0}^{n-1}\hat{\zeta}_k^{\alpha_i} \overset{\mathbb{P}^{\alpha^0,\beta^0}}{\longrightarrow} 0.$

Proof. As in Proposition 2.2, define $U_{l_1}=(U_{l_1,t})_{0\leq t\leq\Delta_n}$ as the $\mathbb{R}^d$-valued process with component equal to $U_{l_1l_2,t}=[S^{-1}(\beta^+,X_t^{\alpha_i(l),\beta^+})\,Y_t^{\alpha_i(l),\beta^+}\,(Y_{\Delta_n}^{\alpha_i(l),\beta^+})^{-1}]_{l_1,l_2}$, and set $X_0^{\alpha_i(l),\beta^+}=x$.
The above Lemma is proved if one shows that

$$\delta(\partial_{\alpha_i}X^{\alpha_i(l),\beta^+}_{l_1,\Delta_n}\;U_{l_1}) = \Delta_n\partial_{\alpha_i}b_{l_1}(\alpha_i(l),x)\left[S^{-2}(\beta^+,x)(X^{\alpha_i(l),\beta^+}_{\Delta_n}-m^{\alpha_i(l),\beta^+}(x))\right]_{l_1}+H_{l_1}, \tag{3.23}$$

for $l_1\in\{1,\cdots,d\}$, with $\left(\mathbb{E}_x^{\alpha_i(l),\beta^+}|H_{l_1}|^{\mu}\right)^{1/\mu} = R({\Delta_n}^{3/2}\epsilon_n,x)$ for all $\mu>1$ ($\epsilon_n\to 0$).

Indeed, one has that $\mathbb{E}_x^{\alpha_i(l),\beta^+}[H_{l_1}] = 0$ since both other random variables of equality (3.23) are centered under $\mathbb{P}_x^{\alpha_i(l),\beta^+}$. Thus, Proposition 3.1 applies and after a summation over l_1 of equalities (3.23), one gets the result.

Proof of (3.23). Here, for simplicity, if V is a random variable (possibly multidimensional), we use the notation $V = R'(\epsilon_n, x)$ if for any $\mu > 1$, one has $[\mathbb{E}_x^{\alpha_i(l),\beta^+}|V|^{\mu}]^{1/\mu} = R(\epsilon_n, x)$ uniformly in all variables (except x, μ and n). From (2.15), one has:

$$\delta(\partial_{\alpha_i} X_{l_1,\Delta_n}^{\alpha_i(l),\beta^+}\, U_{l_1}) = \partial_{\alpha_i} X_{l_1,\Delta_n}^{\alpha_i(l),\beta^+}\, \delta(U_{l_1}) - \int_0^{\Delta_n} \mathcal{D}_t\, \partial_{\alpha_i} X_{l_1,\Delta_n}^{\alpha_i(l),\beta^+} .\, U_{l_1,t}\, dt. \tag{3.24}$$

1- First of all, we are going to prove that

$$\int_0^{\Delta_n} \mathcal{D}_t\, \partial_{\alpha_i} X_{l_1,\Delta_n}^{\alpha_i(l),\beta^+} .\, U_{l_1,t}\, dt = R'({\Delta_n}^2, x). \tag{3.25}$$

Indeed, standard computations with Gronwall's lemma yield $\sup_{0\leq s\leq\Delta_n} |\partial_{\alpha_i} X_s^{\alpha_i(l),\beta^+}| = R'(\Delta_n, x)$. Thus, deriving from (2.17) the equation solved by $(\mathcal{D}_t\ \partial_{\alpha_i} X_{\Delta_n}^{\alpha_i(l),\beta^+})_{0\leq t\leq\Delta_n}$, one can easily obtained $\mathcal{D}_t\ \partial_{\alpha_i} X_{\Delta_n}^{\alpha_i(l),\beta^+} = R'(\Delta_n, x)$ using the above estimates on $\partial_{\alpha_i} X_s^{\alpha_i(l),\beta^+}$ and (2.19). It remains to take into account estimates (2.19) to complete the proof of (3.25).

2- Second, using standard Itô's calculus, one gets from equation (2.17) that

$$\partial_{\alpha_i} X_{l_1,\Delta_n}^{\alpha_i(l),\beta^+} - \Delta_n \partial_{\alpha_i} b_{l_1}(\alpha_i(l), x) = R'({\Delta_n}^{3/2}, x). \tag{3.26}$$

3- At last, set $\hat{U}_{l_1 l_2,t} = (S^{-1})_{l_1,l_2}(\beta^+, X_t^{\alpha_i(l),\beta^+})$ and write $\delta(U_{l_1}) = \delta(\hat{U}_{l_1}) + \delta(U_{l_1} - \hat{U}_{l_1})$: using (2.14) and estimates (2.19), it readily follows that $\delta(U_{l_1} - \hat{U}_{l_1}) = R'(\Delta_n, x)$.
Furthermore, since $\hat{U}_{l_1}$ is an adapted process, $\delta(\hat{U}_{l_1})$ is simply an Itô integral. The matrix S is invertible, thus one has

$$\begin{aligned} dW_t &= S^{-1}(\beta, X_t^{\alpha,\beta})\, dX_t^{\alpha,\beta} - S^{-1}(\beta, X_t^{\alpha,\beta})\, b(\alpha, X_t^{\alpha,\beta})\, dt \\ &= S^{-1}(\beta, x)\, dX_t^{\alpha,\beta} + (\mathrm{I}_d - S^{-1}(\beta, x) S(\beta, X_t^{\alpha,\beta}))\, dW_t - S^{-1}(\beta, x)\, b(\alpha, X_t^{\alpha,\beta})\, dt, \end{aligned} \tag{3.27}$$

for any (α, β). Consequently, easy computations yield

$$\begin{aligned} \delta(\hat{U}_{l_1}) &= \sum_{l_2=1}^{d} \int_0^{\Delta_n} (S^{-1})_{l_1,l_2}(\beta^+, X_t^{\alpha_i(l),\beta^+})\, dW_{l_2,t} = \sum_{l_2=1}^{d} \int_0^{\Delta_n} (S^{-1})_{l_1,l_2}(\beta^+, x)\, dW_{l_2,t} + R'(\Delta_n, x) \\ &= \sum_{l_3=1}^{d} (S^{-2})_{l_1,l_3}(\beta^+, x) \int_0^{\Delta_n} dX_{l_3,t}^{\alpha_i(l),\beta^+} + R'(\Delta_n, x) \\ &= \left[S^{-2}(\beta^+, x)(X_{\Delta_n}^{\alpha_i(l),\beta^+} - m^{\alpha_i(l),\beta^+}(x))\right]_{l_1} + R'(\Delta_n, x). \end{aligned} \tag{3.28}$$

where we used in particular that $m^{\alpha_i(l),\beta^+}(x) = x + R(\Delta_n, x)$. Combining estimates (3.25), (3.26) and (3.28) in (3.24), one completes the proof of (3.23) taking $\epsilon_n = \sqrt{\Delta_n}$.

$\square$

3.2.2 Contributions of the diffusion coefficient

Now, we focus on the approximation of the sum $\sum_{k=0}^{n-1} \zeta_k^{\beta_j}$ in (2.11).

Lemma 3.4. *Assume* **(R)**, **(D)** *and* **(E)**. *Set* $j \in \{1, \cdots, n_\beta\}$. *If one defines*

$$\begin{aligned}
\hat{\zeta}_k^{\beta_j} &= \frac{1}{\sqrt{n}} \int_0^1 dl \, \frac{1}{\Delta_n} \mathrm{Tr} \Big\{ (\partial_{\beta_j} S \, S^{-3})(\beta_j(l), X_{k\Delta_n}) \\
&\Big[(X_{(k+1)\Delta_n} - m^{\alpha,\beta_j(l)}(X_{k\Delta_n}))(X_{(k+1)\Delta_n} - m^{\alpha,\beta_j(l)}(X_{k\Delta_n}))^* - V^{\alpha,\beta_j(l)}(X_{k\Delta_n}) \Big] \Big\},
\end{aligned}$$

then one has $\displaystyle \sum_{k=0}^{n-1} \zeta_k^{\beta_j} - v_j \sum_{k=0}^{n-1} \hat{\zeta}_k^{\beta_j} \overset{\mathbb{P}^{\alpha^0,\beta^0}}{\longrightarrow} 0.$

Proof. The techniques are very similar to those of Lemma 3.3, thus we expose a shortened proof, voluntarily omitting some details (see also [9], section 4.3 for many analogies).
As before, set $X_0^{\alpha,\beta_j(l)} = x$ and define U_{l_1} as the $\mathbb{R}^d$-valued process with component equal to $U_{l_1 l_2,t} = [S^{-1}(\beta_j(l), X_t^{\alpha,\beta_j(l)}) \, Y_t^{\alpha,\beta_j(l)} \, (Y_{\Delta_n}^{\alpha,\beta_j(l)})^{-1}]_{l_1,l_2}$. The lemma is proved if

$$\begin{aligned}
\text{(3.29)} \quad \delta(\partial_{\beta_j} X_{l_1,\Delta_n}^{\alpha,\beta_j(l)} \, U_{l_1}) &= \Big[\partial_{\beta_j} S \, S^{-1}(\beta_j(l), x)(X_{\Delta_n}^{\alpha,\beta_j(l)} - m^{\alpha,\beta_j(l)}(x)) \Big]_{l_1} \\
&\quad \times \Big[S^{-2}(\beta_j(l), x)(X_{\Delta_n}^{\alpha,\beta_j(l)} - m^{\alpha,\beta_j(l)}(x)) \Big]_{l_1} \\
&\quad - \Big(\partial_{\beta_j} S \, S^{-1}(\beta_j(l), x) V^{\alpha,\beta_j(l)}(x) S^{-2}(\beta_j(l), x) \Big)_{l_1,l_1} + H_{l_1},
\end{aligned}$$

for $l_1 \in \{1, \cdots, d\}$, with $\Big(\mathbb{E}_x^{\alpha,\beta_j(l)} |H_{l_1}|^\mu \Big)^{1/\mu} = R(\Delta_n \epsilon_n, x)$ for all $\mu > 1$ ($\epsilon_n \to 0$).
Indeed, easy algebra in equality (3.29) shows that $\mathbb{E}_x^{\alpha,\beta_j(l)}[H_{l_1}] = 0$: thus, Proposition 3.2 applies. Then, if we sum over l_1 equalities (3.29) and remind of Proposition 2.2, we obtain the result taking into account that for A and B some $d \times d$-matrixes and y some vector of $\mathbb{R}^d$, one has $Ay.By = \mathrm{Tr}(A^* B y y^*)$.

Proof of (3.29). For simplicity, we write $V = R'(\epsilon_n, x)$ if the random variable V satisfies for any $\mu > 1$, $[\mathbb{E}_x^{\alpha,\beta_j(l)} |V|^\mu]^{1/\mu} = R(\epsilon_n, x)$ uniformly in all variables (except x, μ and n). From (2.15), one has:

$$\delta(\partial_{\beta_j} X_{l_1,\Delta_n}^{\alpha,\beta_j(l)} \, U_{l_1}) = \partial_{\beta_j} X_{l_1,\Delta_n}^{\alpha,\beta_j(l)} \delta(U_{l_1}) - \int_0^{\Delta_n} \mathcal{D}_t \partial_{\beta_j} X_{l_1,\Delta_n}^{\alpha,\beta_j(l)} \, . \, U_{l_1,t} \, dt. \tag{3.30}$$

From equation (2.18), it readily follows

$$\begin{aligned}
\partial_{\beta_j} X_{l_1,\Delta_n}^{\alpha,\beta_j(l)} &= \sum_{l_2=1}^{d} \int_0^{\Delta_n} \partial_{\beta_j} S_{l_1,l_2}(\beta_j(l), x) \, dW_{l_2,t} + R'(\Delta_n, x) \\
&= \Big[\partial_{\beta_j} S \, S^{-1}(\beta_j(l), x)(X_{\Delta_n}^{\alpha,\beta_j(l)} - m^{\alpha,\beta_j(l)}(x)) \Big]_{l_1} + R'(\Delta_n, x)
\end{aligned} \tag{3.31}$$

where we used at the last equality the same arguments as for (3.28).
As in the proof of Lemma 3.3, one has

$$\delta(U_{l_1}) = \Big[S^{-2}(\beta_j(l), x)(X_{\Delta_n}^{\alpha,\beta_j(l)} - m^{\alpha,\beta_j(l)}(x)) \Big]_{l_1} + R'(\Delta_n, x). \tag{3.32}$$

Moreover, one checks that

$$\int_0^{\Delta_n} \mathcal{D}_t\partial_{\beta_j} X^{\alpha,\beta_j(l)}_{l_1,\Delta_n} \,.\, U_{l_1,t}\, dt = \int_0^{\Delta_n} \sum_{l_2=1}^{d} \partial_{\beta_j} S_{l_1,l_2}(\beta_j(l), X^{\alpha,\beta_j(l)}_t)(S^{-1})_{l_1,l_2}(\beta_j(l), X^{\alpha,\beta_j(l)}_t)\, dt + R'({\Delta_n}^{3/2}, x)$$
$$= \Delta_n \left[\partial_{\beta_j} S S^{-1}(\beta_j(l), x)\right]_{l_1,l_1} + R'({\Delta_n}^{3/2}, x).$$

Besides, standard computations yield $V^{\alpha,\beta_j(l)}(x) = \Delta_n S^2(\beta_j(l), x) + R({\Delta_n}^{3/2}, x)$, so that one gets

$$\int_0^{\Delta_n} \mathcal{D}_t\partial_{\beta_j} X^{\alpha,\beta_j(l)}_{l_1,\Delta_n} \,.\, U_{l_1,t}\, dt = \Big(\partial_{\beta_j} S\; S^{-1}(\beta_j(l), x) V^{\alpha,\beta_j(l)}(x) S^{-2}(\beta_j(l), x)\Big)_{l_1,l_1} + R'({\Delta_n}^{3/2}, x).$$

Plug this last equality, estimates (3.31) and (3.32) into (3.30) to complete the proof of (3.29). Lemma 3.4 is proved.

□

3.3 About an explicit approximation of the log-likelihood

To conclude this section on the expansion of the local log-likelihood ratio, we would like to give an answer to the following question:

"Which explicit (or quasi-explicit) log-likelihood should we have to consider from the beginning to find the same expansion that those given by Lemmas 3.3 and 3.4 combined with equality (2.11)?"

Reasonable explicit likelihoods can be derived from Gaussian Markov chains and in this setting, it is tempting to consider those given by the Euler scheme: nevertheless, as it is underlined by Kessler [16], it does work only under some restrictive assumptions of the decreasing rate of Δ_n.
To get the ad hoc log-likelihood, let us denote by $(Y^{\alpha,\beta}_k)_{0\leq k\leq n}$ the $\mathbb{R}^d$-valued Gaussian Markov chain, which fits the two first conditional moments of $(X^{\alpha,\beta}_{k\Delta_n})_{0\leq k\leq n}$, i.e. defined by $Y^{\alpha,\beta}_0 = x_0$ and $Y^{\alpha,\beta}_{k+1} = Y^{\alpha,\beta}_k + \epsilon_{k+1}$, where ϵ_{k+1} is a Gaussian random variable, independent of $\epsilon_1, \cdots, \epsilon_k$, with mean equal to $m^{\alpha,\beta}(Y^{\alpha,\beta}_k)$ and variance equal to $V^{\alpha,\beta}(Y^{\alpha,\beta}_k)$.
Under our hypotheses, $V^{\alpha,\beta}(x)$ is invertible and the transition density of $Y^{\alpha,\beta}$ is equal to $q^{\alpha,\beta}(x,y) = \frac{1}{\sqrt{(2\pi)^d \det V^{\alpha,\beta}(x)}} \exp\left(-\frac{1}{2}(y - m^{\alpha,\beta}(x)).[(V^{\alpha,\beta}(x))^{-1}(y - m^{\alpha,\beta}(x))]\right)$. The local log-likelihood ratio function associated to Y, in which we have replaced the observed diffusion process, is thus given by $\overline{z}_n = \sum_{k=0}^{n-1} \log(\frac{q^{\alpha^+,\beta^+}}{q^{\alpha^0,\beta^0}})(X_{k\Delta_n}, X_{(k+1)\Delta_n})$. This quantity (explicit up the knowledge of $m^{\alpha,\beta}$ and $V^{\alpha,\beta}$) is our candidate to give the same limit than the true local log-likelihood ratio z_n, defined in 2.11. Indeed, one can prove that $z_n - \overline{z}_n \overset{\mathbb{P}^{\alpha^0,\beta^0}}{\longrightarrow} 0$. This can be done from Lemmas 3.3 and 3.4: we omit the details of the computations, which are somehow standard since everything is explicit.

Of course, this result is not surprising: it confirms in some sense that the approach of Kessler [16] was appropriate. Actually, it is not very interesting to obtain the result now, while we have almost finished to prove the LAN property: it would have been more efficient to have this approximation result from the beginning, but we do not have good ideas to obtain it by direct arguments.

4 LAN property

4.1 Statement of the result

The main result of the paper is:

Theorem 4.1. *Under* **(R)**, **(D)** *and* **(E)**, *one has*

$$\log\left(\frac{d\mathbb{P}_n^{\alpha^0+\frac{u}{\sqrt{n\Delta_n}},\beta^0+\frac{v}{\sqrt{n}}}}{d\mathbb{P}_n^{\alpha^0,\beta^0}}\right)((X_{k\Delta_n})_{0\le k\le n}) \overset{\mathcal{L}(\mathbb{P}^{\alpha^0,\beta^0})}{\longrightarrow} \begin{pmatrix} u \\ v \end{pmatrix}.\mathcal{N}^{\alpha^0,\beta^0} - \frac{1}{2}\begin{pmatrix} u \\ v \end{pmatrix}.\Gamma^{\alpha^0,\beta^0}\begin{pmatrix} u \\ v \end{pmatrix},$$

where $\mathcal{N}^{\alpha^0,\beta^0}$ *is a centered* $\mathbb{R}^{n_\alpha+n_\beta}$*-valued Gaussian variable, with covariance matrix*

$$\Gamma^{\alpha^0,\beta^0} = \left(\begin{array}{c|c} \Gamma_b^{\alpha^0,\beta^0} & 0 \\ \hline 0 & \Gamma_S^{\alpha^0,\beta^0} \end{array}\right)$$

where the elements of matrix $\Gamma_b^{\alpha^0,\beta^0} \in \mathbb{R}^{n_\alpha}\otimes\mathbb{R}^{n_\alpha}$ *and* $\Gamma_S^{\alpha^0,\beta^0} \in \mathbb{R}^{n_\beta}\otimes\mathbb{R}^{n_\beta}$ *are given by*

$$(\Gamma_b^{\alpha^0,\beta^0})_{i,j} = \int_{R^d} \partial_{\alpha_i} b(\alpha^0,x).[S^{-2}(\beta^0,x)\partial_{\alpha_j} b(\alpha^0,x)]\mu^{\alpha^0,\beta^0}(dx),$$

$$(\Gamma_S^{\alpha^0,\beta^0})_{i,j} = 2\int_{R^d} \mathrm{Tr}\left[\partial_{\beta_i} S(\beta^0,x)S^{-1}(\beta^0,x)\partial_{\beta_j} S(\beta^0,x)S^{-1}(\beta^0,x)\right]\mu^{\alpha^0,\beta^0}(dx).$$

First, it is worth noticing that $\Gamma_b^{\alpha^0,\beta^0}$ and $\Gamma_S^{\alpha^0,\beta^0}$ are the asymptotic Fisher information matrixes for the continuous time diffusion (see Prakasa Rao [22], Dacunha-Castelle et al. [2], Florens-Zmirou [4], Genon-Catalot [6], Yoshida [23], Kessler [16]). Second, $\mathcal{N}^{\alpha^0,\beta^0}$ has no correlation between the components involving a perturbation on the drift coefficient and a perturbation on the diffusion coefficient: the efficient estimation of the drift and diffusion parameters are asymptotically independent (see Florens-Zmirou [4], Yoshida [23], Kessler [16]).

4.2 Proof

We are going to prove the following estimates:

$$\text{(4.33)}\qquad \sum_{k=0}^{n-1} \mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\alpha_i} \mid \mathcal{G}_{k\Delta_n}] \overset{\mathbb{P}^{\alpha^0,\beta^0}}{\longrightarrow} -\frac{1}{2}u_i(\Gamma_b^{\alpha^0,\beta^0})_{i,i} - u_{i+1}(\Gamma_b^{\alpha^0,\beta^0})_{i,i+1} - \cdots - u_{n_\alpha}(\Gamma_b^{\alpha^0,\beta^0})_{i,n_\alpha},$$

$$\text{(4.34)}\qquad \sum_{k=0}^{n-1} \mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\alpha_i}\,\hat{\zeta}_k^{\alpha_j} \mid \mathcal{G}_{k\Delta_n}] - \mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\alpha_i} \mid \mathcal{G}_{k\Delta_n}]\,\mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\alpha_j} \mid \mathcal{G}_{k\Delta_n}] \overset{\mathbb{P}^{\alpha^0,\beta^0}}{\longrightarrow} (\Gamma_b^{\alpha^0,\beta^0})_{i,j},$$

$$\text{(4.35)}\qquad \sum_{k=0}^{n-1} \mathbb{E}^{\alpha^0,\beta^0}[(\hat{\zeta}_k^{\alpha_i})^4 \mid \mathcal{G}_{k\Delta_n}] \overset{\mathbb{P}^{\alpha^0,\beta^0}}{\longrightarrow} 0,$$

$$\text{(4.36)}\qquad \sum_{k=0}^{n-1} \mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\beta_i} \mid \mathcal{G}_{k\Delta_n}] \overset{\mathbb{P}^{\alpha^0,\beta^0}}{\longrightarrow} -\frac{1}{2}v_i(\Gamma_S^{\alpha^0,\beta^0})_{i,i} - v_{i+1}(\Gamma_S^{\alpha^0,\beta^0})_{i,i+1} - \cdots - v_{n_\beta}(\Gamma_S^{\alpha^0,\beta^0})_{i,n_\beta},$$

$$\text{(4.37)}\qquad \sum_{k=0}^{n-1} \mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\beta_i}\,\hat{\zeta}_k^{\beta_j} \mid \mathcal{G}_{k\Delta_n}] - \mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\beta_i} \mid \mathcal{G}_{k\Delta_n}]\,\mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\beta_j} \mid \mathcal{G}_{k\Delta_n}] \overset{\mathbb{P}^{\alpha^0,\beta^0}}{\longrightarrow} (\Gamma_S^{\alpha^0,\beta^0})_{i,j},$$

$$\sum_{k=0}^{n-1} \mathbb{E}^{\alpha^0,\beta^0}[(\hat{\zeta}_k^{\alpha_i})^4 \mid \mathcal{G}_{k\Delta_n}] \overset{\mathbb{P}^{\alpha^0,\beta^0}}{\longrightarrow} 0, \tag{4.38}$$

$$\sum_{k=0}^{n-1} \mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\alpha_i}\, \hat{\zeta}_k^{\beta_j} \mid \mathcal{G}_{k\Delta_n}] - \mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\alpha_i} \mid \mathcal{G}_{k\Delta_n}]\, \mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\beta_j} \mid \mathcal{G}_{k\Delta_n}] \overset{\mathbb{P}^{\alpha^0,\beta^0}}{\longrightarrow} 0. \tag{4.39}$$

If we admit for a while these estimates, it is easy to derive Theorem 4.1 by an application of Theorem VII-5-2 from Jacod *et al.* [14] e.g., combined with equality (1.10), (2.11), Lemmas 3.3 and 3.4.
In the following computations, Lemma 3.1 is going to be frequently used without being quoted. Furthermore, the notation ϵ_n refers to any subsequence converging to 0: most of the time, it is equal to some positive power of $\frac{1}{\sqrt{n\Delta_n}}$ or $\sqrt{\Delta_n}$, the power possibly depending of the Hölder exponent γ.

Proof of (4.33). It is clear that

$$\mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\alpha_i} \mid \mathcal{G}_{k\Delta_n}] = \frac{1}{\sqrt{n\Delta_n}} \int_0^1 dl\; \partial_{\alpha_i} b(\alpha_i(l), X_{k\Delta_n}).\left[S^{-2}(\beta^+, X_{k\Delta_n})(m^{\alpha^0,\beta^0}(X_{k\Delta_n}) - m^{\alpha_i(l),\beta^+}(X_{k\Delta_n}))\right].$$

From $m^{\alpha,\beta}(x) = x + \int_0^{\Delta_n} \mathbb{E}_x^{\alpha,\beta}(b(\alpha, X_t^{\alpha,\beta}))\, dt$, it readily follows using equations (2.17) and (2.18) that the difference $m^{\alpha^0,\beta^0}(x) - m^{\alpha_i(l),\beta^+}(x)$ is equal to

$$-\Delta_n \partial_{\alpha_i} b(\alpha^0, x) \frac{l u_i}{\sqrt{n\Delta_n}} - \Delta_n \partial_{\alpha_{i+1}} b(\alpha^0, x) \frac{u_{i+1}}{\sqrt{n\Delta_n}} - \cdots - \Delta_n \partial_{\alpha_{n_\alpha}} b(\alpha^0, x) \frac{u_{n_\alpha}}{\sqrt{n\Delta_n}} + R\left(\epsilon_n \sqrt{\frac{\Delta_n}{n}}, x\right).$$

The completion of proof of (4.33) is now straightforward.

Proof of (4.34). With the previous arguments, one justifies that $\mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\alpha_i} \mid \mathcal{G}_{k\Delta_n}]\, \mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\alpha_j} \mid \mathcal{G}_{k\Delta_n}] = R(n^{-2}, X_{k\Delta_n})$; thus, this term has a negligible contribution. On the other hand, one easily gets

$$\begin{aligned}
\mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\alpha_i}\, \hat{\zeta}_k^{\alpha_j} \mid \mathcal{G}_{k\Delta_n}] &= \frac{1}{n\Delta_n} \sum_{l_1,l_2} \int_0^1 \int_0^1 dl\, dl'\, [S^{-2}(\beta^+, X_{k\Delta_n}) \partial_{\alpha_i} b(\alpha_i(l), X_{k\Delta_n})]_{l_1} \\
&\quad \times [S^{-2}(\beta^+, X_{k\Delta_n}) \partial_{\alpha_j} b(\alpha_j(l'), X_{k\Delta_n})]_{l_2} \times \Big[V_{l_1,l_2}^{\alpha^0,\beta^0}(X_{k\Delta_n}) + \\
&\quad (m^{\alpha^0,\beta^0}(X_{k\Delta_n}) - m^{\alpha_i(l),\beta^+}(X_{k\Delta_n}))_{l_1} (m^{\alpha^0,\beta^0}(X_{k\Delta_n}) - m^{\alpha_j(l'),\beta^+}(X_{k\Delta_n}))_{l_2}\Big] \\
&= \frac{1}{n}\, \partial_{\alpha_i} b(\alpha^0, X_{k\Delta_n}).[S^{-2}(\beta^0, X_{k\Delta_n}) \partial_{\alpha_j} b(\alpha^0, X_{k\Delta_n})] + \frac{\epsilon_n}{n} R(1, x),
\end{aligned}$$

so that convergence (4.34) holds true.

Proof of (4.35). Basic estimates yield $\mathbb{E}^{\alpha^0,\beta^0}[(\hat{\zeta}_k^{\alpha_i})^4 \mid \mathcal{G}_{k\Delta_n}] = R(n^{-2}, x)$ and the result follows.

Proof of (4.36). One has that

$$\begin{aligned}
\mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\beta_i} \mid \mathcal{G}_{k\Delta_n}] = \frac{1}{\sqrt{n\Delta_n}} \int_0^1 dl\; \mathrm{Tr}\Big\{&(\partial_{\beta_i} S\, S^{-3})(\beta_i(l), X_{k\Delta_n}) \\
&\times \Big[\Big((m^{\alpha^0,\beta^0} - m^{\alpha,\beta_i(l)})(m^{\alpha^0,\beta^0} - m^{\alpha,\beta_i(l)})^*\Big)(X_{k\Delta_n}) + \Big(V^{\alpha^0,\beta^0} - V^{\alpha,\beta_i(l)}\Big)(X_{k\Delta_n})\Big]\Big\}.
\end{aligned}$$

Terms involving the difference with $m^{\alpha,\beta}$ are clearly negligible. For the others, use the equality $V_{l_1,l_2}^{\alpha,\beta}(x) = x_{l_1}x_{l_2} + \int_0^{\Delta_n} \mathbb{E}_x^{\alpha,\beta}((S^2)_{l_1,l_2}(\beta, X_t^{\alpha,\beta}) + b_{l_1}(\alpha, X_t^{\alpha,\beta}) X_{l_2,t}^{\alpha,\beta} + b_{l_2}(\alpha, X_t^{\alpha,\beta}) X_{l_1,t}^{\alpha,\beta})\, dt - m_{l_1}^{\alpha,\beta}(x) m_{l_2}^{\alpha,\beta}(x)$, and equations (2.17), (2.18) to obtain that the difference $V^{\alpha^0,\beta^0}(x) - V^{\alpha,\beta_i(l)}(x)$ is equal to

$$-2\Delta_n(\partial_{\beta_i}SS)(\beta^0,x)\frac{lv_i}{\sqrt{n}} - 2\Delta_n(\partial_{\beta_{i+1}}SS)(\beta^0,x)\frac{v_{i+1}}{\sqrt{n}} - \cdots - 2\Delta_n(\partial_{\beta_{n_\beta}}SS)(\beta^0,x)\frac{v_{n_\beta}}{\sqrt{n}} + R\left(\epsilon_n\frac{\Delta_n}{\sqrt{n}},x\right).$$

This completes the proof of (4.36).

Proof of (4.37). We neglect the second product since $\mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\beta_i} \mid \mathcal{G}_{k\Delta_n}]\,\mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\beta_j} \mid \mathcal{G}_{k\Delta_n}] = R(n^{-2},x)$. For the first term, we immediately obtain

$$\begin{aligned}\mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\beta_i}\,\hat{\zeta}_k^{\beta_j} \mid \mathcal{G}_{k\Delta_n}] = \frac{1}{n{\Delta_n}^2}\sum_{l_1,l_2,l_3,l_4}\int_0^1\int_0^1 dl\,dl'\,[\partial_{\beta_i}SS^{-3}(\beta_i(l),X_{k\Delta_n})]_{l_1,l_2}[\partial_{\beta_j}SS^{-3}(\beta_j(l'),X_{k\Delta_n})]_{l_3,l_4}\\ \times\mathbb{E}^{\alpha^0,\beta^0}\Big[\Big((X_{(k+1)\Delta_n} - m^{\alpha,\beta_i(l)}(X_{k\Delta_n}))_{l_1}(X_{(k+1)\Delta_n} - m^{\alpha,\beta_i(l)}(X_{k\Delta_n}))_{l_2} - V^{\alpha,\beta_i(l)}_{l_1,l_2}(X_{k\Delta_n})\Big)\\ \times\Big((X_{(k+1)\Delta_n} - m^{\alpha,\beta_j(l')}(X_{k\Delta_n}))_{l_3}(X_{(k+1)\Delta_n} - m^{\alpha,\beta_j(l')}(X_{k\Delta_n}))_{l_4} - V^{\alpha,\beta_j(l')}_{l_3,l_4}(X_{k\Delta_n})\Big) \mid \mathcal{G}_{k\Delta_n}\Big].\end{aligned}$$

Long but standard computations give that the expectation inside the sum satisfies

$$\mathbb{E}^{\alpha^0,\beta^0}[\cdots \mid \mathcal{G}_{k\Delta_n}] = {\Delta_n}^2\left[(S^2)_{l_1,l_3}(S^2)_{l_2,l_4} + (S^2)_{l_1,l_4}(S^2)_{l_2,l_3}\right](\beta^0,X_{k\Delta_n}) + R({\Delta_n}^2\epsilon_n,X_{k\Delta_n}).$$

The end of the proof of (4.37) now follows easily.

Proof of (4.38). It is clear since $\mathbb{E}^{\alpha^0,\beta^0}[(\hat{\zeta}_k^{\beta_i})^4 \mid \mathcal{G}_{k\Delta_n}] = R(n^{-2},x)$.

Proof of (4.39). Using standard estimates, one has

$$\begin{aligned}\mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\alpha_i}\,\hat{\zeta}_k^{\beta_j} \mid \mathcal{G}_{k\Delta_n}] &= \frac{1}{n{\Delta_n}^{3/2}}\sum_{l_1,l_2,l_3}\int_0^1\int_0^1 dl\,dl'\,[S^{-2}(\beta^+,X_{k\Delta_n})\partial_{\alpha_i}b(\alpha_i(l),X_{k\Delta_n})]_{l_1}\\ &\quad[\partial_{\beta_j}SS^{-3}(\beta_j(l'),X_{k\Delta_n})]_{l_2,l_3}\times\mathbb{E}^{\alpha^0,\beta^0}\Big[(X_{(k+1)\Delta_n} - m^{\alpha_i(l),\beta^+}(X_{k\Delta_n}))_{l_1}\\ &\quad\times\Big((X_{(k+1)\Delta_n} - m^{\alpha,\beta_j(l')}(X_{k\Delta_n}))_{l_2}(X_{(k+1)\Delta_n} - m^{\alpha,\beta_j(l')}(X_{k\Delta_n}))_{l_3} - V^{\alpha,\beta_j(l')}_{l_2,l_3}(X_{k\Delta_n})\Big) \mid \mathcal{G}_{k\Delta_n}\Big]\\ &= R(n^{-1}\sqrt{\Delta_n},X_{k\Delta_n}).\end{aligned}$$

Furthermore, it is clear that $\mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\alpha_i} \mid \mathcal{G}_{k\Delta_n}]\,\mathbb{E}^{\alpha^0,\beta^0}[\hat{\zeta}_k^{\beta_j} \mid \mathcal{G}_{k\Delta_n}] = R(n^{-2},X_{k\Delta_n})$. This completes the proof of (4.39).

□

5 Validity of the LAN property under other assumptions

In this section, we consider a new set of hypotheses, different of **(R)**, **(D)** and **(E)**, and we discuss the validity of the result of previous sections under these assumptions. Our motivation is to extend the class of ergodic models that we may consider for the LAN property, to a class of SDE's with bounded drift coefficient (for which **(D)** can not be fulfilled). Assumptions **(R)** and **(D)** have to be replaced by the following ones.

Assumption (R'): this is the same assumption as **(R)**, except that $b(\alpha,x)| \leq c(1+|x|)$ is replaced by $|b(\alpha,x)| \leq c$.

Assumption (D'): there are constants $K_0 > 0$ and $c_0 > 0$ such that

$$\forall (\alpha, x) \in \Theta_\alpha \times \mathbb{R}^d \quad (|x| \geq K_0 \Longrightarrow b(\alpha, x).x \leq -c_0|x|)\,.$$

An analogous assumption to **(D')** is made by Florens-Zmirou in [4] (see also Has'minskii [11]). Now, we are going to briefly justify than under **(R')**, **(D')** and **(E)**, $X^{\alpha,\beta}$ is ergodic: the main tool is time uniform controls on exponential moments which we now state.

Proposition 5.1. *Let $f_C(x)$ be a smooth function which coincides with $exp(C|x|)$ for $|x| \geq 1$. Under* **(R')**, **(D')** *and* **(E)**, *there is a constant $C_e > 0$ such that*

1. *for any $C \in [0, C_e)$, one has for some constants $\lambda = \lambda(C) > 0$ and $K = K(C)$:*

$$\forall\, t \geq 0 \qquad \mathbb{E}^{\alpha,\beta}_{x_0} f_C(X_t) \leq f_C(x_0)\ \exp(-\lambda\, t) + K. \tag{5.40}$$

2. *$(X^{\alpha,\beta}_t)_{t\geq 0}$ is ergodic and its unique invariant measure $\mu^{\alpha,\beta}$ satisfies for any $C < C_e$:*

$$\int_{\mathbb{R}^d} \exp(C|x|)\ \mu^{\alpha,\beta}(dx) < \infty. \tag{5.41}$$

Proof. We apply the same arguments as for the proof of Proposition 1.1. Using assumptions **(R')** and **(D')**, check that for $|x| \geq 1$, one has $L^{\alpha,\beta} f_C(x) \leq C\ f_C(x)(\frac{b(\alpha,x).x}{|x|} + \frac{K_1}{|x|} + K_2 C)$, hence $L^{\alpha,\beta} f_C(x) \leq -C\ c_0 f_C(x)/2$ for $|x| \geq (4K_1/c_0) \vee K_0 \vee 1$ and $C \leq c_0/(4K_2)$. Thus, if $g(t) = \mathbb{E}^{\alpha,\beta}_{x_0} f_C(X_t)$, one has proved that $g'(t) \leq -\lambda g(t) + K$ (with $\lambda = C\ c_0/2$) and (5.40) easily follows.
Since one gets time uniform control on moments, the existence of an unique invariant measure is a consequence (see Has'minskii [11]) of the strict positivity of the transition density, this fact being clear under **(R')** and **(E')**. The proof of (5.41) is obtained as for (1.5).

□

We now state that the LAN property is also valid for this class of models.

Theorem 5.1. *Under* **(R')**, **(D')** *and* **(E)**, *the conclusion of Theorem 4.1 remains true.*

Proof. Apply exactly the same arguments as for Theorem 4.1. The main difference comes from the estimates of Proposition 1.2, which have to be adapted to the new hypotheses. Actually, one can prove, without difficulty, that estimates (1.6), (1.7), (1.8) and (1.9) are valid without the factor $\exp(\pm ct|x|^2)$: clearly, this modification does not change the result, since μ^{α^0,β^0} has polynomial moments of any order.

□

The reader may have understood than weaker forms of assumption **(R')** and **(D')** are available: the crucial fact is to ensure that μ^{α^0,β^0} has enough moments to control the growth of the derivatives of b and S. For instance, if one replaces $b(\alpha, x).x \leq -c_0|x|$ by $b(\alpha, x).x \leq -c_0|x|^{\gamma'}$ with $\gamma' \in (0, 1)$ (this ensures polynomial moments for μ^{α^0,β^0} up to some order q_0), one can explicit the maximal polynomial growth order which is allowed for the derivatives of b and S.

A Estimates on the transition density function

This Appendix is devoted to the proof of Proposition 1.2, which assumptions we assume.

A.1 Proof of (1.6) and (1.7).

Owing the Markov property, note that it is sufficient to prove these estimates only for $t \le T_0$, where $T_0 > 0$ is an arbitrary small positive constant depending only on b and S.

Our techniques are based on a Girsanov transformation. We introduce some notation and recall some well known results.
For sake of simplicity, $p^{\alpha,\beta}(t,x,y)$ (resp. $\mathbb{E}^{\alpha,\beta}$) is simply denoted by $p(t,x,y)$ (resp. $\mathbb{E}$). We also omit the parameters α and β in the coefficients b and S. $\mathbb{E}^0$ and $p^0(t,x,y)$ refers to the law of the SDE's (1.3) where the drift coefficient is removed, i.e. $X_t = x + \int_0^t S(X_s)dB_s$ (B being a Brownian motion under $\mathbb{E}^0$). We set $Z_t = \exp(\int_0^t S^{-1}(X_s)b(X_s)dB_s - \frac{1}{2}\int_0^t |S^{-1}(X_s)b(X_s)|^2 ds)$. Since $S^{-1}b$ has a linear growth, $(Z_t)_{t\ge 0}$ is a martingale (see Benes' criterion, [15] p.200) and this allows a Girsanov transformation.
Furthermore, it is well known (see Aronson [1], Friedman [5]) that $p^0(t,x,y)$ is smooth and satisfies

$$\text{(A.42)} \qquad \frac{1}{K\, t^{d/2}} \exp\left(-c\frac{|x-y|^2}{t}\right) \le p^0(t,x,y) \le \frac{K}{t^{d/2}} \exp\left(-\frac{|x-y|^2}{c\,t}\right),$$

$$\text{(A.43)} \qquad \left|\nabla_x p^0(t,x,y)\right| \le \frac{K}{t^{(d+1)/2}} \exp\left(-\frac{|x-y|^2}{c\,t}\right)$$

for some uniform constants. We are going to derive (1.6) and (1.7), from (A.42) and (A.43) using the announced Girsanov transformation. The following Lemma gives the other necessary estimates.

Lemma A.1. *For any $\mu_1 > 1$, any $q \ge 0$, there are some constants $T_0 > 0$, $c > 0$, $K > 0$ such that for $t \le T_0$, one has*

$$\mathbb{E}_x^0\left(Z_t^{\mu_1}(1+|X_t|)^q\right) + \mathbb{E}_x^0\left(Z_t^{-\mu_1}(1+|X_t|)^q\right) \le K\exp(c\,t\,|x|^2)\,(1+|x|)^q.$$

Proof. Since for any $r \ge 0$, $\mathbb{E}_x^0(1+|X_t|)^r \le K(1+|x|)^r$, it suffices to prove Lemma A.1 when $q = 0$. Fix $\lambda \ge 0$. One has that $\lambda\int_0^t |S^{-1}(X_s)b(X_s)|^2 ds \le \lambda K t(|x|^2 + \sup_{s\in[0,t]}|X_s - x|^2)$, and besides, one easily checks $\mathbb{E}_x^0 \exp(\lambda K t \sup_{s\in[0,t]}|X_s - x|^2) \le K_1$ for t small enough (use e.g. a time-changed Brownian motion coordinate-wise); thus, for $t \le T_0(\lambda)$, one obtains that

$$\text{(A.44)} \qquad \mathbb{E}_x^0 \exp\left(\lambda\int_0^t |S^{-1}(X_s)b(X_s)|^2 ds\right) \le K\exp(ct|x|^2).$$

Write $Z_t^{\mu_1} = \exp\left(\mu_1\int_0^t S^{-1}(X_s)b(X_s)dB_s - \mu_1^2\int_0^t |S^{-1}(X_s)b(X_s)|^2 ds\right)\exp\left(\mu_1^2\int_0^t |S^{-1}(X_s)b(X_s)|^2 ds\right)$, take the expectation and apply the Cauchy-Schwarz inequality: the first term is equal to 1 and the second one is estimated by (A.44) for t small enough. This completes the proof of the estimate for $\mathbb{E}_x^0(Z_t^{\mu_1})$. Same arguments apply for $\mathbb{E}_x^0(Z_t^{-\mu_1})$.

□

A.1.1 Proof of (1.6).

Owing the Girsanov transformation, one has

$$\text{(A.45)} \qquad p(t,x,y) = p^0(t,x,y)\mathbb{E}_x^0\left(Z_t \mid X_t = y\right).$$

To deal with the above conditioning, we invoke the law of the diffusion bridge from $X_0 = x$ to $X_t = y$ (see Lyons et al. [20] e.g.), i.e. an other Girsanov transformation, which transforms the Brownian motion $B_.$ in $B_. + \int_0^{\cdot} S(X_u) \frac{\nabla_x p^0(t-u, X_u, y)}{p^0(t,x,y)} du$. Hence, since $Z_t = 1 + \int_0^t Z_s\, S^{-1}(X_s) b(X_s)\, dB_s$, one gets

$$\mathbb{E}_x^0\left(Z_t \mid X_t = y\right) = 1 + \frac{1}{p^0(t,x,y)} \int_0^t \mathbb{E}_x^0\left[Z_s\, b(X_s).\nabla_x p^0(t-s, X_s, y)\right] ds.$$

Applying Hölder's inequality (with μ_1 and μ_2 conjugate), Lemma A.1, upper bounds (A.42) and (A.43), one obtains (for t small enough) that $\left|\mathbb{E}_x^0\left[Z_s\, b(X_s).\nabla_x p^0(t-s, X_s, y)\right]\right|$ is bounded by

$$\begin{aligned} &K\ \exp(c\, t\, |x|^2)\ (1+|x|) \left[\int_{\mathbb{R}^d} \frac{dz}{s^{d/2}(t-s)^{(d+1)\mu_2/2}}\ \exp\left(-\frac{|x-z|^2}{c\, s} - \mu_2 \frac{|z-y|^2}{c\,(t-s)}\right)\right]^{1/\mu_2} \\ &\le K\ \exp(c\, t\, |x|^2)\ (1+|x|) \frac{1}{t^{d/(2\mu_2)}(t-s)^{(d+1)/2 - d/(2\mu_2)}}\ \exp\left(-\frac{|x-y|^2}{c'\, t}\right). \end{aligned}$$

We now choose μ_2 closed to 1 to ensure that $(d+1)/2 - d/(2\mu_2) < 1$: it readily follows that

$$\mathbb{E}_x^0\left(Z_t \mid X_t = y\right) \le 1 + \frac{K}{p^0(t,x,y)} \frac{\exp(c\, t\, |x|^2)\ (1+|x|)}{t^{(d-1)/2}} \exp\left(-\frac{|x-y|^2}{c'\, t}\right).$$

Using $\exp(c\ t\ |x|^2)\ |x|/t^{(d-1)/2} \le K \exp(c'\ t\ |x|^2)/t^{d/2}$ combined with the inequality above, equality (A.45) and upper bound (A.42), one completes the proof of (1.6) for t small enough.

A.1.2 Proof of (1.7).

From equality (A.45), Jensen's inequality yields

$$\text{(A.46)} \qquad \frac{1}{p(t,x,y)} \le \frac{1}{p^0(t,x,y)} \mathbb{E}_x^0\left(Z_t^{-1} \mid X_t = y\right)$$

with $Z_t^{-1} = 1 - \int_0^t Z_s^{-1}\ S^{-1}(X_s) b(X_s)\ dB_s + \int_0^t Z_s^{-1}\ |S^{-1}(X_s) b(X_s)|^2\ ds$. Introducing the diffusion bridge as before, we can prove that

$$\text{(A.47)} \qquad \left|\mathbb{E}_x^0\left(\int_0^t Z_s^{-1}\ S^{-1}(X_s) b(X_s)\ dB_s \mid X_t = y\right)\right| \le \frac{K}{p^0(t,x,y)} \frac{\exp(c\, t\, |x|^2)}{t^{d/2}} \exp\left(-\frac{|x-y|^2}{c\, t}\right).$$

Besides, using Hölder's inequality (with μ_1 and μ_2 conjugate), one gets for $s < t$

$$\begin{aligned} \mathbb{E}_x^0\left(Z_s^{-1}\ |S^{-1}(X_s) b(X_s)|^2 \mid X_t = y\right) &= \frac{1}{p^0(t,x,y)} \mathbb{E}_x^0\left(Z_s^{-1}\ |S^{-1}(X_s) b(X_s)|^2\ p^0(t-s, X_s, y)\right) \\ &\le \frac{K\ \exp(c\, t\, |x|^2)\ (1+|x|^2)}{p^0(t,x,y)} \frac{\exp\left(-\frac{|x-y|^2}{c\, t}\right)}{t^{d/(2\mu_2)}(t-s)^{d/2 - d/(2\mu_2)}}, \end{aligned}$$

so that choosing μ_2 closed to 1, one obtains that

$$\text{(A.48)} \quad \mathbb{E}_x^0\left(\int_0^t ds\ Z_s^{-1}\ |S^{-1}(X_s) b(X_s)|^2 \mid X_t = y\right) \le \frac{1}{p^0(t,x,y)} \frac{K\ \exp(c\, t\, |x|^2)}{t^{d/2}} \exp\left(-\frac{|x-y|^2}{c'\, t}\right).$$

Combining (A.46), (A.47), (A.48) and (A.42), one completes the proof of the lower bound of $p(t,x,y)$ for t small enough.

$\square$

A.2 Proof of (1.8) and (1.9).

The arguments being similar for both estimates, we only detail the proof of (1.8). Using Jensen's inequality and Proposition 2.2, one obtains:

$$\mathbb{E}_x^{\overline{\alpha},\overline{\beta}}\left|\frac{\partial_{\alpha_i}p^{\alpha,\beta}}{p^{\alpha,\beta}}(t,x,X_t)\right|^\nu \le \int_{\mathbb{R}^d} dy\, p^{\overline{\alpha},\overline{\beta}}(t,x,y)\frac{1}{t^\nu}\,\mathbb{E}_x^{\alpha,\beta}\left[\left|\sum_{l_1=1}^d \delta(\partial_{\alpha_i}X_{l_1,t}^{\alpha,\beta}\,U_{l_1})\right|^\nu \mid X_t^{\alpha,\beta}=y\right]$$

$$\le \frac{1}{t^\nu}\mathbb{E}_x^{\alpha,\beta}\left[\frac{p^{\overline{\alpha},\overline{\beta}}(t,x,X_t^{\alpha,\beta})}{p^{\alpha,\beta}(t,x,X_t^{\alpha,\beta})}\left|\sum_{l_1=1}^d \delta(\partial_{\alpha_i}X_{l_1,t}^{\alpha,\beta}\,U_{l_1})\right|^\nu\right].$$

Apply Hölder's inequality (with μ_1 and μ_2 conjugate). On one hand, check that $\mathbb{E}_x^{\alpha,\beta}\left[\frac{p^{\overline{\alpha},\overline{\beta}}(t,x,X_t^{\alpha,\beta})}{p^{\alpha,\beta}(t,x,X_t^{\alpha,\beta})}\right]^{\mu_1}$ is bounded by $\exp(c\,t|x|^2)$ up to choosing μ_1 closed to 1 (see the arguments used to prove (A.48)). On the other hand, $\mathbb{E}_x^{\alpha,\beta}[|\sum_{l_1=1}^d \delta(\partial_{\alpha_i}X_{l_1,t}^{\alpha,\beta}\,U_{l_1})|^{\nu\mu_2}]$ is estimated by $t^{3\nu\mu_2/2}(1+|x|)^q$, applying the arguments as in the proof of Lemma 3.3. We are finished.

□

References

[1] D.G. Aronson : Bounds for the fundamental solution of a parabolic equation - Bulletin of the American Mathematical society - Vol.73 - 1967 - pp. 890-903.

[2] D. Dacunha-Castelle and D. Florens-Zmirou: Estimation of the coefficient of a diffusion from discrete observations - Stochastics - Vol.19 - 1986 - pp. 263-284.

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

[4] D. Florens-Zmirou: Approximate discrete time schemes for statistics of diffusion processes - Statistics - Vol.20 - 1989 - pp. 547-557.

[5] A. Friedman : Stochastic differential equations and applications, volume 1, Academic Press, New York - 1975.

[6] V. Genon-Catalot: Maximum contrast estimation for diffusion processes from discrete observations - Statistics - Vol.21 - 1990 - pp. 99-116.

[7] 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.

[8] V. Genon-Catalot and J. Jacod: Estimation of the diffusion coefficient for diffusion processes: random sampling - Scandinavian Journal of Statistics - Vol.21 - 1994 - pp. 193–221.

[9] E. Gobet: LAMN property for elliptic diffusion: a Malliavin calculus approach - Submitted - 2000.

[10] J. Hajek: Local asymptotic minimax and admissibility in estimation - Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley and Los Angeles, University of California Press - 1971 - pp. 175-194.

[11] R.Z. Has'minskii: Stochastic stability of differential equations - Sijthoff & Noordhoff - 1980.

[12] R. Höpfner, J. Jacod and L. Ladelli: Local asymptotic normality and mixed normality for Markov statistical models - Probab. Theory Rel. Fields - Vol.86 - 1990 - pp. 105-129.

[13] J. Jacod: Private communication.

[14] J. Jacod and A.N. Shiryaev: Limit theorems for stochastic processes, Springer Verlag, Berlin - 1987.

[15] I. Karatzas and S.E. Shreve: Brownian motion and stochastic calculus, Springer Verlag - 1988.

[16] M. Kessler: Estimation of an ergodic diffusion from discrete observations - Scandinavian Journal of Statistics - Vol.24 - 1997 - pp. 211-229.

[17] 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.

[18] L. Le Cam: Limits of experiments - Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley and Los Angeles, University of California Press - 1971 - pp. 245-261.

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

[20] T.J. Lyons and W.A. Zheng: On conditional diffusion processes - Proceedings of the Royal Society of Edinburgh - Vol.115A - 1990 - pp. 243-255.

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

[22] B.L.S. Prakasa Rao: Asymptotic theory for non-linear least square estimator for diffusion processes - Math. Operationsforsch. Statist. Ser. Stat - Vol.14 - 1983 - pp. 195-209.

[23] N. Yoshida: Estimation for diffusion processes from discrete observations - Journal of Multivariate Analysis - Vol. 41 - 1992 - pp. 220-242.