Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2008, Vol. 44, No. 1, 104–128
DOI: 10.1214/07-AIHP111

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

LAMN property for hidden processes: The case of integrated diffusions

Arnaud Gloter[a] and Emmanuel Gobet[b]

[a] *Université de Marne–la–Vallée, Laboratoire d'Analyse et de Mathématiques Appliquées UMR 8050, 5 Boulevard Descartes, 77454 Marne–la–Vallée Cedex 2, France. E-mail: arnaud.gloter@univ-mlv.fr.*
[b] *ENSIMAG, INP Grenoble, Laboratoire de Modélisation et Calcul UMR 5523, B.P. 53, 38041 Grenoble Cedex 9, France. E-mail: emmanuel.gobet@imag.fr.*

Received 10 May 2006; accepted 20 November 2006

Abstract. In this paper we prove the Local Asymptotic Mixed Normality (LAMN) property for the statistical model given by the observation of local means of a diffusion process X. Our data are given by $\int_0^1 X_{(s+i)/n}\,\mathrm{d}\mu(s)$ for $i=0,\ldots,n-1$ and the unknown parameter appears in the diffusion coefficient of the process X only. Although the data are neither Markovian nor Gaussian we can write down, with help of Malliavin calculus, an explicit expression for the log-likelihood of the model, and then study the asymptotic expansion. We actually find that the asymptotic information of this model is the same one as for a usual discrete sampling of X.

Résumé. Dans ce papier nous démontrons la propriété LAMN pour le modèle statistique constitué par l'observation des moyennes locales d'une diffusion X. Nos données sont définies comme $\int_0^1 X_{(s+i)/n}\,\mathrm{d}\mu(s)$ avec $i=0,\ldots,n-1$ et le paramètre inconnu apparaît seulement dans le coefficient de diffusion du processus X. Bien que cette observation ne soit ni Gaussienne ni Markovienne nous pouvons, par le calcul de Malliavin, obtenir une expression pour la log-vraisemblance du modèle. Nous sommes alors capables de calculer l'information asymptotique et montrons qu'elle est la même que pour l'observation ponctuelle de la diffusion.

MSC: 60F99; 60H07; 62F12; 62M09

Keywords: Diffusion processes; Parametric estimation; LAMN property; Malliavin calculus; Non-Markovian data

1. Statement of the problem and main results

1.1. Introduction

Model
Let us consider the family of strong solutions X^θ to the following scalar equation

$$\mathrm{d}X_t^\theta = a\big(X_t^\theta,\theta\big)\,\mathrm{d}B_t + b\big(X_t^\theta\big)\,\mathrm{d}t, \tag{1}$$

$$X_0^\theta = \xi_0, \tag{2}$$

where B is a one-dimensional Brownian motion. We suppose that θ lies in some compact interval Θ of $\mathbb{R}$ and that ξ_0 is a real constant, which does not depend on θ, and thus is known to the statistician.

Observations
We consider μ some probability measure on [0, 1] and assume that our observation of the process is given by the local means of X associated with this measure, with sampling of size $1/n$:

$$\overline{X}_j = \overline{X}_{j,n} := \int_0^1 X_{(s+j)/n} \, \mathrm{d}\mu(s), \quad \text{for } j = 0, \ldots, n-1. \tag{observations}$$

In the sequel this case is referred to as the integrated diffusion case. This is an indirect observation of the process X and the observation is no more the realization of a Markov chain. Thus, this framework is deeply related to the inference of hidden processes. We assume that μ does not depend on θ and is known by the statistician. When μ is equal to the Lebesgue measure, the observation is the discrete sampling of $I_t = \int_0^t X_s \, \mathrm{d}s$. This is presumably the simplest case of the observation of only one component of a bidimensional diffusion process $(X_t, I_t)_{0 \leqslant t \leqslant 1}$, which is known in the literature as the standard integrated diffusion case. Clearly, the usual case of pointwise observation of X is obtained if μ is some Dirac measure. However, we will exclude that the measure has mass only on the end points of the interval, and hence make the assumption:

$$\mu\big((0, 1)\big) > 0. \tag{3}$$

This paper is concerned with the Local Asymptotic Mixed Normality property of this statistical model.

Motivation
Taking as the observation the integrated process is actually quite natural. For instance, it arises when the realization of the process has been observed after passage through an electronic filter. Also, in random mechanics (see [19]), X models the velocity of the system and in general, we observe its position, i.e. the integral of X. The modeling of ice-core data can be made through an integrated diffusion process (see [2]). Integrated processes also play an important role in finance, when modeling the stochastic volatility (see, for instance, [1] and references therein).

Literature background
Despite of these numerous motivations, few statistical studies deal with this situation. Gloter [7,8] provides an estimator in the multiplicative case $a(x, \theta) = \theta a(x)$ and proves its consistency and asymptotic normality. The case of a low frequency observation (local means over interval of length 1) is studied by Ditlevsen and Sørensen [3], using prediction-based estimating functions. On the other hand, for a direct observation of the diffusion X, there are many contributions in the literature: see [5,24] and references therein. None of these works deal with the problem of optimal estimation in the integrated diffusion model.

Here, we directly address the problem of the LAMN property, whose fundamental consequence is to provide information on the minimal dispersion for an estimator of the parameter θ (see [14,16,17,21,24]). Such properties, for the observation of a discrete sampling of the diffusion, have been established in the one-dimensional setting by Dohnal [4], and then extended by Gobet [10,11] to the multidimensional setting, both in the high frequency and ergodic framework. For this, Malliavin calculus techniques were used and paved the way to possibly handle more general situations than Markovian observations. This is exactly this way we follow in this work, to tackle the case of integrated diffusion.

Outlook
We guess that this model captures the main difficulty of most hidden models: the lack of Markov property for the observation. Hence, the method developed below (augmented observation, Malliavin calculus representation, Gaussian approximation) may be useful to treat more general situations. Among the natural situations coming from applications, one can think of the measurement of a stochastic phenomenon blurred by some noise, or stochastic volatility models widely used in finance [6]. This can be formalized as follows: the system $\mathcal{X}^\theta$ is governed by the $(d+d')$-dimensional stochastic differential equation

$$\mathcal{X}_t^\theta = \mathcal{X}_0^\theta + \int_0^t \mathcal{A}\big(\mathcal{X}_s^\theta, \theta\big) \, \mathrm{d}B_s + \int_0^t \mathcal{B}\big(\mathcal{X}_s^\theta\big) \, \mathrm{d}s,$$

where only a discrete sampling of the first d components is observable. This is left to further research.

1.2. Main results

Before going into the details of our results, we present a very simple example which gives some insight on the type of results that one can expect.

Example 1 (Multiplicative Brownian case). *Assume that the model is*

$$X_t^\theta = \theta B_t$$

(*corresponding to* $b \equiv 0$ *and* $a(\cdot, \theta) = \theta, \xi_0 = 0$).

1. *Consider a first situation where one observes the diffusion at discrete times. Hence, the observation is* $(X_{i/n})_{0 \leqslant i \leqslant n}$, *or equivalently* $(Z_i = \theta(B_{i/n} - B_{(i-1)/n}) = \theta G_i)_{1 \leqslant i \leqslant n}$, *where* G_i *are independent centered Gaussian variables, with a known variance. Thus, the estimation of* θ^2 *is achieved at rate* $\sqrt{n}$, *with a minimal variance equal to* $2\theta^4$.
2. *Now consider a second situation where one observes only the integrated diffusion at discrete times. Hence, the observation is* $(\overline{X}_i = \theta \int_0^1 B_{(s+i)/n} \mu(\mathrm{d}s) = \theta G'_i)_{1 \leqslant i \leqslant n}$, *where* $(G'_i)_i$ *is a centered Gaussian vector, with a known covariance matrix. In addition, this matrix is invertible and thus,* θ^2 *can be estimated with the same rate and asymptotic variance as before.*

This means that observing the process at discrete times or its integrated version lead to the same accuracy in the parameter estimation. The results of this paper state that this is true, even for the more general models (1) *and* (2), *which is far from intuitive.*

Before stating our main results, we define the working assumptions of this paper. The coefficients $a : \mathbb{R} \times \Theta \to \mathbb{R}$ and $b : \mathbb{R} \to \mathbb{R}$, are assumed to satisfy the following set of conditions (as usual, derivatives w.r.t. θ are denoted with a dot: for instance, $\partial_\theta a = \dot{a}$).

Assumption (R).

(1) *The function* $a : \mathbb{R} \times \Theta \to \mathbb{R}$ *is* $\mathcal{C}^{1+\gamma}$ *for some* $\gamma \in (0, 1)$ (*it admits a derivative which is* γ*-Hölder*). *The one-dimensional functions* $x \mapsto a(x, \theta)$, $x \mapsto \dot{a}(x, \theta)$, $x \mapsto b(x)$ *are assumed to be* $\mathcal{C}^3(\mathbb{R})$.
(2) *The functions* a, $\dot{a}$ *and* b *and all their derivatives with respect to* x *are bounded uniformly in* θ.
(3) *We have the nondegeneracy condition, for some* $\underline{a}$: $a(x, \theta) > \underline{a} > 0$ *for all* x, θ.

Actually, the uniform controls in (R) can be weakened to local ones, using extra techniques of space localization (see Lemma 4.1 in [10]). We omit further details. An extension of our results to a multidimensional parameter θ and to time dependent coefficients is straightforward, in the same way as it is done in [10] and [11].

We denote by $\mathbb{P}^\theta$ the law on $\mathcal{C}([0, 1])$ of the process X^θ, and then simply denote X the canonical process on $\mathcal{C}([0, 1])$. We let $\mathbf{p}^{\mathbf{n},\theta}$ denote the law on $\mathbb{R}^n$ of the observation $\mathcal{O}^n := (\overline{X}_j)_{j=0,\ldots,n-1}$, when the true value of the parameter is θ. And for θ_0, θ_1 two values of the parameter we introduce the likelihood ratio,

$$Z^n_{\theta_0,\theta_1} = \frac{\mathrm{d}\mathbf{p}^{\mathbf{n},\theta_1}}{\mathrm{d}\mathbf{p}^{\mathbf{n},\theta_0}}(\mathcal{O}^n). \tag{4}$$

The main result is that this statistical model satisfies the so called LAMN property. For this denote the sequence $u_n := n^{-1/2}$, and let $\theta_0 \in \Theta$ and $h \in \mathbb{R}$ such that $\theta_0 + u_n h \in \Theta$, $\forall n$. Then, by the following theorem, the model has the LAMN property for the likelihood at point θ_0, with rate u_n and conditional information:

$$\mathcal{I}_{\theta_0} = 2 \int_0^1 \left(\frac{\dot{a}}{a}\right)^2 (X_s, \theta_0) \, \mathrm{d}s.$$

Theorem 1. *Assume* (R), *then we have the expansion*,

$$\log Z^n_{\theta_0,\theta_0+u_n h} = h N_n - \frac{1}{2} h^2 I_n + R_n,$$

where $I_n \xrightarrow[n\to\infty]{\mathbb{P}^{\theta_0}} \mathcal{I}_{\theta_0}$, $R_n \xrightarrow[n\to\infty]{\mathbb{P}^{\theta_0}} 0$ *and there exists an extra random variable* $N \sim \mathcal{N}(0,1)$ *independent of the process* X *such that,* N_n *converges in law under* $\mathbb{P}^{\theta_0}$ *to* $N\sqrt{\mathcal{I}_{\theta_0}}$.

Moreover this convergence is stable: *for any random variable* F *measurable with respect to the* X, *we have the convergence in law* $(F, N_n) \xrightarrow[n\to\infty]{law} (F, N\sqrt{\mathcal{I}_{\theta_0}})$. *In particular, it implies the joint convergence under* $\mathbb{P}^{\theta_0}$:

$$(I_n, N_n) \xrightarrow[n\to\infty]{law} \big(\mathcal{I}_{\theta_0}, N\sqrt{\mathcal{I}_{\theta_0}}\big).$$

Remark 1. *Let us stress that the rate* $u_n = n^{-1/2}$ *and the information* $\mathcal{I}_{\theta_0}$ *are the same one as for the pointwise observation* (*see* [5]). *This corroborates the intuition from Example* 1.

We will not be able to prove directly this result, instead we shall consider first the easier problem where one can observe additionally the exact value of the diffusion at some instants. This device was proved to be useful in [9] for the study of a Gaussian diffusion process observed with noise that leads to non-Markovian observations, too.

Let $k = k_n$ be an integer in $\{1, \dots, n\}$ and define $L = L_n := \lfloor n/k \rfloor$, then we consider the set of random variables:

$$\mathcal{O}^{n,\mathrm{aug}} = \mathcal{O}^n \cup \{X_{kl/n}, l = 1, \dots, L\} \cup \{X_1\}.$$

Since this set of variables contains more data than the initial set, we call it the *augmented observation* set. Clearly, we can split this set into blocks, $\mathbb{B}_0, \dots, \mathbb{B}_L$, where for $l = 0, \dots, L-1$,

$$\mathbb{B}_l = \{\overline{X}_{kl}, \dots, \overline{X}_{kl+k-1}, X_{k(l+1)/n}\}$$

and $\mathbb{B}_L = \{\overline{X}_{kL}, \dots, \overline{X}_{n-1}, X_1\}$. Note that if $kL = n$ we consider that the last block is empty, and (immediate) modifications should take care of this in the sequel, however, to have shorter notations we will not explicitly write these modifications.

The advantage of this set of augmented observation is that using the Markov property of X, the law the block $\mathbb{B}_l$ conditional to the previous blocks $(\mathbb{B}_{l'})_{l'<l}$ only depends on the last variable, $X_{kl/n}$, of the block $\mathbb{B}_{l-1}$.

Denote by $\mathbf{p}^{\mathbf{n},\mathrm{aug},\theta}$ the law of $\mathcal{O}^{n,\mathrm{aug}}$ on $\mathbb{R}^{n+L+1}$ and introduce the likelihood ratio for the augmented observation:

$$Z^{n,\mathrm{aug}}_{\theta_0,\theta_1} = \frac{\mathrm{d}\mathbf{p}^{\mathbf{n},\mathrm{aug},\theta_1}}{\mathrm{d}\mathbf{p}^{\mathbf{n},\mathrm{aug},\theta_0}}\big(\mathcal{O}^{n,\mathrm{aug}}\big). \tag{5}$$

Theorem 2. *There exists a sequence* $k_n \to \infty$, *such that the* augmented model *satisfies the LAMN property*:

$$\log Z^{n,\mathrm{aug}}_{\theta_0,\theta_0+u_n h} = hN_n^{\mathrm{aug}} - \frac{1}{2}h^2 I_n^{\mathrm{aug}} + R_n^{\mathrm{aug}},$$

where $I_n^{\mathrm{aug}} \xrightarrow[n\to\infty]{\mathbb{P}^{\theta_0}} \mathcal{I}_{\theta_0}$, $R_n^{\mathrm{aug}} \xrightarrow[n\to\infty]{\mathbb{P}^{\theta_0}} 0$ *and there exists an extra random variable* $N \sim \mathcal{N}(0,1)$ *independent of the process* X *such that,* N_n^{aug} *converges in law under* $\mathbb{P}^{\theta_0}$ *to* $N\sqrt{\mathcal{I}_{\theta_0}}$. *Moreover, this convergence is stable.*

From Theorem 2 and from the consequences of the LAMN property, an *asymptotically optimal* estimator θ_n in the augmented model should satisfy that $\sqrt{n}(\theta_n - \theta_0)$ is asymptotically distributed under $\mathbb{P}_{\theta_0}$ as $\frac{1}{\sqrt{\mathcal{I}_{\theta_0}}}N$. However, any estimator in the initial model of observation $\mathcal{O}^n$ can be seen as an estimator in the augmented model, hence the Theorem 2 is sufficient by itself to imply a lower bound for estimation in the initial model.

Remark 2. *The fact that* $k_n \to \infty$ *means that the data added in the observation are sparse compared to the initial data. Actually, the Theorem* 2 *holds for any sequence* k_n *whose growth to* ∞ *is slow enough.*

If we assume now that $k_n = k \in \mathbb{N}$ remains fixed as $n \to \infty$, the number of data $(X_{kl/n})_{l=0,\dots,L}$ added to the model is not negligible compared to the number of initial data. Hence, the statistical properties of the augmented model shall depend on k and thus differ from the statistical properties of the initial model given in Theorem 1. Actually we have the following LAMN property for the augmented model in that case.

Theorem 3. *If a sequence $k_n = k$ is fixed, then the* augmented model *satisfies the LAMN property with rate $u_n = n^{-1/2}$ and conditional information equal to*:

$$\mathcal{I}_{k,\theta_0} = 2\left(\frac{k+1}{k}\right)\int_0^1 \left(\frac{\dot{a}}{a}\right)^2 (X_s,\theta_0)\,\mathrm{d}s.$$

As expected, the conditional information is greater by a factor $(k+1)/k$, due to the nonnegligibility of the added observations. Actually this factor should be read as $1+\frac{1}{k}$, meaning that an addition of $\frac{1}{k}\%$ of data increases the information in the same way. Local means and values at discrete points are not redundant (as expected from the multiplicative Brownian case, see Example 1) and moreover, they bring an equal information. Considering $k=1$ is interesting, since we observe then on each block $[i/n,(i+1)/n]$ both the exact value $X_{i/n}$ and a mean $\overline{X}_i$. It appears that the asymptotic information is then twice the information given by the observation of only the exact values (or only the means).

1.3. Outline of the paper

In Section 2 we study the score function given by the observation of only one block of data ($\mathbb{B}_0$ for instance). We first focus on the existence of a density for a block of data; and in the case of a block of size 2, $(n^{1/2}\int_0^1 (X_{s/n} - X_0)\,\mathrm{d}\mu(s), n^{1/2}(X_{1/n}-X_0))$ we give original lower and upper bounds of Gaussian type for the density. It is useful for our proof of the LAMN property, but it is also interesting for itself.

In Section 2.2 we present an exact expression for the score function of a block of data $\mathbb{B}_0$ (see Theorem 5). This result is the key point in the proof of the LAMN property, it extends a former result of Gobet [10,11] which gave the score function for the observation of $X^\theta_{1/n}$. In Section 2.3 we study an explicit approximation for the score function when the sampling interval $1/n$ tends to zero and the length of the block k/n remains moderate so that one can consider the coefficients of the diffusion X almost constant on the interval $[0,k/n]$. The key point is the Gaussian approximation for the diffusion given in Section 2.3.1.

In Section 3 we deduce from the previous results a proof of Theorems 2 and 3 and Section 4 deals on how to deduce Theorem 1 from Theorem 2.

Finally, the Appendix contains the proof of some results of Section 2.1 together with some useful lemmas.

Notations
In our proofs, we will keep the same notation for constants which may change from one line to another. In particular, the constants $c, c(k), c(p), c(p,k)$ will stand for all finite, nonnegative and nondecreasing deterministic functions of an index p (arising from $\mathbf{L}^p$-norm) and of the block size k. These constants are independent of n,θ and depend on the process X^θ, only through the bounds on the coefficients a, b and their derivatives.

2. Score function for a block of data

In this section we shall study the law of the blocks of data $\mathbb{B}_l$; recalling the Markov property of the process X it is sufficient to focus on $\mathbb{B}_0 = \{\overline{X}_0,\dots,\overline{X}_{k-1}, X_{k/n}\}$ assuming that the diffusion X now starts from some value x_0. In this section it is convenient to transform the short time asymptotic $k/n \to 0$ into an almost stationarity property of the coefficients. To this end, we introduce the rescaled process $\mathcal{X}^{n,\theta}_t = n^{1/2}(X^\theta_{t/n} - x_0)$ (where X^θ solves (1) with $X^\theta_0 = x_0$). It solves the equation

$$\mathrm{d}\mathcal{X}^{n,\theta}_t = a_n(\mathcal{X}^{n,\theta}_t,\theta)\,\mathrm{d}W_t + b_n(\mathcal{X}^{n,\theta}_t)\,\mathrm{d}t, \quad \mathcal{X}^{n,\theta}_0 = 0, \tag{6}$$

where W is a standard Brownian motion (arising from the rescaling of B), and

$$a_n(x,\theta) = a(x_0 + n^{-1/2}x,\theta), \qquad b_n(x) = n^{-1/2}b(x_0+n^{-1/2}x). \tag{7}$$

Since for the score we are only concerned with the law of $\mathcal{X}^{n,\theta}$, we can assume that W is independent of the rescaling coefficient n.

2.1. *The density of an integrated diffusion*

In this section, we will present preliminary results on the density of the law of the mean of a diffusion process. However, the proofs are postponed to Appendix A.1. To our knowledge, the lower and upper bounds for this density are new results.

2.1.1. *Existence of the density*

Our first result actually deals with the two-dimensional variable given by solely one local mean and the exact value:

$$\begin{aligned}\left(U^{n,\theta}, V^{n,\theta}\right) &:= \left(\int_0^1 \mathcal{X}_s^{n,\theta}\,\mathrm{d}\mu(s), \mathcal{X}_1^{n,\theta}\right)\\ &\overset{\text{law}}{=} \left(n^{1/2}\int_0^1 \left(X_{s/n}^{\theta} - x_0\right)\mathrm{d}\mu(s), n^{1/2}\left(X_{1/n}^{\theta} - x_0\right)\right).\end{aligned} \tag{8}$$

Notice that by the Markov property, the preliminary study of this bi-dimensional variable will be a key step to obtain results on the observation vector $\mathcal{O}^n$.

Theorem 4. *Assume* (R), *then the vector* $(U^{n,\theta}, V^{n,\theta})$ *admits a density* $\mathfrak{p}_{x_0}^n(\cdot,\cdot,\theta)$ *on* $\mathbb{R}^2$, *and there exist two constants* $c_1 > c_2 > 0$, *such that*,

$$c_1^{-1}\mathrm{e}^{-c_1(u^2+v^2)} \leqslant \mathfrak{p}_{x_0}^n(u,v,\theta) \leqslant c_2^{-1}\mathrm{e}^{-c_2(u^2+v^2)}. \tag{9}$$

The constants c_1 *and* c_2 *only depend on the bounds on the coefficients* a, b *and their derivatives.*

The proof of this theorem is given in Appendix A.1. The existence of the density is obtained by means of the Malliavin calculus. On the other hand, the upper and lower bounds rely on the direct study of $(U^{n,\theta}, V^{n,\theta})$ around its skeleton (see [12,13] for related works; and [18] for different methods involving Malliavin calculus).

The following is a direct corollary of Theorem 4:

Corollary 1. *The vector* $\mathbb{B}_0 = \{\overline{X}_0, \dots, \overline{X}_{k-1}, X_{k/n}\}$ *admits a positive density.*

Proof. The bi-dimensional process $(\overline{X}_l, X_{(l+1)/n})_{l=0,\dots,k-1}$ is a Markov chain with transition density $p_{x_l}(\overline{x}_l, x_{l+1}, \theta) = n\mathfrak{p}_{x_l}^n(n^{1/2}(\overline{x}_{l+1} - x_l), n^{1/2}(x_{l+1} - x_l), \theta)$. Then it is clear that the vector $\mathbb{B}_0$ admits a positive density. $\square$

2.1.2. *Invertibility of the Malliavin covariance matrix of a block*

Actually the existence of a density for the law of the random variable $\mathbb{B}_0$ will not be sufficient, and we need a nondegeneracy condition for this variable.

Before this, let us precise briefly a few notations from the Malliavin calculus, used in the sequel (see [22,23] for details). We let H be the Hilbert space $\mathbf{L}^2([0,\infty))$ so that the Brownian motion $(W_t)_{t\in[0,\infty)}$, appearing in (6), is canonically associated to this Hilbert space via the standard $\mathbf{L}^2$ isometry. In this setting, for any $p \geqslant 1$ and natural number q, recall that the set $\mathbb{D}^{q,p}$ denotes the space of real valued Wiener functionals with q derivatives and whose derivatives belong to $\mathbf{L}^p(\Omega)$. If we denote by D the derivative operator then the space $\mathbb{D}^{q,p}$ is endowed with the norm, $\|F\|_{q,p} = [E(|F|^p) + \sum_{j=1}^q E(\|D^j F\|^p_{\mathbf{L}^2([0,\infty)^j)})]^{1/p}$. The space of variable with q derivatives in any $\mathbf{L}^p(\Omega)$ is denoted $\mathbb{D}^{q,\infty} = \bigcap_{p\geqslant 1}\mathbb{D}^{q,p}$. These definitions can be extended to random variables with values in any Hilbert space V and the corresponding spaces are denoted $\mathbb{D}^{q,p}(V)$, $\mathbb{D}^{q,\infty}(V)$ (see Section 1.5 in [22]). In particular, the operator D is then well defined from $\mathbb{D}^{q,\infty}$ to $\mathbb{D}^{q-1,\infty}(H)$. Finally, the adjoint operator of D is the Skohorod integral δ, and the Malliavin covariance matrix of an element $F \in \mathbb{D}^{1,\infty}(\mathbb{R}^d)$ is defined as the matrix $\gamma_{F_1,\dots,F_d} = [\langle Df_i, D\cdot F_j\rangle_H]_{1\leqslant i,j\leqslant d}$.

Now, we consider the variables

$$U_0^{n,\theta} := \int_0^1 \mathcal{X}_s^{n,\theta}\,\mathrm{d}\mu(s), \tag{10}$$

$$U_1^{n,\theta} := \int_0^1 \big(\mathcal{X}_{s+1}^{n,\theta} - \mathcal{X}_s^{n,\theta}\big)\,\mathrm{d}\mu(s), \tag{11}$$

$$\vdots$$

$$U_{k-1}^{n,\theta} := \int_0^1 \big(\mathcal{X}_{s+k-1}^{n,\theta} - \mathcal{X}_{s+k-2}^{n,\theta}\big)\,\mathrm{d}\mu(s), \tag{12}$$

$$U_k^{n,\theta} := \int_0^1 \big(\mathcal{X}_k^{n,\theta} - \mathcal{X}_{s+k-1}^{n,\theta}\big)\,\mathrm{d}\mu(s). \tag{13}$$

Note that the joint law of these $k+1$ variables is, by rescaling, the same as the law of the vector composed with variables of the first block $\mathbb{B}_0$: $n^{1/2}(\overline{X}_0^\theta - x_0, \overline{X}_1^\theta - \overline{X}_0^\theta, \ldots, \overline{X}_{k-1}^\theta - \overline{X}_{k-2}^\theta, X_{k/n}^\theta - \overline{X}_{k-1}^\theta)$. These variables satisfy the following nondegeneracy property whose proof is postponed to Appendix A.1.3.

Proposition 1. *Under* (R), $(U_0^{n,\theta}, \ldots, U_k^{n,\theta}) \in \mathbb{D}^{3,\infty}$. *Denote by $K(\theta)$ the Malliavin covariance matrix of $(U_0^{n,\theta}, \ldots, U_k^{n,\theta})$. It is a.s. an invertible matrix and for all $p \geqslant 1$, we have*

$$E\big(\big|\det\big(K(\theta)\big)\big|^{-p}\big) \leqslant c(p,k).$$

2.2. An exact expression using Malliavin calculus

In this section we intend to give an exact expression for the score function of the observation of $\mathbb{B}_0$ or equivalently for the vector $(U_0^{n,\theta}, \ldots, U_k^{n,\theta})$ given by (10)–(13).

Under the condition (R), we know that there exists a version of the solution of (6) such that P-almost surely the function $\theta \to \mathcal{X}_t^\theta$ is continuously differentiable for all t and $\tau_t^{n,\theta} := \frac{\partial \mathcal{X}_t^{n,\theta}}{\partial\theta}$ is a solution of the stochastic equation (see [20]):

$$\begin{aligned} \mathrm{d}\tau_t^{n,\theta} &= \frac{\partial a_n}{\partial x}\big(\mathcal{X}_t^{n,\theta},\theta\big)\tau_t^{n,\theta}\,\mathrm{d}W_t + \frac{\partial a_n}{\partial \theta}\big(\mathcal{X}_t^{n,\theta},\theta\big)\,\mathrm{d}W_t + \frac{\partial b_n}{\partial x}\big(\mathcal{X}_t^{n,\theta}\big)\tau_t^{n,\theta}\,\mathrm{d}t,\\ \tau_0^{n,\theta} &= 0. \end{aligned} \tag{14}$$

The main result of this section is an explicit representation for the derivative of the log-likelihood of one block. This extends a former result given by Gobet (see [10,11]).

Theorem 5. *The random vector $(U_0^{n,\theta}, \ldots, U_k^{n,\theta})$ admits a positive density on $\mathbb{R}^{k+1}$, denoted by $p_{x_0}(u_0, \ldots, u_k, \theta)$. For a.e. $(u_0, \ldots, u_k)$, this density is an absolutely continuous function with respect to the parameter θ and we have the formula*:

$$\frac{\dot{p}_{x_0}}{p_{x_0}}(u_0, \ldots, u_k, \theta) = E\Bigg[\delta\Bigg(\sum_{0\leqslant j,j'\leqslant k} \frac{\partial U_j^{n,\theta}}{\partial\theta} K(\theta)^{-1}_{j,j'} DU_{j'}^{n,\theta}\Bigg)\Big|\big(U_j^{n,\theta} = u_j\big)_{j=0,\ldots,k}\Bigg],$$

where $K(\theta)^{-1}$ is the inverse of the Malliavin covariance matrix of $(U_0^{n,\theta}, \ldots, U_k^{n,\theta})$.

Proof. Denote $U^{n,\theta}$ the Wiener functional, $U^{n,\theta} = (U_0^{n,\theta}, \dots, U_k^{n,\theta})$ and let $f : \mathbb{R}^{k+1} \to \mathbb{R}$ be a smooth function with compact support. Then the function $\theta \mapsto E[f(U^{n,\theta})]$ can be differentiated pointwise and:

$$\frac{\partial}{\partial\theta} E\bigl[f\bigl(U^{n,\theta}\bigr)\bigr] = E\left[\sum_{j=0}^{k} \frac{\partial f}{\partial u_j}\bigl(U^{n,\theta}\bigr)\frac{\partial U_j^{n,\theta}}{\partial\theta}\right].$$

By Proposition 1, the Malliavin covariance matrix of $U^{n,\theta}$ is invertible and a standard computation on Wiener functionals (see formula (2.4), p. 81 in [22]) shows that: $\frac{\partial f}{\partial u_j}(U^{n,\theta}) = \sum_{j'=0}^{k} \langle D(f(U^{n,\theta})), DU_{j'}^{n,\theta}\rangle_H K(\theta)_{j,j'}^{-1}$. It follows that $\frac{\partial}{\partial\theta} E[f(U^{n,\theta})]$ is equal to

$$E\left[\sum_{j=0}^{k}\sum_{j'=0}^{k} \bigl\langle D\bigl(f\bigl(U^{n,\theta}\bigr)\bigr), DU_{j'}^{n,\theta}\bigr\rangle_H K(\theta)_{j,j'}^{-1} \frac{\partial U_j^{n,\theta}}{\partial\theta}\right] = E\bigl[\bigl\langle D\bigl(f\bigl(U^{n,\theta}\bigr)\bigr), L^\theta\bigr\rangle_H\bigr],$$

where L^θ is the H-valued random variable:

$$L^\theta := \sum_{j=0}^{k}\sum_{j'=0}^{k} \frac{\partial U_j^{n,\theta}}{\partial\theta} K(\theta)_{j,j'}^{-1} DU_{j'}^{n,\theta}. \tag{15}$$

Introducing δ the adjoint operator of D, we get

$$\frac{\partial}{\partial\theta} E\bigl[f\bigl(U^{n,\theta}\bigr)\bigr] = E\bigl[f\bigl(U^{n,\theta}\bigr)\delta\bigl(L^\theta\bigr)\bigr]. \tag{16}$$

Let g be any smooth function with compact support on $\mathbb{R}$. Using the integration by part formula and the equation (16) we have:

$$\begin{aligned}\int \mathrm{d}\theta\, \dot g(\theta) E\bigl(f\bigl(U^{n,\theta}\bigr)\bigr) &= -\int \mathrm{d}\theta\, g(\theta)\frac{\partial}{\partial\theta}E\bigl[f\bigl(U^{n,\theta}\bigr)\bigr] = -\int \mathrm{d}\theta\, g(\theta) E\bigl[f\bigl(U^{n,\theta}\bigr)\delta\bigl(L^\theta\bigr)\bigr]\\ &= -\int \mathrm{d}\theta\, g(\theta) E\bigl[f\bigl(U^{n,\theta}\bigr)E\bigl[\delta\bigl(L^\theta\bigr)\big|\bigl(U_0^{n,\theta}, \dots, U_k^{n,\theta}\bigr)\bigr]\bigr].\end{aligned}$$

Introducing the density of the random vector $U^{n,\theta}$ the equation above writes,

$$\begin{aligned}&\int \dot g(\theta)\,\mathrm{d}\theta \int f(u_0, \dots, u_k) p_{x_0}(u_0, \dots, u_k, \theta)\,\mathrm{d}u_0 \cdots \mathrm{d}u_k\\ &\quad = -\int g(\theta)\,\mathrm{d}\theta \int f(u_0, \dots, u_k) E\bigl[\delta\bigl(L^\theta\bigr)\big|\bigl(U_l^{n,\theta} = u_l\bigr)_l\bigr] p_{x_0}(u_0, \dots, u_k, \theta)\,\mathrm{d}u_0 \cdots \mathrm{d}u_k.\end{aligned}$$

Now using Fubini's theorem it can be seen that $\mathrm{d}u_0 \cdots \mathrm{d}u_k$-almost everywhere the function $\theta \to p_{x_0}(u_0, \dots, u_k, \theta)$ is absolutely continuous with

$$\dot p_{x_0}(u_0, \dots, u_k, \theta) = E\bigl[\delta\bigl(L^\theta\bigr)\big|\bigl(U_l^{n,\theta} = u_l\bigr)_l\bigr] p_{x_0}(u_0, \dots, u_k, \theta).$$

Hence the theorem is proved. □

Remark that the proof of Theorem 5 does not rely on the specific expression (10)–(13) and thus an analogous representation for the score function seems achievable in many situations.

2.3. *A Gaussian approximation for the log-likelihood*

In this section we intend to give a tractable approximation for the score function of $(U_0^{n,\theta}, \dots, U_k^{n,\theta})$.

2.3.1. Approximation for the diffusion

We introduce $\tilde{\mathcal{X}}_t^\theta = a(x_0, \theta) W_t$ and $\tilde{\tau}_t^\theta = \dot{a}(x_0, \theta) W_t$ which stand – by (6) and (14) – for the first-order approximations of $\mathcal{X}_t^{n,\theta}$ and $\tau_t^{n,\theta} = \frac{\partial \mathcal{X}_t^{n,\theta}}{\partial \theta}$. Then, we consider the quantities obtained by replacing in (10)–(13) the process $\mathcal{X}$ by this Gaussian approximation

$$\widetilde{U}_0^\theta := a(x_0, \theta) \int_0^1 W_s \, \mathrm{d}\mu(s) = a(x_0, \theta) \int_0^1 \mu\big([s, 1]\big) \, \mathrm{d}W_s, \tag{17}$$

$$\begin{aligned} \widetilde{U}_j^\theta &:= a(x_0, \theta) \int_0^1 (W_{j+s} - W_{j-1+s}) \, \mathrm{d}\mu(s), \quad \text{for } j = 1, \ldots, k-1, \\ &= a(x_0, \theta) \int_{j-1}^{j} \mu\big(\big[0, s - (j-1)\big]\big) \, \mathrm{d}W_s + a(x_0, \theta) \int_j^{j+1} \mu\big([s - j, 1]\big) \, \mathrm{d}W_s, \end{aligned} \tag{18}$$

$$\widetilde{U}_k^\theta := a(x_0, \theta) \int_0^1 (W_k - W_{k-1+s}) \, \mathrm{d}\mu(s) = a(x_0, \theta) \int_{k-1}^{k} \mu\big(\big[0, s - (k-1)\big]\big) \, \mathrm{d}W_s, \tag{19}$$

where we have repeatedly used the Fubini theorem for stochastic integrals (see [25], p. 176). In the next lemma, we control the difference between the $U_j^{n,\theta}$ and their approximation in terms of Sobolev norm.

Lemma 1. *For all $k, p > 1$, there exist constants $c(k, p)$, $c(p)$ such that for all $j \in \{0, \ldots, k\}$:*

$$\big\| U_j^{n,\theta} - \widetilde{U}_j^\theta \big\|_{2,p} \leqslant c(k, p) n^{-1/2}, \qquad \big\| \widetilde{U}_j^\theta \big\|_{3,p} \leqslant c(p), \tag{20}$$

$$\left\| \frac{\partial U_j^{n,\theta}}{\partial \theta} - \frac{\partial \widetilde{U}_j^{n,\theta}}{\partial \theta} \right\|_{2,p} \leqslant c(k, p) n^{-1/2}, \qquad \left\| \frac{\partial \widetilde{U}_j^\theta}{\partial \theta} \right\|_{3,p} \leqslant c(p), \tag{21}$$

$$\forall 0 \leqslant j, j' \leqslant k, \quad \big| E\big(U_j^{n,\theta} U_{j'}^{n,\theta} - \widetilde{U}_j^\theta \widetilde{U}_{j'}^\theta\big) \big| \leqslant c(k) n^{-1}. \tag{22}$$

Proof. The inequalities on the right-hand side of (20), (21) are immediate by the definition of $\widetilde{U}_j^\theta$.

Comparing expressions of (10)–(13) with (17)–(19), the two remaining bounds in (20) and (21) will be a consequence of the Minkowski inequality – for the Sobolev norm – and of the control on the diffusions:

$$\sup_{t \leqslant k} \big\| \mathcal{X}_t^{n,\theta} - \tilde{\mathcal{X}}_t^\theta \big\|_{2,p} + \sup_{t \leqslant k} \big\| \tau_t^{n,\theta} - \tilde{\tau}_t^\theta \big\|_{2,p} \leqslant n^{-1/2} c(k, p).$$

We only prove the control on $\mathcal{X}^{n,\theta}$ since the proof for $\tau^{n,\theta}$ is analogous. Recalling (6) and (7), we can write

$$\begin{aligned} \mathcal{X}_t^{n,\theta} - \tilde{\mathcal{X}}_t^\theta &= \int_0^t \big[a_n\big(\mathcal{X}_s^{n,\theta}, \theta\big) - a(x_0, \theta)\big] \, \mathrm{d}W_s + \int_0^t b_n\big(\mathcal{X}_s^{n,\theta}\big) \, \mathrm{d}s \\ &= \frac{1}{\sqrt{n}} \int_0^t \int_0^1 a'_x\left(x_0 + \frac{u \mathcal{X}_s^{n,\theta}}{\sqrt{n}}, \theta\right) \mathcal{X}_s^{n,\theta} \, \mathrm{d}u \, \mathrm{d}W_s + \frac{1}{\sqrt{n}} \int_0^t b\left(x_0 + \frac{\mathcal{X}_s^{n,\theta}}{\sqrt{n}}\right) \mathrm{d}s. \end{aligned} \tag{23}$$

But we know [22] that under Assumption (R) the variables $\mathcal{X}^{n,\theta}$ belong to $\mathbb{D}^{3,\infty}$ with a control (independent of θ, n): $\sup_{u_1, u_2 \leqslant s \leqslant k} E(|D^2_{u_1, u_2} \mathcal{X}_s^{n,\theta}|^p) \leqslant c(p, k)$. This is sufficient to deduce $\| \mathcal{X}_t^{n,\theta} - \tilde{\mathcal{X}}_t^\theta \|_{2,p} \leqslant n^{-1/2} c(p, k)$ after a few computations.

To obtain (22) note that by (20) it is sufficient to show $E((U_j^{n,\theta} - \widetilde{U}_j^\theta) \widetilde{U}_{j'}^\theta) \leqslant c(k) n^{-1}$. This property will follow again from an analogous relation on the diffusion,

$$\sup_{t, t' \leqslant k} \big| E\big(\big(\mathcal{X}_t^{n,\theta} - \tilde{\mathcal{X}}_t^\theta\big) \tilde{\mathcal{X}}_{t'}^\theta\big) \big| \leqslant c(k) n^{-1}.$$

Indeed, from (23), the above expectation is equal to

$$n^{-1/2}\int_0^{t\wedge t'}\int_0^1 E\big[a_x'\big(x_0+n^{-1/2}u\mathcal{X}_s^{n,\theta},\theta\big)\mathcal{X}_s^{n,\theta}\big]\,\mathrm{d}u\,a(x_0,\theta_0)\,\mathrm{d}s$$
$$+\,n^{-1/2}\int_0^t E\big[b\big(x_0+n^{-1/2}\mathcal{X}_s^{n,\theta}\big)W_{t'}\big]a(x_0,\theta)\,\mathrm{d}s.$$

Using $|E[a_x'(x_0,\theta)\mathcal{X}_s^{n,\theta}]|=|\int_0^s a_x'(x_0,\theta)E[b_n(\mathcal{X}_u^{n,\theta})]\,\mathrm{d}u|\leqslant cn^{-1/2}$, $E[b(x_0)W_{t'}]=0$ and the boundedness of a_{xx}'' and b', we get the required estimate. □

2.3.2. Approximation for the log-likelihood

Let us denote the deterministic tridiagonal matrix $\widetilde{K}$ of size $(k+1)\times(k+1)$,

$$\widetilde{K}=\begin{bmatrix} v_1 & c & 0 & 0 & 0\\ c & v_1+v_2 & \ddots & 0 & 0\\ 0 & \ddots & \ddots & \ddots & 0\\ 0 & 0 & \ddots & v_1+v_2 & c\\ 0 & 0 & 0 & c & v_2\end{bmatrix},$$

where the entries of the matrix are:

$$v_1=\int_0^1\mu\big([s,1]\big)^2\mathrm{d}s,\qquad v_2=\int_0^1\mu\big([0,s]\big)^2\,\mathrm{d}s,\qquad c=\int_0^1\mu\big([0,s]\big)\mu\big([s,1]\big)\,\mathrm{d}s.$$

It can be easily checked that $a^2(x_0,\theta)\widetilde{K}$ is the covariance matrix of the Gaussian vector $(\widetilde{U}_0^\theta,\dots,\widetilde{U}_k^\theta)$ and that it is invertible using (3). Now the idea is to introduce the score function that would be produced from the observation of this Gaussian vector. Hence we let:

$$\mathcal{L}_{x_0}(u_0,\dots,u_k,\theta)=\frac{\dot{a}}{a}(x_0,\theta)\bigg\{a(x_0,\theta)^{-2}\sum_{0\leqslant j,j'\leqslant k}u_j\widetilde{K}_{j,j'}^{-1}u_{j'}-(k+1)\bigg\}. \tag{24}$$

In this section, we will show that this quantity is an approximation for the true score function $\frac{\dot{p}}{p}$.

Theorem 6. *Let us consider the difference*

$$\frac{\dot{p}_{x_0}}{p_{x_0}}(u_0,\dots,u_k,\theta)-\mathcal{L}_{x_0}(u_0,\dots,u_k,\theta):=r_{x_0}(u_0,\dots,u_k,\theta). \tag{25}$$

Then we have the following bounds:

$$\big|E\big[r_{x_0}\big(U_0^{n,\theta},\dots,U_k^{n,\theta},\theta\big)\big]\big|\leqslant c(k)n^{-1}, \tag{26}$$

$$\forall p\geqslant 1,\quad E\big[\big|r_{x_0}\big(U_0^{n,\theta},\dots,U_k^{n,\theta},\theta\big)\big|^p\big]^{1/p}\leqslant c(k,p)n^{-1/2}. \tag{27}$$

Proof. Keeping in mind the definition of L^θ (see (15)), we introduce its approximation based on the Gaussian quantities defined above:

$$\widetilde{L}^\theta:=\sum_{j=0}^k\sum_{j'=0}^k\frac{\partial\widetilde{U}_j^\theta}{\partial\theta}a(x_0,\theta)^{-2}\widetilde{K}_{j,j'}^{-1}D\widetilde{U}_{j'}^\theta.$$

The first step is to obtain the following control on the difference $r_1:=L^\theta-\widetilde{L}^\theta$:

$$\forall p>1,\quad \|r_1\|_{\mathbb{D}^{1,p}(H)}\leqslant c(k,p)n^{-1/2}. \tag{28}$$

Actually, it is a easy consequence of Lemma 1, Proposition 1 and the invertibility of $\widetilde{K}$, noting that the Malliavin covariance matrix of $\widetilde{U}^\theta$ coincides with the covariance matrix $a^2(x_0,\theta)\widetilde{K}$ of the Gaussian vector $\widetilde{U}^\theta$. We omit further details.

The second step is to obtain a simple expression for $\delta(\widetilde{L}^\theta)$. To see this, we first use the relation for $F \in \mathbb{D}^{1,\infty}$, $u \in \mathbb{D}^{1,\infty}(H)$, $\delta(Fu) = F\delta(u) - \langle D\cdot F, u\rangle_H$ (see [22]):

$$\delta(\widetilde{L}^\theta) = \sum_{j=0}^{k}\sum_{j'=0}^{k} \frac{\partial \widetilde{U}_j^\theta}{\partial\theta} a(x_0,\theta)^{-2}\widetilde{K}^{-1}_{j,j'}\delta\big(D\big(\widetilde{U}^\theta_{j'}\big)\big) - \sum_{j=0}^{k}\sum_{j'=0}^{k} a(x_0,\theta)^{-2}\widetilde{K}^{-1}_{j,j'}\left\langle D\frac{\partial \widetilde{U}_j^\theta}{\partial\theta}, D\widetilde{U}^\theta_{j'}\right\rangle_H.$$

On the one hand, $\delta(D(\widetilde{U}^\theta_{j'})) = \widetilde{U}^\theta_{j'}$ ($\delta\circ D$ is the identity operator on the first chaos space). On the other hand, one has $\frac{\partial \widetilde{U}_j^\theta}{\partial\theta} = \frac{\dot a(x_0,\theta)}{a(x_0,\theta)}\widetilde{U}_j^\theta$ by (17)–(19). We deduce

$$\begin{aligned}\delta(\widetilde{L}^\theta) &= \frac{\dot a(x_0,\theta)}{a(x_0,\theta)}\sum_{j=0}^{k}\sum_{j'=0}^{k}\widetilde{U}_j^\theta a(x_0,\theta)^{-2}\widetilde{K}^{-1}_{j,j'}\widetilde{U}^\theta_{j'} - \frac{\dot a(x_0,\theta)}{a(x_0,\theta)}\sum_{j=0}^{k}\sum_{j'=0}^{k} a(x_0,\theta)^{-2}\widetilde{K}^{-1}_{j,j'}\big\langle D\widetilde{U}_j^\theta, D\widetilde{U}^\theta_{j'}\big\rangle_H \\ &= \frac{\dot a(x_0,\theta)}{a(x_0,\theta)}\sum_{j=0}^{k}\sum_{j'=0}^{k}\widetilde{U}_j^\theta a(x_0,\theta)^{-2}\widetilde{K}^{-1}_{j,j'}\widetilde{U}^\theta_{j'} - \frac{\dot a(x_0,\theta)}{a(x_0,\theta)}(k+1).\end{aligned}$$

Now set

$$r_2 = \frac{\dot a(x_0,\theta)}{a^3(x_0,\theta)}\sum_{0\leqslant j,j'\leqslant k}\widetilde{U}_j^\theta\widetilde{K}^{-1}_{j,j'}\widetilde{U}^\theta_{j'} - \frac{\dot a(x_0,\theta)}{a^3(x_0,\theta)}\sum_{0\leqslant j,j'\leqslant k}U_j^{n,\theta}\widetilde{K}^{-1}_{j,j'}U^{n,\theta}_{j'},$$

and take the conditional expectation in the relation $\delta(L^\theta) = \delta(\widetilde{L}^\theta) + \delta(r_1)$: by Theorem 5, we get (25) with $r_{x_0}(u_0,\ldots,u_k,\theta) = E(\delta(r_1)|(U_j^{n,\theta})_j = (u_j)_j) + E(r_2|(U_j^{n,\theta})_j = (u_j)_j)$.

The final step in the proof is to show that r_{x_0} satisfies conditions (26), (27). For the first condition, since the Skorohod integral has zero mean, we have $E[r_{x_0}(U_0^{n,\theta},\ldots,U_k^{n,\theta},\theta)] = E(r_2)$ and we conclude using (22).

We now prove (27). The conditional expectation being a contraction on $\mathbf{L}^p$ it is sufficient to prove

$$E\big(\big|\delta(r_1)\big|^p\big)^{1/p} \leqslant c(p,k)n^{-1/2}, \qquad E\big(|r_2|^p\big)^{1/p} \leqslant c(p,k)n^{-1/2}.$$

The first estimate follows from (28) and the continuity of the operator δ from $\mathbb{D}^{1,p}(H)$ to $\mathbf{L}^p$. The second one is an immediate consequence of Lemma 1. □

Remark 3. *Let us note that the constants $c(k), c(k,p)$ in Theorem 6 should increase as the block length k goes to infinity since the Gaussian approximation ceases to be valid in that case. However, in the sequel we shall not need a precise evaluation of this dependence on k since we will have the possibility to conveniently choose the growth rate of $k = k_n$.*

In the following sections we will need this corollary of Theorem 6.

Corollary 2. *We have for all $p > 1$,*

$$E\left[\left|\frac{\dot p_{x_0}}{p_{x_0}}\big(U_0^{n,\theta},\ldots,U_k^{n,\theta},\theta\big)\right|^p\right] \leqslant c(k,p).$$

Proof. By Theorem 6 it is sufficient to show that $E[|\mathcal{L}_{x_0}(U_0^{n,\theta},\ldots,U_k^{n,\theta},\theta)|^p] \leqslant c(k,p)$. But from the expression of $\mathcal{L}_{x_0}$, this estimate is clear. □

3. Asymptotic study for the augmented model

In this section we establish Theorem 2. Let us recall some notations: we now deal with the diffusion given by (1) and (2); k_n is some integer in $\{1, \dots, n\}$, $L_n = \lfloor n/k_n \rfloor$ and our observation consists of the $L_n + 1$ blocks $\mathbb{B}_0, \dots, \mathbb{B}_{L_n}$ described in Section 1. The length of the block $\mathbb{B}_l$ is $k_{n,l} + 1$, where $k_{n,l} = k_n$ if $l \leqslant L_n - 1$ and $k_{n,L_n} = n - L_n k_n$. For sake of simplicity in the sequel we sometimes omit the dependence with respect to n and l of the block size, and let $k_{n,l} = k$ with a slight abuse of notation in particular for the last block of data.

To be able to use the results of the Section 2, we introduce on each block the random variables corresponding to the definitions (10)–(13) for the first block. Hence for $l \in \{0, \dots, L_n\}$, we define the $k_{n,l} + 1$ following variables:

$$
\begin{aligned}
U_{0,l} &= n^{1/2}(\overline{X}_{kl} - X_{kl/n}),\\
U_{1,l} &= n^{1/2}(\overline{X}_{kl+1} - \overline{X}_{kl}),\\
&\ \vdots\\
U_{k-1,l} &= n^{1/2}(\overline{X}_{kl+k-1} - \overline{X}_{kl+k-2}),\\
U_{k,l} &= n^{1/2}(X_{k(l+1)/n} - \overline{X}_{kl+k-1}).
\end{aligned}
$$

Clearly the observation of the $(U_{j,l})$ for $l \in \{0, \dots, L_n\}$, $j \in \{0, \dots, k_{n,l}\}$ is equivalent to the observation of the $L_n + 1$ blocks. Using the Markov property for the process X it appears that the law of the vector $(U_{j,l})_{j=0,\dots,k_{n,l}}$ conditionally to all the variables $U_{j,l'}$ with $l' < l$, $j \in \{0, \dots, k_{n,l'}\}$ is the same as conditionally to $X_{kl/n}$ only; moreover this law – conditionally to $X_{kl/n} = x_0$ – coincides with that of the vector $(U_0^{n,\theta}, \dots, U_k^{n,\theta})$ studied in Section 2. Thus it admits the density $p_{X_{kl/n}}(u_0, \dots, u_k, \theta)$ studied in Sections 2.2 and 2.3. Hence, the log-likelihood of the augmented model admits the additive structure:

$$
\begin{aligned}
\ln\bigl(Z^{n,\mathrm{aug}}_{\theta_0,\theta_0+u_nh}\bigl(\mathcal{O}^{n,\mathrm{aug}}\bigr)\bigr) &= \sum_{l=0}^{L_n} \ln \frac{p_{X_{kl/n}}(U_{0,l}, \dots, U_{k,l}, \theta_0 + u_n h)}{p_{X_{kl/n}}(U_{0,l}, \dots, U_{k,l}, \theta_0)}\\
&= \sum_{l=0}^{L_n} \int_{\theta_0}^{\theta_0+u_nh} \frac{\dot{p}_{X_{kl/n}}(U_{0,l}, \dots, U_{k,l}, s)}{p_{X_{kl/n}}(U_{0,l}, \dots, U_{k,l}, s)}\,\mathrm{d}s.
\end{aligned}
$$

Owing to Theorem 6, we deduce the decomposition

$$
\ln\bigl(Z^{n,\mathrm{aug}}_{\theta_0,\theta_0+u_nh}\bigl(\mathcal{O}^{n,\mathrm{aug}}\bigr)\bigr) = \sum_{l=0}^{L_n} \int_{\theta_0}^{\theta_0+u_nh} \mathcal{L}_{X_{kl/n}}(U_{0,l}, \dots, U_{k,l}, s)\,\mathrm{d}s + \sum_{l=0}^{L_n} \int_{\theta_0}^{\theta_0+u_nh} r_{X_{kl/n}}(U_{0,l}, \dots, U_{k,l}, s)\,\mathrm{d}s.
$$

In the above decomposition, we will show in Sections 3.1 and 3.2 that the explicit term involving $\mathcal{L}_{x_0}$ governs the asymptotic behavior of the log-likelihood ratio; the other term does not contribute in the limit.

3.1. Proof of Theorem 2: the explicit term

Let us introduce a slight modification of $\mathcal{L}_{X_{kl/n}}$, which has the advantage of being a smoother function w.r.t. θ:

$$
\xi_{l,n}(\theta) = \frac{\dot{a}}{a}(X_{kl/n}, \theta_0)\left\{a(X_{kl/n}, \theta)^{-2} \sum_{0 \leqslant j,j' \leqslant k} U_{j,l} \widetilde{K}^{-1}_{j,j'} U_{j',l} - (k+1)\right\}, \tag{29}
$$

and we set $N_n^{\mathrm{aug}} = u_n \sum_{l=0}^{L_n} \xi_{l,n}(\theta_0)$ and $I_n^{\mathrm{aug}} = -u_n^2 \sum_{l=0}^{L_n} \frac{\partial \xi_{l,n}}{\partial \theta}(\theta_0)$.

Proposition 2. *If $k_n \to \infty$ slowly enough,*

$$
\sum_{l=0}^{L_n} \int_{\theta_0}^{\theta_0+u_nh} \mathcal{L}_{X_{kl/n}}(U_{0,l}, \dots, U_{k,l}, s)\,\mathrm{d}s = h N_n^{\mathrm{aug}} - \frac{h^2}{2} I_n^{\mathrm{aug}} + R_n, \tag{30}
$$

where $I_n^{\mathrm{aug}} \xrightarrow[n\to\infty]{\mathbb{P}^{\theta_0}} \mathcal{I}_{\theta_0}$, $R_n \xrightarrow[n\to\infty]{\mathbb{P}^{\theta_0}} 0$ *and there exists an extra random variable* $N \sim \mathcal{N}(0,1)$ *independent of the process* X *such that*, N_n^{aug} *converges stably in law under* $\mathbb{P}^{\theta_0}$ *to* $N\sqrt{\mathcal{I}_{\theta_0}}$.

Proof. Comparing (24) with the definition of $\xi_{l,n}(\theta)$ above and using a Taylor expansion for $\xi_{l,n}(\theta)$ around θ_0, we get Eq. (30) with a remainder term $R_n = R_n^{(1)} + R_n^{(2)}$ satisfying:

$$R_n^{(1)} = \sum_{l=0}^{L_n} \int_{\theta_0}^{\theta_0+u_n h} \left[\frac{\dot a}{a}(X_{kl/n}, s) - \frac{\dot a}{a}(X_{kl/n}, \theta_0)\right] \left\{ \sum_{0\leqslant j,j'\leqslant k} \frac{U_{j,l}\widetilde{K}_{j,j'}^{-1} U_{j',l}}{a(X_{kl/n}, s)^2} - (k+1)\right\} \mathrm{d}s, \tag{31}$$

$$\left|R_n^{(2)}\right| \leqslant c \sum_{l=0}^{L_n} u_n^{2+\gamma} \left\{ \sum_{0\leqslant j,j'\leqslant k} \left|U_{j,l}\widetilde{K}_{j,j'}^{-1} U_{j',l}\right| \right\} \tag{32}$$

(for $R_n^{(2)}$ we have used that $\theta \mapsto \dot a(x,\theta)$ is γ-Hölder continuous). To complete the proof, we repeatedly use the following classical convergence result about triangular arrays of random variables.

Lemma 2 ([5], Lemma 9). *Let* $(\chi_l^n)_{0\leqslant l\leqslant L_n}$, U *be random variables, with* χ_l^n *being* $\mathcal{F}_{l+1}^n$*-measurable. The two following conditions imply* $\sum_{l=0}^{L_n} \chi_l^n \xrightarrow{\mathbb{P}} U$:

$$\sum_{l=0}^{L_n} E\left[\chi_l^n \middle| \mathcal{F}_l^n\right] \xrightarrow{\mathbb{P}} U \quad \textit{and} \quad \sum_{l=0}^{L_n} E\left[\left(\chi_l^n\right)^2 \middle| \mathcal{F}_l^n\right] \xrightarrow{\mathbb{P}} 0.$$

- *We first focus on* N_n^{aug}. Let us introduce the sigma field $\mathcal{F}_l^n = \sigma(X_0; B_s, s \leqslant \frac{kl}{n})$ for $l = 0, \dots, L_n$ and $\mathcal{F}_{L_n+1}^n = \sigma(X_0; B_s, s \leqslant 1)$. Then the variable $\xi_{l,n}(\theta_0)$ is $\mathcal{F}_{l+1}^n$-measurable and the asymptotic behavior of N_n^{aug} will follow from Lemma 2. To make clearer this point we introduce the following approximation based on conditionally Gaussian variables:

$$\widetilde{\xi}_{l,n}(\theta) = \frac{\dot a}{a}(X_{kl/n}, \theta_0) \left\{ a(X_{kl/n}, \theta)^{-2} \sum_{0\leqslant j,j'\leqslant k} \widetilde{U}_{j,l} \widetilde{K}_{j,j'}^{-1} \widetilde{U}_{j',l} - (k+1) \right\}. \tag{33}$$

Here, $\widetilde{U}_{j,l}$ is the Gaussian approximation under $\mathbb{P}^\theta$ of $U_{j,l}$ corresponding on the block $\mathbb{B}_l$ to the variables (17)–(19) on the block $\mathbb{B}_0$:

$$\widetilde{U}_{0,l} := a(X_{kl/n}, \theta) n^{1/2} \int_0^1 (B_{(kl+s)/n} - B_{kl/n})\, \mathrm{d}\mu(s),$$

$$\widetilde{U}_{j,l} := a(X_{kl/n}, \theta) n^{1/2} \int_0^1 (B_{(kl+j+s)/n} - B_{(kl+j-1+s)/n})\, \mathrm{d}\mu(s) \quad \text{for } j = 1, \dots, k-1,$$

$$\widetilde{U}_{k,l} := a(X_{kl/n}, \theta) n^{1/2} \int_0^1 (B_{k(l+1)/n} - B_{(kl+k-1+s)/n})\, \mathrm{d}\mu(s).$$

Observe that this vector $(\widetilde{U}_{j,l})_{j=0,\dots,k}$ has, under $\mathbb{P}^\theta$ and conditionally to $X_{kl/n} = x_0$, the same law as the vector $(\widetilde{U}_j^\theta)_{j=0,\dots,k}$ defined in Section 2.3. Thus its conditional law is Gaussian with covariance matrix $a(X_{kl/n}, \theta)^2 \widetilde{K}$. Hence, the variable $\widetilde{\xi}_{l,n}(\theta_0)$ is $\mathcal{F}_{l+1}^n$-measurable and under $\mathbb{P}^{\theta_0}$, it is conditionally (to $X_{kl/n}$) distributed as a recentered $\chi^2(k+1)$ variable. Thus, we deduce the following four properties:

(1) $u_n \sum_{l=0}^{L_n} E_{\theta_0}[\widetilde{\xi}_{l,n}(\theta_0) | \mathcal{F}_l^n] = 0$;

(2) Using $u_n^2 = 1/n$, $L_n \sim n/k_n \to \infty$ and $k_n \to \infty$, one has

$$\begin{aligned} u_n^2 \sum_{l=0}^{L_n} E_{\theta_0}\big[\big(\widetilde{\xi}_{l,n}(\theta_0)\big)^2 \big| \mathcal{F}_l^n\big] &= u_n^2 \sum_{l=0}^{L_n} 2(k_n+1)\left(\frac{\dot a}{a}\right)^2 (X_{kl/n}, \theta_0) \\ &= \frac{2(k_n+1)}{k_n} \int_0^1 \left(\frac{\dot a}{a}\right)^2 (X_s, \theta_0)\,\mathrm{d}s + \mathrm{o}_{\mathbb{P}^{\theta_0}}(1) \\ &\xrightarrow{\mathbb{P}^{\theta_0}} 2\int_0^1 \left(\frac{\dot a}{a}\right)^2 (X_s, \theta_0)\,\mathrm{d}s = \mathcal{I}_{\theta_0}; \end{aligned} \tag{34}$$

(3) $u_n^4 \sum_{l=0}^{L_n} E_{\theta_0}[|\widetilde{\xi}_{l,n}(\theta_0)|^4 | \mathcal{F}_l^n] \leqslant cn^{-2} L_n k_n^4 \leqslant cn^{-1} k_n^3 \to 0$, if k_n goes to ∞ slowly enough;
(4) $u_n \sum_{l=0}^{L_n} E_{\theta_0}[\widetilde{\xi}_{l,n}(\theta_0)[B_{(k+1)l/n} - B_{kl/n}] | \mathcal{F}_l^n] = 0$.

From these four properties, it follows (see [15]) that $u_n \sum_{l=0}^{L_n} \widetilde{\xi}_{l,n}(\theta_0)$ converges stably under $\mathbb{P}^{\theta_0}$ to a mixed Gaussian variable as in the statement of the proposition. To obtain the limit for N_n^{aug}, it is sufficient to prove that

$$N_n^{\mathrm{aug}} - u_n \sum_{l=0}^{L_n} \widetilde{\xi}_{l,n}(\theta_0) \xrightarrow{\mathbb{P}^{\theta_0}} 0. \tag{35}$$

Due to Lemma 2, a sufficient condition consists in the two following points:

(1) $u_n \sum_{l=0}^{L_n} E_{\theta_0}[\widetilde{\xi}_{l,n}(\theta_0) - \xi_{l,n}(\theta_0) | \mathcal{F}_l^n] \to 0$ in probability;
(2) $u_n^2 \sum_{l=0}^{L_n} E_{\theta_0}[(\widetilde{\xi}_{l,n}(\theta_0) - \xi_{l,n}(\theta_0))^2 | \mathcal{F}_l^n] \to 0$ in probability.

But these two points can be shown using (20) and (22) of Lemma 1 (for k_n slowly increasing).

- *We now study* I_n^{aug}. A direct differentiation of $\xi_{l,n}(\theta)$ (recall (29)) gives

$$\dot\xi_{l,n}(\theta) = \frac{\dot a}{a}(X_{kl/n}, \theta_0) \frac{-2\dot a}{a^3}(X_{kl/n}, \theta) \sum_{0 \leqslant j, j' \leqslant k} U_{l,j} \widetilde{K}_{j,j'}^{-1} U_{l,j'}.$$

Then, with a few computations similar to the study of N_n^{aug}, we obtain (for appropriate k_n):

(1)

$$\begin{aligned} u_n^2 \sum_{l=0}^{L_n} E_{\theta_0}\big[\dot\xi_{l,n}(\theta_0) | \mathcal{F}_l^n\big] &= u_n^2 \sum_{l=0}^{L_n} -2(k_{n,l}+1) \frac{\dot a^2}{a^2}(X_{kl/n}, \theta_0) + \mathrm{O}_{\mathbb{P}^{\theta_0}}\left(\frac{c(k_n)}{\sqrt{n}}\right) \\ &= -2\frac{(k_n+1)}{k_n} \int_0^1 \left(\frac{\dot a}{a}(X_s, \theta_0)\right)^2 \mathrm{d}s + \mathrm{o}_{\mathbb{P}^{\theta_0}}(1) \xrightarrow{\mathbb{P}^{\theta_0}} -\mathcal{I}_{\theta_0}; \end{aligned} \tag{36}$$

(2) $u_n^4 \sum_{l=0}^{L_n} E_{\theta_0}[[\dot\xi_{l,n}(\theta_0)]^2 | \mathcal{F}_l^n] \leqslant cn^{-1} k_n^4 \to 0$.

Combined with Lemma 2, these two convergences imply that of I_n^{aug} to $\mathcal{I}_{\theta_0}$ under $\mathbb{P}^{\theta_0}$.

- *The remainder term* R_n. Firstly, a direct use of (20) gives $E(|R_n^{(2)}|) \leqslant c(k_n) n^{-\gamma/2} \to 0$ if k_n slowly goes to ∞. Secondly, the convergence to zero of $R_n^{(1)} = \sum_{l=0}^{L_n} R_{n,l}^{(1)}$ is more delicate and Lemma 2 is helpful for this. To this end we evaluate the conditional expectation of $R_{n,l}^{(1)}$ using (22) and the fact the $(\widetilde{U}_{j,l})_j$ have the conditional covariance matrix $a(X_{kl/n}, \theta)^2 \widetilde{K}$:

$$E_{\theta_0}\big[R_{n,l}^{(1)} | \mathcal{F}_l^n\big] = \int_{\theta_0}^{\theta_0 + u_n h} \left[\frac{\dot a}{a}(X_{kl/n}, s) - \frac{\dot a}{a}(X_{kl/n}, \theta_0)\right] \left\{\frac{a(X_{kl/n}, \theta_0)^2}{a(X_{kl/n}, s)^2} - 1\right\} (k_n+1)\,\mathrm{d}s + \mathrm{O}\big(n^{-1} u_n c(k_n)\big).$$

The function a being $\mathcal{C}^{1+\gamma}$ in θ, one gets: $\sum_{l=0}^{L_n} |E_{\theta_0}[R_{n,l}^{(1)}|\mathcal{F}_l^n]| \leqslant cn^{-\gamma/2} + \frac{c(k_n)}{k_n} n^{-1/2} \to 0$ for appropriate k_n. With similar considerations we evaluate the second conditional moment and obtain $u_n^2 \sum_{l=0}^{L_n} E_{\theta_0}[(R_n^{(1)})^2|\mathcal{F}_l^n] \leqslant c(k_n) L_n u_n^{2+2\gamma} \overset{n\to\infty}{\longrightarrow} 0$. □

3.2. Proof of Theorem 2: the negligible terms

It remains to prove that, as announced, there is convergence to zero of $\sum_{l=0}^{L_n} \eta_l$ with $\eta_l = \int_{\theta_0}^{\theta_0+u_n h} r_{X_{kl/n}}(U_{0,l}, \ldots, U_{k,l}, s)\,\mathrm{d}s$. We aim at applying Lemma 2 by computing the first two conditional moments of η_l under $\mathbb{P}^{\theta_0}$. The main difficulty here comes from the fact that we do not have an explicit expression for $r_{x_0}((u_j)_j, \theta)$. Indeed by Theorem 6 we know bounds for the moments $E_{\theta,x_0}^n(|r_{x_0}((U_j)_j, \theta)|^p)$ where by $E_{x_0,\theta}^n$ we denote the expectation with respect to the law of $\mathcal{X}^{n,\theta}$ solution of (6). This is a priori insufficient to compute the conditional moments of η_l under $\mathbb{P}^{\theta_0}$ which involve quantities such as $E_{\theta_0,x_0}^n(|r_{x_0}((U_j)_j, s)|^p)$ for $s \neq \theta_0$. Thus in Lemmas 7 and 8 in the Appendix we study the transformation of such moments under change of measure.

Firstly, we evaluate the conditional expectation of η_l,

$$E_{\theta_0}\big[\eta_l|\mathcal{F}_l^n\big] = \int_{\theta_0}^{\theta_0+u_n h} E_{\theta_0,x}^n\big[r_x\big((U_j)_j, s\big)\big]_{|x=X_{kl/n}}\,\mathrm{d}s.$$

But $|E_{\theta_0,x}^n[r_x((U_j)_j, s)]| \leqslant |E_{s,x}^n[r_x((U_j)_j, s)]| + |E_{\theta_0,x}^n[r_x((U_j)_j, s)] - E_{s,x}^n[r_x((U_j)_j, s)]|$ can be bounded using (26) and Lemma 8 in the Appendix by $c(k)n^{-1} + |s-\theta_0| E_{s,x}^n[|r_x((U_j)_j, s)|^\alpha]^{1/\alpha}$ for some $\alpha \geqslant 1$. Then by (27) we deduce $|E[\eta_l|\mathcal{F}_l^n]| \leqslant c(k_n)[u_n n^{-1} + u_n^2 n^{-1/2}]$. Finally, a block length k_n slowly increasing guarantees $\sum_{l=0}^{L_n} E_{\theta_0}[\eta_l|\mathcal{F}_l^n] \overset{\mathbb{P}^{\theta_0}}{\longrightarrow} 0$.

Secondly and similarly, owing to Theorem 6 and Lemma 7 in the Appendix, we get $E[\eta_l^2|\mathcal{F}_l^n] \leqslant c(k_n) u_n^2 n^{-1} \to 0$. Therefore, by Lemma 2, we have proved $\sum_{l=0}^{L_n} \eta_l \overset{\mathbb{P}^{\theta_0}}{\longrightarrow} 0$. This ends the proof of Theorem 2.

3.3. Proof of Theorem 3

The proof is essentially the same as that of Theorem 2, the difference in the asymptotic information comes from the difference in the limit of the quantities (34) and (36) when k is fixed.

4. LAMN property for the initial model

In this section we are back to the model where the observation is only $\mathcal{O}^n = (\overline{X}_j)_{j=0,\ldots,n-1}$ and we will prove Theorem 1 by relying on the LAMN property for the augmented model.

A first intermediate result is that one can approximate the log-likelihood of the augmented model by a function of the observation $\mathcal{O}^n$.

Proposition 3. *There exist random variables Γ_n measurable with respect to $\mathcal{O}^n$ such that*:

$$\ln\big(Z_{\theta_0,\theta_0+u_n h}^{n,\mathrm{aug}}\big(\mathcal{O}^{n,\mathrm{aug}}\big)\big) - \Gamma_n \underset{\mathbb{P}^{\theta_0}}{\overset{n\to\infty}{\longrightarrow}} 0.$$

Proof. We have seen in Section 3 that $\ln(Z_{\theta_0,\theta_0+u_n h}^{n,\mathrm{aug}}(\mathcal{O}^{n,\mathrm{aug}})) = hN_n^{\mathrm{aug}} - 1/2h^2 I_n^{\mathrm{aug}} + \mathrm{o}_{\mathbb{P}^{\theta_0}}(1)$ where the quantities N_n^{aug} and I_n^{aug} were defined in Section 3.1.

Thus, the proof of the proposition consists in introducing a proper modification of these quantities which only depends on the observations. We let for $l = 0, \ldots, k_n$

$$\xi_{l,n}^{\mathrm{obs}}(\theta) = \frac{\dot{a}}{a}(\overline{X}_{kl-1}, \theta_0)\Bigg\{a^{-2}(\overline{X}_{kl-1}, \theta) \sum_{1\leqslant j,j'\leqslant k-1} U_{l,j}\widehat{K}_{j,j'}^{-1} U_{l,j'} - (k-1)\Bigg\},$$

with the convention $\overline{X}_{-1} = \xi_0$ is the known initial value of the diffusion and the matrix $a^2(x_0,\theta)\widehat{K}$ is the covariance matrix of the conditionally Gaussian vector $(\widetilde{U}_1^\theta, \dots, \widetilde{U}_{k-1}^\theta)$:

$$\widehat{K} = \begin{bmatrix} v_1+v_2 & c & 0 & 0 \\ c & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & c \\ 0 & 0 & c & v_1+v_2 \end{bmatrix}.$$

Clearly, $\xi_{l,n}^{\mathrm{obs}}(\theta)$ only depends on the observation $\mathcal{O}^n$ since we have suppressed all occurrences of the variables $U_{0,l}$ and $U_{k,l}$ and replaced $X_{kl/n}$ by $\overline{X}_{kl-1}$ in the expression of $\xi_{l,n}(\theta)$ (compare with (29)). Then we let $N_n^{\mathrm{obs}} = u_n \sum_{l=0}^{L_n} \xi_{l,n}^{\mathrm{obs}}(\theta_0)$ and $I_n^{\mathrm{obs}} = -u_n^2 \sum_{l=0}^{L_n} \frac{\partial \xi_{l,n}^{\mathrm{obs}}}{\partial\theta}(\theta_0)$.

- *Study of* $N_n^{\mathrm{aug}} - N_n^{\mathrm{obs}}$. The first step is to consider the conditionally recentered chi-square approximation of $\xi_{l,n}^{\mathrm{obs}}(\theta)$ that we define as:

$$\widetilde{\xi}_{l,n}^{\mathrm{obs}}(\theta) = \frac{\dot{a}}{a}(X_{kl/n},\theta_0)\left\{a^{-2}(X_{kl/n},\theta) \sum_{1\leqslant j,j'\leqslant k-1} \widetilde{U}_{l,j}\widehat{K}_{j,j'}^{-1}\widetilde{U}_{l,j'} - (k-1)\right\}. \tag{37}$$

The first step is to prove the validity of the approximation:

$$u_n \sum_{l=0}^{L_n} \left\{\xi_{l,n}^{\mathrm{obs}}(\theta_0) - \widetilde{\xi}_{l,n}^{\mathrm{obs}}(\theta_0)\right\} \xrightarrow{\mathbb{P}^{\theta_0}} 0. \tag{38}$$

This is done similarly to the proof of $N_n^{\mathrm{aug}} - u_n \sum_{l=0}^{L_n} \widetilde{\xi}_{l,n}(\theta_0) \to 0$ in Proposition 2, by considering the first two conditional moments, but here the first moment is more delicate to handle: the conditional moment $E_{\theta_0}[\xi_{l,n}^{\mathrm{obs}}(\theta_0) - \widetilde{\xi}_{l,n}^{\mathrm{obs}}(\theta_0)|\mathcal{F}_l^n]$ is of the form $(k-1)\{g(X_{kl/n}) - g(\overline{X}_{kl-1})\}h(\overline{X}_{kl-1}) + \mathrm{O}(c(k_n)/n)$ for g and h two $\mathcal{C}^2$ functions. If we abruptly use the relation $\|X_{kl/n} - \overline{X}_{kl-1}\|_{\mathbf{L}^p} \leqslant c(p)n^{-1/2}$ then we only deduce that $u_n \sum_{l=0}^{L_n} E_{\theta_0}[\xi_{l,n}^{\mathrm{obs}}(\theta_0) - \widetilde{\xi}_{l,n}^{\mathrm{obs}}(\theta_0)|\mathcal{F}_l^n]$ remains bounded in probability. To show that it actually converges to zero, we have to apply again Lemma 2 to the new triangular array of variables, $u_n \sum_{l=0}^{L_n}(k-1)\{g(X_{kl/n}) - g(\overline{X}_{kl-1})\}h(\overline{X}_{kl-1})$. Then by rather long computations, using that $\|\overline{X}_{kl-1} - \overline{X}_{kl-2}\|_{\mathbf{L}^p} \leqslant c(p)(k/n)^{1/2}$ and $|E_{\theta_0}[X_{kl/n} - \overline{X}_{kl-1}|\mathcal{F}_{l-1}^n]| \leqslant cn^{-1}$, we can prove,

$$u_n \sum_{l=0}^{L_n} (k-1)\left|E_{\theta_0}\left[\left\{g(X_{kl/n}) - g(\overline{X}_{kl-1})\right\}h(\overline{X}_{kl-1})\middle|\mathcal{F}_{l-1}^n\right]\right| \leqslant c(k)n^{-1/2} \xrightarrow{\mathbb{P}^{\theta_0}} 0,$$

$$u_n^2 \sum_{l=0}^{L_n} (k-1)^2 E_{\theta_0}\left[\left\{g(X_{kl/n}) - g(\overline{X}_{kl-1})\right\}^2 h(\overline{X}_{kl-1})^2\middle|\mathcal{F}_{l-1}^n\right] \leqslant c(k)n^{-1} \xrightarrow{\mathbb{P}^{\theta_0}} 0.$$

Thus, we deduce $u_n \sum_{l=0}^{L_n} E_{\theta_0}[\xi_{l,n}^{\mathrm{obs}}(\theta_0) - \widetilde{\xi}_{l,n}^{\mathrm{obs}}(\theta_0)|\mathcal{F}_l^n] \to 0$. The second condition $u_n^2 \sum_{l=0}^{L_n} E_{\theta_0}[(\xi_{l,n}^{\mathrm{obs}}(\theta_0) - \widetilde{\xi}_{l,n}^{\mathrm{obs}}(\theta_0))^2|\mathcal{F}_l^n] \leqslant c(k)n^{-1} \to 0$ is easily obtained and we deduce (38).

Thus, in view of Eq. (35), it remains to prove that $u_n \sum_{l=0}^{L_n}\{\widetilde{\xi}_{l,n}^{\mathrm{obs}}(\theta_0) - \widetilde{\xi}_{l,n}(\theta_0)\}$ is negligible. But by Lemma 10 in the Appendix, comparing expressions (33) and (37), it appears that conditionally to $\mathcal{F}_l^n$ the random variable $\widetilde{\xi}_{l,n}^{\mathrm{obs}}(\theta_0) - \widetilde{\xi}_{l,n}(\theta_0)$ is a recentered $\chi^2(2)$ variable, and hence the following properties hold:

$$u_n \sum_{l=0}^{L_n} E_{\theta_0}\left(\widetilde{\xi}_{l,n}^{\mathrm{obs}}(\theta_0) - \widetilde{\xi}_{l,n}(\theta_0)\middle|\mathcal{F}_l^n\right) = 0,$$

$$u_n^2 \sum_{l=0}^{L_n} E_{\theta_0}\big(\{\widetilde{\xi}_{l,n}^{\mathrm{obs}}(\theta_0) - \widetilde{\xi}_{l,n}(\theta_0)\}^2 \big| \mathcal{F}_l^n\big) = \sum_{l=0}^{L_n} u_n^2 4 \frac{\dot{a}^2}{a^2}(X_{kl/n}, \theta_0) \leqslant \frac{c}{k_n} \to 0.$$

These two properties imply by Lemma 2 the convergence to 0 under $\mathbb{P}^{\theta_0}$ of $u_n \sum_{l=0}^{L_n} \{\widetilde{\xi}_{l,n}^{\mathrm{obs}}(\theta_0) - \widetilde{\xi}_{l,n}(\theta_0)\}$, and thus $N_n^{\mathrm{aug}} - N_n^{\mathrm{obs}} \xrightarrow{\mathbb{P}^{\theta_0}} 0$.

- *Study of* $I_n^{\mathrm{aug}} - I_n^{\mathrm{obs}}$. Exactly as we proved that I_n^{aug} tends to $\mathcal{I}_{\theta_0}$ we can show that $I_n^{\mathrm{obs}} \to \mathcal{I}_{\theta_0}$. Thus the difference is negligible.

Finally the proposition is obtained by setting $\Gamma_n = h N_n^{\mathrm{obs}} - h^2/2 I_n^{\mathrm{obs}}$. □

Then Theorem 1 is a consequence of the following proposition combined with Proposition 3 and Theorem 2.

Proposition 4. *We have the convergence*,

$$Z_{\theta_0,\theta_0+u_n h}^n - \mathrm{e}^{\Gamma_n} \underset{\mathbb{P}^{\theta_0}}{\overset{n\to\infty}{\longrightarrow}} 0.$$

Proof. The starting point is the relation between the likelihood of the initial and of the augmented model: $Z_{\theta_0,\theta_0+u_n h}^n = E_{\theta_0}[Z_{\theta_0,\theta_0+u_n h}^{n,\mathrm{aug}} | \mathcal{O}^n]$. By Proposition 3 we can write $Z_{\theta_0,\theta_0+u_n h}^{n,\mathrm{aug}} = \mathrm{e}^{\Gamma_n}\mathrm{e}^{\varepsilon_n}$ where ε_n tends to zero in $\mathbb{P}_{\theta_0}$ probability. Using that Γ_n is $\mathcal{O}^n$ measurable we deduce,

$$Z_{\theta_0,\theta_0+u_n h}^n - \mathrm{e}^{\Gamma_n} = E_{\theta_0}\big[\mathrm{e}^{\Gamma_n}\big(\mathrm{e}^{\varepsilon_n} - 1\big)\big|\mathcal{O}^n\big].$$

We now use the inequality $|\mathrm{e}^u - 1| \leqslant (|u| \wedge 1)(\mathrm{e}^u + 1)$ to obtain that $|Z_{\theta_0,\theta_0+u_n h}^n - \mathrm{e}^{\Gamma_n}| \leqslant \alpha_n + \beta_n$ with:

$$\alpha_n = E_{\theta_0}\big[\big(|\varepsilon_n| \wedge 1\big)\mathrm{e}^{\Gamma_n}\big|\mathcal{O}^n\big] = E_{\theta_0}\big[|\varepsilon_n| \wedge 1 |\mathcal{O}^n\big]\mathrm{e}^{\Gamma_n},$$
$$\beta_n = E_{\theta_0}\big[\big(|\varepsilon_n| \wedge 1\big)\mathrm{e}^{\Gamma_n}\mathrm{e}^{\varepsilon_n}\big|\mathcal{O}^n\big] = E_{\theta_0}\big[\big(|\varepsilon_n| \wedge 1\big)Z_{\theta_0,\theta_0+u_n h}^{n,\mathrm{aug}}\big|\mathcal{O}^n\big].$$

It now remains to show the convergence to zero of α_n and β_n.

For α_n, let us notice that $(\mathrm{e}^{\Gamma_n})_n$ is a tight sequence and that $E_{\theta_0}[|\varepsilon_n| \wedge 1 | \mathcal{O}^n]$ converges in $\mathbf{L}^1(\mathbb{P}_{\theta_0})$ norm to zero since,

$$E_{\theta_0}\big[E_{\theta_0}\big[|\varepsilon_n| \wedge 1 |\mathcal{O}^n\big]\big] = E_{\theta_0}\big[|\varepsilon_n| \wedge 1\big] \overset{n\to\infty}{\longrightarrow} 0.$$

For β_n, we have $E_{\theta_0}[\beta_n] = E_{\theta_0+u_n h}[|\varepsilon_n| \wedge 1]$. But the sequence of probabilities $\mathbb{P}^{\theta_0}$ and $\mathbb{P}^{\theta_0+u_n h}$ restricted to the sigma fields $\mathcal{O}^{n,\mathrm{aug}}$ are contiguous (this is a consequence of the LAMN property for the augmented model, see e.g. Proposition 1 in [16]); hence the sequence $(\varepsilon_n)_n$ which is measurable with respect to $\mathcal{O}^{n,\mathrm{aug}}$ and converges to zero in $\mathbb{P}^{\theta_0}$-probability converges also in $\mathbb{P}^{\theta_0+u_n h}$-probability. This implies $E_{\theta_0}[\beta_n] = E_{\theta_0+u_n h}[|\varepsilon_n| \wedge 1] \to 0$. □

Appendix

A.1. Proof of results of Section 2.1

Since the results of Section 2.1 concern only the study of a density for fixed values of θ, we omit the dependence upon θ in our notations. We will prove the results in the following order. First in Appendix A.1.1, we show that the law of the Wiener functional $(U^n, V^n) = (\int_0^1 \mathcal{X}_s^n \,\mathrm{d}\mu(s), \mathcal{X}_1^n)$ admits a density. Then we prove the lower and upper bounds given in Theorem 4 (Appendix A.1.2) and eventually we deduce the Proposition 1 (Appendix A.1.3).

A.1.1. Existence of the density $\mathfrak{p}^n_{x_0}$

We know [22] that under (R) the random variable $\mathcal{X}^n_t$ is an element of $\mathbb{D}^{3,\infty}$ and its first derivative is equal to

$$D_t \mathcal{X}^n_s = \mathbf{1}_{\{t \leqslant s\}} \mathcal{Y}^n_s \left(\mathcal{Y}^n_t\right)^{-1} a_n\left(\mathcal{X}^n_t\right), \tag{39}$$

where $\mathcal{Y}^n$ is the solution of

$$\mathrm{d}\mathcal{Y}^n_t = a_n'\left(\mathcal{X}^n_t\right)\mathcal{Y}^n_t \,\mathrm{d}W_t + b_n'\left(\mathcal{X}^n_t\right)\mathcal{Y}^n_t \,\mathrm{d}t, \qquad \mathcal{Y}^n_0 = 1. \tag{40}$$

In the sequel, we will repeatedly use the positivity of $\mathcal{Y}^n$ and the control

$$E\Bigl(\sup_{t\in[0,1]} \left(\mathcal{Y}^n_t\right)^p\Bigr) + E\Bigl(\sup_{t\in[0,1]} \left(\mathcal{Y}^n_t\right)^{-p}\Bigr) \leqslant c(p). \tag{41}$$

From this we can see that the random variables U^n and V^n are elements of $\mathbb{D}^{3,\infty}$ and using (39) with the linearity of the operator D, we have

$$D_t U^n = \int_0^1 \mathbf{1}_{\{t \leqslant s\}} \mathcal{Y}^n_s \left(\mathcal{Y}^n_t\right)^{-1} a_n\left(\mathcal{X}^n_t\right) \mathrm{d}\mu(s) = a_n\left(\mathcal{X}^n_t\right)\left(\mathcal{Y}^n_t\right)^{-1} \mathbf{1}_{\{t \leqslant 1\}} \int_{[t,1]} \mathcal{Y}^n_s \,\mathrm{d}\mu(s),$$

$$D_t V^n = a_n\left(\mathcal{X}^n_t\right)\mathcal{Y}^n_1\left(\mathcal{Y}^n_t\right)^{-1}\mathbf{1}_{\{t \leqslant 1\}}.$$

Using Theorem 2.1.2, p. 86 in [22], a sufficient condition for the existence of a density for (U^n, V^n) is that its Malliavin covariance matrix γ_{U^n,V^n} satisfies a nondegeneracy condition given, for instance, by the following lemma.

Lemma 3. *γ_{U^n,V^n} is an a.s. invertible matrix and for all $p \geqslant 1$, we have*

$$E\left(\left|\det(\gamma_{U^n,V^n})\right|^{-p}\right) \leqslant c(p).$$

Proof. To have shorter notations, during the proof we will denote by c_* any generic positive random variable which satisfies $E(c_*^{-p}) \leqslant c(p)$. By direct computations we have,

$$\langle U^n, U^n\rangle_H = \int_0^1 a_n^2\left(\mathcal{X}^n_t\right)\left(\mathcal{Y}^n_t\right)^{-2}\left(\int_{[t,1]} \mathcal{Y}^n_s \,\mathrm{d}\mu(s)\right)^2 \mathrm{d}t, \tag{42}$$

$$\langle U^n, V^n\rangle_H = \int_0^1 a_n^2\left(\mathcal{X}^n_t\right)\left(\mathcal{Y}^n_t\right)^{-2}\left(\int_{[t,1]} \mathcal{Y}^n_s \,\mathrm{d}\mu(s)\right) \mathrm{d}t\, \mathcal{Y}^n_1, \tag{43}$$

$$\langle V^n, V^n\rangle_H = \int_0^1 a_n^2\left(\mathcal{X}^n_t\right)\left(\mathcal{Y}^n_t\right)^{-2} \mathrm{d}t \left(\mathcal{Y}^n_1\right)^2. \tag{44}$$

Now, define the probability density on [0, 1]

$$m^n_t = a_n^2\left(\mathcal{X}^n_t\right)\left(\mathcal{Y}^n_t\right)^{-2}\left(\int_0^1 a_n^2\left(\mathcal{X}^n_s\right)\left(\mathcal{Y}^n_s\right)^{-2} \mathrm{d}s\right)^{-1}, \tag{45}$$

and set $f^n(t) := \int_{[t,1]} \mathcal{Y}^n_s \,\mathrm{d}\mu(s)$. Thus we can write:

$$\det(\gamma_{U^n,V^n}) = \langle V^n, V^n\rangle_H^2 \left(\mathcal{Y}^n_1\right)^{-2}\left[\int_0^1 m^n_t f^n(t)^2 \,\mathrm{d}t - \left(\int_0^1 m^n_s f^n(s)\,\mathrm{d}s\right)^2\right].$$

Hence, the above bracket can be interpreted as the variance of the function $f^n(t)$ under the probability measure $m^n_t \,\mathrm{d}t$ and hence:

$$\det(\gamma_{U^n,V^n}) = \langle V^n, V^n\rangle_H^2 \left(\mathcal{Y}^n_1\right)^{-2} \int_0^1 m^n_t\left[f^n(t) - \left(\int_0^1 m^n_r f^n(r)\,\mathrm{d}r\right)\right]^2 \mathrm{d}t.$$

But clearly under Assumption (R), $\langle V^n, V^n\rangle_H^2 \geqslant \underline{a}^2 \inf_{t\in[0,1]}(\mathcal{Y}_t^n)^{-2} \inf_{t\in[0,1]}(\mathcal{Y}_t^n)^2$ and hence by (41) this yields, $\langle V^n, V^n\rangle_H^2 \geqslant c_*$, using our convention about generic positive random variables c_*. Similarly, by (45), we have $m_t^n \geqslant c_*$ and thus,

$$\det(\gamma_{U^n,V^n}) \geqslant c_* \int_0^1 \left[f^n(t) - \left(\int_0^1 m_r^n f^n(r)\,\mathrm{d}r \right) \right]^2 \mathrm{d}t.$$

Then, writing the integral above as

$$\int_0^{1/2} \left[f^n(t) - \left(\int_0^1 m_r^n f^n(r)\,\mathrm{d}r \right) \right]^2 + \left[f^n\left(t+\frac{1}{2}\right) - \left(\int_0^1 m_r^n f^n(r)\,\mathrm{d}r \right) \right]^2 \mathrm{d}t,$$

and using the simple inequality $x^2+y^2 \geqslant (x-y)^2/2$, we get: $\det(\gamma_{U^n,V^n}) \geqslant c_* \int_0^{1/2} (\int_{[t,t+1/2)} \mathcal{Y}_s^n \,\mathrm{d}\mu(s))^2\,\mathrm{d}t$. Using again $\inf_{s\in[0,1]} \mathcal{Y}_s^n \geqslant c_*$, we obtain: $\det(\gamma_{U^n,V^n}) \geqslant c_* \int_0^{1/2} \mu([t,t+1/2))^2\,\mathrm{d}t$. But this integral is positive as soon as $\mu((0,1)) > 0$ which is the case by assumption (3). Thus the lemma is proved. □

A.1.2. Bounds for the density

For the proof of (9), we make a crucial use of the fact that the diffusion process $\mathcal{X}^n$ is one-dimensional by introducing the classical transformation:

$$s_n(x) := \int_0^x a_n^{-1}(y)\,\mathrm{d}y, \qquad \mathcal{W}_t^n := s_n\big(\mathcal{X}_t^n\big).$$

By the assumptions on a, the function s_n is one to one on $\mathbb{R}$ and the derivatives of s_n and s_n^{-1} are bounded independently of n. By Itô's formula, $\mathcal{W}^n$ solves the equation $\mathrm{d}\mathcal{W}_t^n = \mathrm{d}W_t + \tilde{b}_n(\mathcal{W}_t^n)\,\mathrm{d}t$ where $\tilde{b}_n(w) := \frac{b_n}{a_n} \circ s_n^{-1}(w) - \frac{1}{2}a_n' \circ s_n^{-1}(w)$ and the initial value is $\mathcal{W}_0^n = s_n(\mathcal{X}_0^n) = 0$. We let $\tilde{P}$ be the probability defined on $(\Omega, \mathcal{A})$ by

$$\frac{\mathrm{d}\tilde{P}}{\mathrm{d}P} = \exp\biggl(-\int_0^1 \tilde{b}_n\big(\mathcal{W}_u^n\big)\,\mathrm{d}W_u - \frac{1}{2}\int_0^1 \tilde{b}_n^2\big(\mathcal{W}_u^n\big)\,\mathrm{d}u \biggr).$$

The Girsanov theorem implies that the process $\mathcal{W}^n$ is under $\tilde{P}$ a standard Brownian motion. Note that the random variables (U^n, V^n) have the following expressions with respect to this $\tilde{P}$-Brownian motion:

$$U^n = \int_0^1 s_n^{-1}\big(\mathcal{W}_r^n\big)\,\mathrm{d}\mu(r), \tag{46}$$

$$V^n = s_n^{-1}\big(\mathcal{W}_1^n\big). \tag{47}$$

Now let h_0 and h_1 be nonnegative real functions, then:

$$E_P\big[h_0\big(U^n\big)h_1\big(V^n\big)\big] = E_{\tilde{P}}\big[h_0\big(U^n\big)h_1\big(V^n\big)L^n\big], \tag{48}$$

where $L^n = \exp(\int_0^1 \tilde{b}_n(\mathcal{W}_r^n)\,\mathrm{d}\mathcal{W}_r^n - \frac{1}{2}\int_0^1 \tilde{b}_n^2(\mathcal{W}_r^n)\,\mathrm{d}r)$. But using Itô's formula, $L^n = \exp(\tilde{B}_n(\mathcal{W}_1^n) - \frac{1}{2}\int_0^1 (\tilde{b}_n^2 + \tilde{b}_n')(\mathcal{W}_r^n)\,\mathrm{d}r)$ where $\tilde{B}_n$ is the primitive function of $\tilde{b}_n$ vanishing at zero. Since $\tilde{b}_n$ and $\tilde{b}_n'$ are clearly bounded by $cn^{-1/2}$ for some constant c only depending on a and b and $|\tilde{B}_n(x)| \leqslant cn^{-1/2}|x|$ we have: $c^{-1}\exp(-cn^{-1/2}|\mathcal{W}_1^n|) \leqslant L^n \leqslant c\exp(cn^{-1/2}|\mathcal{W}_1^n|)$. By (47) and the boundedness of s_n' we deduce $c^{-1}\exp(-cn^{-1/2}|V^n|) \leqslant L^n \leqslant c\exp(cn^{-1/2}|V^n|)$. From this and (48), we obtain:

$$c^{-1}E_{\tilde{P}}\big[h_0\big(U^n\big)h_1\big(V^n\big)\mathrm{e}^{-cn^{-1/2}|V^n|}\big] \leqslant E_P\big[h_0\big(U^n\big)h_1\big(V^n\big)\big] \leqslant cE_{\tilde{P}}\big[h_0\big(U^n\big)h_1\big(V^n\big)\mathrm{e}^{cn^{-1/2}|V^n|}\big].$$

Hence, we have transformed the problem of finding bounds for the density of the law of (U^n, V^n) under P into an analogous problem under $\tilde{P}$. Consequently, the bounds for $\mathfrak{p}_{x_0}^n$ stated in (9) will follow from the next lemma.

Lemma 4. *Let h_0, h_1 be two nonnegative functions. There exist some constants $c_1 > c_2 > 0$, depending only on the coefficients a and b such that*:

$$c_1^{-1} \iint h_0(u)h_1(v)\mathrm{e}^{-c_1(u^2+v^2)}\,\mathrm{d}u\,\mathrm{d}v \leqslant E_{\tilde{P}}\big[h_0\big(U^n\big)h_1\big(V^n\big)\big] \leqslant c_2^{-1} \iint h_0(u)h_1(v)\mathrm{e}^{-c_2(u^2+v^2)}\,\mathrm{d}u\,\mathrm{d}v.$$

Proof. We first show the lower bound. Using that the random variable V^n is measurable with respect to $\mathcal{W}_1^n$ (by (47)), we can write:

$$\begin{aligned} E_{\tilde{P}}\big[h_0\big(U^n\big)h_1\big(V^n\big)\big] &= E_{\tilde{P}}\big[h_1\big(V^n\big)E_{\tilde{P}}\big[h_0\big(U^n\big)\big|\mathcal{W}_1^n\big]\big] \\ &= \int \mathfrak{g}(w)h_1\big(s_n^{-1}(w)\big)E_{\tilde{P}}\big[h_0\big(U^n\big)\big|\mathcal{W}_1^n = w\big]\,\mathrm{d}w, \end{aligned} \tag{49}$$

where $\mathfrak{g}$ is the density of the standard Gaussian law. Now let us admit temporarily the following relation on the conditional law of U^n:

$$E_{\tilde{P}}\big[h_0\big(U^n\big)\big|\mathcal{W}_1^n\big] \geqslant c^{-1}\mathrm{e}^{-c(\mathcal{W}_1^n)^2} \int h_0(u)\mathrm{e}^{-cu^2}\,\mathrm{d}u. \tag{50}$$

Then $E_{\tilde{P}}[h_0(U^n)h_1(V^n)]$ is greater than:

$$c^{-1} \int h_0(u)\mathrm{e}^{-cu^2}\,\mathrm{d}u \times \int \mathfrak{g}(w)h_1\big(s_n^{-1}(w)\big)\mathrm{e}^{-cw^2}\,\mathrm{d}w.$$

The change of variable $v = s_n^{-1}(w)$ in the second integral above, the inequalities $|w| \leqslant c|v|$ and $s_n'(v) \geqslant c$ give the new lower bound

$$c^{-1} \int h_0(u)\mathrm{e}^{-cu^2}\,\mathrm{d}u \times \int \mathfrak{g}\big(s_n(v)\big)h_1(v)\mathrm{e}^{-cv^2}\,\mathrm{d}v,$$

with a new constant c. Since $\mathfrak{g}$ is the Gaussian kernel and thanks to the inequality $|s_n(v)| \leqslant c|v|$, we deduce the required lower bound for $E_{\tilde{P}}[h_0(U^n)h_1(V^n)]$.

We obtain the upper bound quite similarly. Let us temporarily admit that for all ε small enough there exists $c(\varepsilon)$ such that:

$$E_{\tilde{P}}\big[h_0\big(U^n\big)\big|\mathcal{W}_1^n\big] \leqslant c(\varepsilon)^{-1}\mathrm{e}^{\varepsilon(\mathcal{W}_1^n)^2} \int h_0(u)\mathrm{e}^{-c(\varepsilon)u^2}\,\mathrm{d}u. \tag{51}$$

Plugging this in Eq. (49), we deduce that $E_{\tilde{P}}[h_0(U^n)|\mathcal{W}_1^n]$ is smaller than

$$c(\varepsilon)^{-1} \int h_0(u)\mathrm{e}^{-c(\varepsilon)u^2}\,\mathrm{d}u \times \int \mathfrak{g}(w)\mathrm{e}^{\varepsilon w^2}h_1\big(s_n^{-1}(w)\big)\,\mathrm{d}w.$$

Since $\mathfrak{g}(w) = \exp(-w^2/2)/\sqrt{2\pi}$, any choice of ε smaller than $1/4$ implies that the second integral in the equation above is bounded by $c \int \mathrm{e}^{-(1/4)w^2}h_1(s_n^{-1}(w))\,\mathrm{d}w$. As for the lower bound, we conclude by the change of variable $v = s_n^{-1}(w)$. □

It remains to show (50) and (51). This is done in the following lemma.

Lemma 5. *For some constant $c > 0$ and $\overline{\varepsilon} > 0$, we have*

$$E_{\tilde{P}}\big[h_0\big(U^n\big)\big|\mathcal{W}_1^n\big] \geqslant c^{-1}\mathrm{e}^{-c(\mathcal{W}_1^n)^2} \int h_0(u)\mathrm{e}^{-cu^2}\,\mathrm{d}u. \tag{52}$$

For all $\varepsilon \in)0, \overline{\varepsilon}($, there exists $c(\varepsilon) > 0$ such that,

$$E_{\tilde{P}}\big[h_0\big(U^n\big)\big|\mathcal{W}_1^n\big] \leqslant c(\varepsilon)^{-1}\mathrm{e}^{\varepsilon(\mathcal{W}_1^n)^2} \int h_0(u)\mathrm{e}^{-c(\varepsilon)u^2}\,\mathrm{d}u. \tag{53}$$

Proof. Let us recall that the process $\mathcal{W}^*_t := \mathcal{W}^n_t - t\mathcal{W}^n_1$ is a Brownian bridge on [0, 1], independent of the variable $\mathcal{W}^n_1$. Thus, we can evaluate the conditional expectation $E_{\tilde{P}}[h_0(U^n)|\mathcal{W}^n_1 = w]$ as the expectation (recall (46)),

$$E\left[h_0\left(\int_0^1 s_n^{-1}(\mathcal{W}^*_t + tw)\,\mathrm{d}\mu(t)\right)\right], \tag{54}$$

for $\mathcal{W}^*$ some Brownian bridge. This Brownian bridge itself admits a decomposition

$$\mathcal{W}^*_t = \xi\eta_t + \mathcal{W}^{**}_t, \tag{55}$$

where ξ is a $\mathcal{N}(0, 1)$ variable, η is the deterministic triangle shaped function:

$$\eta_t = \begin{cases} t & \text{if } t \in [0, 1/2], \\ (1-t) & \text{if } t \in [1/2, 1], \end{cases}$$

and $\mathcal{W}^{**}$ is the process on [0, 1] constructed as the concatenation of two independent Brownian bridges, one on [0, 1/2] and another on [1/2, 1]. Furthermore in this decomposition the r.v. η and the process $\mathcal{W}^{**}$ are independent.

For any realization of $\mathcal{W}^{**}$ we can introduce the real function,

$$x \mapsto g_{\mathcal{W}^{**}}(x) = \int_0^1 \left\{s_n^{-1}\left(x\eta_t + \mathcal{W}^{**}_t + tw\right)\right\}\mathrm{d}\mu(t).$$

Using (55) and the independence of ξ and $\mathcal{W}^{**}$, the quantity (54) now writes,

$$E_{(\mathcal{W}^{**})}E_{(\xi)}\left[h_0\left(g_{\mathcal{W}^{**}}(\xi)\right)\right], \tag{56}$$

where the inner expectation denotes the expectation with respect to the random variable ξ and the outer one with respect to the process $\mathcal{W}^{**}$.

First we evaluate the inner expectation. Using that ξ is a standard Gaussian variable we have

$$E_{(\xi)}\left[h_0\left(g_{\mathcal{W}^{**}}(\xi)\right)\right] = (2\pi)^{-1/2}\int h_0\left(g_{\mathcal{W}^{**}}(x)\right)\mathrm{e}^{-x^2/2}\,\mathrm{d}x. \tag{57}$$

Note now that for any realization of $\mathcal{W}^{**}$, the function $x \mapsto g_{\mathcal{W}^{**}}(x)$ is differentiable and using that $\frac{1}{c} \leqslant (s_n^{-1})' \leqslant c$ we get

$$\frac{1}{c}\int_0^1 \eta_t\,\mathrm{d}\mu(t) \leqslant g'_{\mathcal{W}^{**}}(x) \leqslant c\int_0^1 \eta_t\,\mathrm{d}\mu(t).$$

By assumption (3) on the measure μ the integral $\int_0^1 \eta_t\,\mathrm{d}\mu(t)$ is positive. Thus, the function $x \mapsto g_{\mathcal{W}^{**}}(x)$ is invertible on $\mathbb{R}$, with a derivative bounded from above and from below by some constant independent of $\mathcal{W}^{**}$ and n. This allows us to make a change of variable in (57) to obtain the bounds

$$c^{-1}\int h_0(u)\mathrm{e}^{-(g^{-1}_{\mathcal{W}^{**}}(u))^2/2}\,\mathrm{d}u \leqslant E_{(\xi)}\left[h_0\left(g_{\mathcal{W}^{**}}(\xi)\right)\right] \leqslant c\int h_0(u)\mathrm{e}^{-(g^{-1}_{\mathcal{W}^{**}}(u))^2/2}\,\mathrm{d}u. \tag{58}$$

Now the proofs of (52) and (53) are treated separately.

- For the lower bound, we have seen that $g^{-1}_{\mathcal{W}^{**}}$ is globally Lipschitz with a constant independent of $\mathcal{W}^{**}$ and thus $|g^{-1}_{\mathcal{W}^{**}}(u)| \leqslant c|u| + |g^{-1}_{\mathcal{W}^{**}}(0)| \leqslant c|u| + c|g_{\mathcal{W}^{**}}(0)|$. In addition, a simple computation from the definition of $g_{\mathcal{W}^{**}}$ and then boundedness of $(s_n^{-1})'$ show that

$$\left|g_{\mathcal{W}^{**}}(0)\right| \leqslant c\left[|w| + \sup_{t\in[0,1]}\left|\mathcal{W}^{**}_t\right|\right]. \tag{59}$$

Using this in (58) we find a new lower bound for the inner expectation:

$$E_{(\xi)}\big[h_0\big(g_{\mathcal{W}^{**}}(\xi)\big)\big] \geqslant c^{-1}\mathrm{e}^{-cw^2}\mathrm{e}^{-c\sup_{t\in[0,1]}(W_t^{**})^2}\int h_0(u)\mathrm{e}^{-cu^2}\,\mathrm{d}u.$$

Taking the expectation with respect to $\mathcal{W}^{**}$ proves that (56) is larger than $c^{-1}E_{(\mathcal{W}^{**})}(\mathrm{e}^{-c\sup_{t\in[0,1]}(W_t^{**})^2})\times \mathrm{e}^{-cw^2}\int h_0(u)\mathrm{e}^{-cu^2}\,\mathrm{d}u$. This gives (52).

- For the upper bound, we write using that $g_{\mathcal{W}^{**}}$ is Lipschitz $||u|-|g_{\mathcal{W}^{**}}(0)||\leqslant |g_{\mathcal{W}^{**}}(g^{-1}_{\mathcal{W}^{**}}(u))-g_{\mathcal{W}^{**}}(0)|\leqslant c|g^{-1}_{\mathcal{W}^{**}}(u)|$. Together with the relation $(x-y)^2\geqslant x^2\frac{\varepsilon}{1+\varepsilon}-\varepsilon y^2$ (for $x,y\in\mathbb{R}$, $\varepsilon\in(0,1)$), we deduce that $\exp(-\frac{1}{2}(g^{-1}_{\mathcal{W}^{**}}(u))^2)$ is upper bounded by

$$\exp\biggl(-\frac{\varepsilon u^2}{2c^2(1+\varepsilon)}+\frac{\varepsilon(g_{\mathcal{W}^{**}}(0))^2}{2c^2}\biggr)\leqslant \exp\biggl(-\frac{\varepsilon u^2}{2c^2(1+\varepsilon)}+\varepsilon w^2+\varepsilon\sup_{t\in[0,1]}\big(\mathcal{W}_t^{**}\big)^2\biggr),$$

where we have used (59). Combining this with (58) and taking the expectation with respect to $\mathcal{W}^{**}$, we get that the quantity (56) is smaller than: $c\int h_0(u)\mathrm{e}^{-\varepsilon u^2/(2c(1+\varepsilon))}\mathrm{e}^{\varepsilon w^2}E_{\mathcal{W}^{**}}(\mathrm{e}^{\varepsilon\sup_{t\in[0,1]}(\mathcal{W}_t^{**})^2})$. The last expectation is finite as soon as ε is small enough, and thus (53) holds. □

A.1.3. Proof of Proposition 1

To have shorter notations we set $\overline{\mathcal{X}}^n_j=\int_0^1\mathcal{X}^n_{j+t}\,\mathrm{d}\mu(t)$, for $j\geqslant 0$. First we prove the following lemma.

Lemma 6. *Let us define Γ^k the Malliavin covariance matrix of the vector $((\overline{\mathcal{X}}^n_j,\mathcal{X}^n_{j+1}))_{j=0,\dots,k-1}$ of size $2k$. Then this matrix is a.s invertible and $E(\det(\Gamma^k)^{-p})\leqslant c(p,k)$.*

Proof. In the case $k=1$, the lemma reduces to Lemma 3. For $k\geqslant 2$, we proceed by induction by establishing simple relations between the columns of Γ^k (this simplification follows from the flow property of the process $\mathcal{X}^n$).

To see this, notice that firstly by (39) if $t<k-1$ and $s>k-1$, we have $D_t\mathcal{X}^n_s=D_t\mathcal{X}^n_{k-1}\mathcal{Y}^n_s(\mathcal{Y}^n_{k-1})^{-1}$; and secondly if $t>k-1$ and $s<k-1$, $D_t\mathcal{X}^n_s=0$. Using these two properties, a calculation shows that if $(C_j)_{j=1,\dots,2k}$ denote the columns of Γ^k, we have the relation,

$$[C_{2k-1};C_{2k}]=\biggl[\biggl(\int_0^1\mathcal{Y}^n_{k-1+s}\big(\mathcal{Y}^n_{k-1}\big)^{-1}\,\mathrm{d}\mu(s)\biggr)C_{2k-2};\mathcal{Y}^n_k\big(\mathcal{Y}^n_{k-1}\big)^{-1}C_{2k-2}\biggr]+\begin{pmatrix}0&0\\ \vdots&\vdots\\ 0&0\\ \multicolumn{2}{c}{[\gamma_k]}\end{pmatrix},$$

where γ_k is the matrix of size 2×2 given by

$$\begin{pmatrix}\int_{k-1}^k(D_t\overline{\mathcal{X}}^n_{k-1})^2\,\mathrm{d}t & \int_{k-1}^k(D_t\mathcal{X}^n_k)(D_t\overline{\mathcal{X}}^n_{k-1})\,\mathrm{d}t\\ \int_{k-1}^k(D_t\mathcal{X}^n_k)(D_t\overline{\mathcal{X}}^n_{k-1})\,\mathrm{d}t & \int_{k-1}^k(D_t\mathcal{X}^n_k)^2\,\mathrm{d}t\end{pmatrix}.$$

This proves that $\det\Gamma^k=\det\Gamma^{k-1}\det\gamma_k$. But it can be seen that the matrix γ_k has an expression similar to γ_{U^n,V^n} (but with integration interval shifted from $[0,1]$ to $[k-1,k]$) from which we can prove $E((\det\gamma_k)^{-p})\leqslant c(p)$.

The lemma then follows from induction on k. □

Now we can deduce the Proposition 1. Recalling (11)–(13) we can find an invertible matrix M of size $2k\times 2k$ that maps $((\overline{\mathcal{X}}^n_j,\mathcal{X}^n_{j+1}))_{j=0,\dots,k-1}$ into a vector whose $k+1$ first components are exactly $(U^n_0,\dots,U^n_k)$. Denoting $\hat{\Gamma}^k$ the Malliavin covariance matrix of the image by M of $((\overline{\mathcal{X}}^n_j,\mathcal{X}^n_{j+1}))_{j=0,\dots,k-1}$, we have $\hat{\Gamma}^k=M\Gamma^kM^\star$. Thus, Lemma 6 yields $E((\det(\hat{\Gamma}^k))^{-p})\leqslant c(p,k)$ since M is invertible. Observing that the Malliavin covariance matrix $K(\theta)$ is the matrix extracted from the $k+1$ first rows and columns of $\hat{\Gamma}^k$ we deduce Proposition 1.

A.2. Some estimates on the change of measures

For this section we denote by $\mathcal{X}$ the canonical process on $\mathcal{C}([0,\infty))$ and we consider the random variable on this space defined by $H = f(U_0,\dots,U_k)$, where $(U_0,\dots,U_k)$ is given by (10)–(13) with the canonical process $\mathcal{X}$ in place of $\mathcal{X}^{\theta,n}$; we denote by E^n_{θ,x_0} the expectation with respect to the measure induced on the canonical space by the law of $\mathcal{X}^{\theta,n}$ solution of (6).

Lemma 7. *There exist $r \geqslant 1$ and a constant $c(k) \geqslant 0$, such that $\forall H = f(U_0,\dots,U_k) \geqslant 0$, $\forall \theta,\theta' \in \Theta$, $\forall x_0 \in \mathbb{R}$, we have*

$$E^n_{\theta',x_0}[H] \leqslant c(k) E^n_{\theta,x_0}\big[H^r\big]^{1/r}.$$

Proof. Recalling the notation of Section 2.1.1 we denote $\mathfrak{p}^n_{x_0}(u,v,\theta)$ the density of the vector (8) and for $j = 0,\dots,k-1$ we let

$$Z_{j,\theta,\theta'} = \frac{\mathfrak{p}^n_{\mathcal{X}_j}(\int_0^1(\mathcal{X}_{j+s}-\mathcal{X}_j)\,\mathrm{d}\mu(s), (\mathcal{X}_{j+1}-\mathcal{X}_j), \theta')}{\mathfrak{p}^n_{\mathcal{X}_j}(\int_0^1(\mathcal{X}_{j+s}-\mathcal{X}_j)\,\mathrm{d}\mu(s), (\mathcal{X}_{j+1}-\mathcal{X}_j), \theta)}. \tag{60}$$

Then using the Markov property of the process $\mathcal{X}$ under the laws P^n_θ and $P^n_{\theta'}$, we have

$$E^n_{\theta',x_0}[H] = E^n_{\theta,x_0}\left[H\prod_{j=0}^{k-1} Z_{j,\theta,\theta'}\right] \leqslant E^n_{\theta,x_0}\big[H^r\big]^{1/r} E^n_{\theta,x_0}\left[\prod_{j=0}^{k-1}(Z_{j,\theta,\theta'})^{r'}\right]^{1/r'},$$

where r and r' are conjugate exponents. But we know by Theorem 4 that there exist two constants $0 < c_2 \leqslant c_1$ (uniform w.r.t. θ, x_0, n) such that

$$c_1^{-1}\mathrm{e}^{-c_1(u^2+v^2)} \leqslant \mathfrak{p}^n_{x_0}(u,v,\theta) \leqslant c_2^{-1}\mathrm{e}^{-c_2(u^2+v^2)}.$$

Then one can bound the conditional expectation $E^n_{\theta,x_0}[(Z_{k-1,\theta,\theta'})^{r'}|\mathcal{X}_s, s \leqslant (k-1)]$ by

$$\frac{c_1^{(r'-1)}}{c_2^{r'}}\int_{\mathbb{R}^2}\mathrm{e}^{(u^2+v^2)(-r'c_2+(r'-1)c_1)}\,\mathrm{d}u\,\mathrm{d}v.$$

But if r is chosen large enough such that r' is sufficiently close to 1 the latter integral converges and is equal to some constant κ. Proceeding by induction we get:

$$E^n_{\theta,x_0}\left[\prod_{j=0}^{k-1}(Z_{j,\theta,\theta'})^{r'}\right]^{1/r'} \leqslant E^n_{\theta,x_0}\left[\prod_{j=0}^{k-2}(Z_{j,\theta,\theta'})^{r'}\right]^{1/r'}\kappa^{1/r'} \leqslant \cdots \leqslant \kappa^{k/r'}$$

which gives the result. □

Lemma 8. *There exist $c(k) \geqslant 0$ and $\alpha \geqslant 1$ such that $\forall H = f(U_0,\dots,U_k)$ (with $E^n_{\theta,x_0}|H|^\alpha < +\infty$), $\forall \theta,\theta' \in \Theta$, $\forall x_0 \in \mathbb{R}$, we have*

$$\big|E^n_{\theta',x_0}[H] - E^n_{\theta,x_0}[H]\big| \leqslant c(k)\big|\theta-\theta'\big|\big[E^n_{\theta,x_0}|H|^\alpha\big]^{1/\alpha}. \tag{61}$$

Proof. Using the notations of Lemma 7, we write

$$E^n_{\theta',x_0}[H] - E^n_{\theta,x_0}[H] = E^n_{\theta,x_0}\left[\left(\prod_{j=0}^{k-1}Z_{j,\theta,\theta'} - 1\right)H\right] = \sum_{i=0}^{k-1}E^n_{\theta,x_0}\left[(Z_{i,\theta,\theta'}-1)\prod_{j=i+1}^{k-1}Z_{j,\theta,\theta'}H\right].$$

Thus for conjugate exponents α and β, the left-hand side of (61) is bounded by

$$\sum_{i=0}^{k-1}\big[E^n_{\theta,x_0}|H|^\alpha\big]^{1/\alpha}E^n_{\theta,x_0}\Bigg[|Z_{i,\theta,\theta'}-1|^\beta\prod_{j=i+1}^{k-1}(Z_{j,\theta,\theta'})^\beta\Bigg]^{1/\beta}$$

$$=\sum_{i=0}^{k-1}\big[E^n_{\theta,x_0}|H|^\alpha\big]^{1/\alpha}E^n_{\theta,x_0}\Bigg[|Z_{i,\theta,\theta'}-1|^\beta E^n_{\theta,x_0}\Bigg[\prod_{j=i+1}^{k-1}(Z_{j,\theta,\theta'})^\beta\Big|\mathcal{X}_s, s\leqslant i+1\Bigg]\Bigg]^{1/\beta}.$$

Using the Markov property of $\mathcal{X}$ it can be shown exactly as in Lemma 7 that the conditional expectation in the equation above is finite, as soon as β is small enough and bounded by κ^{k-i-1}. Thus by Lemma 9 below, we deduce $|E_{\theta',x_0}[H]-E_{\theta,x_0}[H]|\leqslant c(\beta)|\theta-\theta'|[E^n_{\theta,x_0}|H|^\alpha]^{1/\alpha}\sum_{i=0}^{k-1}\kappa^{(k-i-1)/\beta}$. □

Lemma 9. *There exists $\overline{\beta}>1$ such that for all $1<\beta\leqslant\overline{\beta}$ we have*:

$$E^n_{\theta,x_0}\big[|Z_{i,\theta,\theta'}-1|^\beta\big]^{1/\beta}\leqslant c(\beta)\big|\theta-\theta'\big|.$$

Proof. Using the expression of $Z_{i,\theta,\theta'}$, and Eq. (8) with the Markov property, it suffices to bound the quantity

$$E^n_{x_0,\theta}\Bigg[\bigg|\frac{\mathfrak{p}^n_{x_0}(U^n,V^n,\theta')-\mathfrak{p}^n_{x_0}(U^n,V^n,\theta)}{\mathfrak{p}^n_{x_0}(U^n,V^n,\theta)}\bigg|^\beta\Bigg]^{1/\beta}. \tag{62}$$

By Theorem 5 with $k=1$ the function $\theta\to\mathfrak{p}_{x_0}(U^n,V^n,\theta)$ is absolutely continuous and we can write the quantity above as: $E^n_{x_0,\theta}[|(\int_\theta^{\theta'}\dot{\mathfrak{p}}^n_{x_0}(U^n,V^n,s)\,\mathrm{d}s)/(\mathfrak{p}^n_{x_0}(U^n,V^n,\theta))|^\beta]^{1/\beta}$. Using first the Minkowski inequality, a change of measure and then the Hölder inequality one finds the following bounds for this quantity:

$$\int_\theta^{\theta'}E^n_{x_0,s}\Bigg[\bigg|\frac{\dot{\mathfrak{p}}^n_{x_0}(U^n,V^n,s)}{\mathfrak{p}^n_{x_0}(U^n,V^n,\theta)}\bigg|^\beta\frac{\mathfrak{p}^n_{x_0}(U^n,V^n,\theta)}{\mathfrak{p}^n_{x_0}(U^n,V^n,s)}\Bigg]^{1/\beta}\mathrm{d}s$$

$$\leqslant\int_\theta^{\theta'}E_{x_0,s}\Bigg[\bigg|\frac{\dot{\mathfrak{p}}^n_{x_0}(U^n,V^n,s)}{\mathfrak{p}^n_{x_0}(U^n,V^n,s)}\bigg|^{\beta\alpha'}\Bigg]^{1/(\beta\alpha')}E_{x_0,s}\Bigg[\bigg|\frac{\mathfrak{p}^n_{x_0}(U^n,V^n,s)}{\mathfrak{p}^n_{x_0}(U^n,V^n,\theta)}\bigg|^{(\beta-1)\beta'}\Bigg]^{1/(\beta\beta')}\mathrm{d}s$$

for two conjugate exponents α' and β'. But the first expectation in the right-hand side above is bounded by Corollary 2 (with $k=1$) for all choices of α', β. The second expectation can be bounded if $(\beta-1)\beta'$ is close enough to zero by using (9) as in the proof of Lemma 7. This gives that (62) is smaller than $c|\theta-\theta'|$. □

A.3. A technical lemma

Lemma 10. *Let $(G_0,\dots,G_k)$ be a centered Gaussian vector with invertible covariance matrix C_{k+1} and let us denote by C_{k-1} the covariance matrix of $(G_1,\dots,G_{k-1})$. Then,*

$$\sum_{0\leqslant j,j'\leqslant k}G_j[C_{k+1}]^{-1}_{j,j'}G_{j'}-\sum_{1\leqslant j,j'\leqslant k-1}G_j[C_{k-1}]^{-1}_{j,j'}G_{j'}, \tag{63}$$

is a $\chi^2(2)$ random variable.

Proof. Write the Gram–Schmidt orthonormalization procedure for the $\mathbf{L}^2$ vectors $G_1,\dots,G_k,G_0$ as:

$$\begin{bmatrix}\mathcal{H}_0\\ \vdots\\ \mathcal{H}_k\end{bmatrix}=P_k\begin{bmatrix}G_0\\ \vdots\\ G_k\end{bmatrix},$$

where the variables $\mathcal{H}_0,\dots,\mathcal{H}_k$ are i.i.d. with standard Gaussian law and P_k is some triangular matrix. Then a few linear algebra shows that (63) is equal to $\sum_{j=0}^k\mathcal{H}_j^2-\sum_{j=1}^{k-1}\mathcal{H}_j^2=\mathcal{H}_0^2+\mathcal{H}_k^2$ and thus is chi-square distributed. □

References

[1] O. E. Barndorff-Nielsen and N. Shephard. Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **63** (2001) 167–241. MR1841412

[2] P. D. Ditlevsen, S. Ditlevsen and K. K. Andersen. The fast climate fluctuations during the stadial and interstadial climate states. *Ann. Glaciology* **35** (2002) 457–462.

[3] S. Ditlevsen and M. Sørensen. Inference for observations of integrated diffusion processes. *Scand. J. Statist.* **31** (2004) 417–429. MR2087834

[4] G. Dohnal. On estimating the diffusion coefficient. *J. Appl. Probab.* **24** (1987) 105–114. MR0876173

[5] V. Genon-Catalot and J. Jacod. On the estimation of the diffusion coefficient for multi-dimensional processes. *Ann. Inst. H. Poincaré Probab. Statist.* **29** (1993) 119–151. MR1204521

[6] V. Genon-Catalot, T. Jeantheau and C. Laredo. Parameter estimation for discretely observed stochastic volatility models. *Bernoulli* **5** (1999) 855–872. MR1715442

[7] A. Gloter. Discrete sampling of an integrated diffusion process and parameter estimation of the diffusion coefficient. *ESAIM Probab. Statist.* **4** (2000) 205–227. MR1808927

[8] A. Gloter. Parameter estimation for a discrete sampling of an integrated Ornstein–Uhlenbeck process. *Statistics* **35** (2001) 225–243. MR1925514

[9] A. Gloter and J. Jacod. Diffusions with measurement errors. I. Local asymptotic normality. *ESAIM Probab. Statist.* **5** (2001) 225–242. MR1875672

[10] E. Gobet. Local asymptotic mixed normality property for elliptic diffusion: a Malliavin calculus approach. *Bernoulli* **7** (2001) 899–912. MR1873834

[11] E. Gobet. LAN property for ergodic diffusions with discrete observations. *Ann. Inst. H. Poincaré Probab. Statist.* **38** (2002) 711–737. MR1931584

[12] F. Hirsch and S. Song. Criteria of positivity for the density of a Wiener functional. *Bull. Sci. Math.* **121** (1997) 261–273. MR1456283

[13] F. Hirsch and S. Song. Properties of the set of positivity for the density of a regular Wiener functional. *Bull. Sci. Math.* **122** (1998) 1–15. MR1617582

[14] I. A. Ibragimov and R. Z. Has'minskii. *Statistical Estimation. Asymptotic Theory.* Springer-Verlag, New York–Berlin, 1981. (Translated from the Russian by Samuel Kotz.) MR0620321

[15] J. Jacod. On continuous conditional Gaussian martingales and stable convergence in law. In *Séminaire de probabilité XXXI*, 232–246. *Lecture Notes in Math.* **1655**. Springer, Berlin, 1997. MR1478732

[16] P. Jeganathan. On the asymptotic theory of estimation when the limit of the log-likelihood ratios is mixed normal. *Sankhyā Ser. A* **44** (1982) 173–212. MR0688800

[17] P. Jeganathan. Some asymptotic properties of risk functions when the limit of the experiment is mixed normal. *Sankhyā Ser. A* **45** (1983) 66–87. MR0749355

[18] A. Kohatsu-Higa. Lower bounds for densities of uniformly elliptic random variables on Wiener space. *Probab. Theory Related Fields* **126** (2003) 421–457. MR1992500

[19] P. Krée and C. Soize. *Mathematics of Random Phenomena. Random Vibrations of Mechanical Structures*. D. Reidel Publishing Co., Reidel, Dordrecht, 1986. MR0873731

[20] H. Kunita. Stochastic differential equations and stochastic flows of diffeomorphisms. *École d'été de probabilités de Saint-Flour, XII – 1982*, 143–303. *Lecture Notes in Math.* **1097**. Springer, Berlin, 1984. MR0876080

[21] L. Le Cam and G. Lo Yang. *Asymptotics in Statistics. Some Basic Concepts*, 2nd edition. Springer-Verlag, New York, 2000. MR1784901

[22] D. Nualart. *The Malliavin Calculus and Related Topics. Probability and Its Application*. Springer-Verlag, New-York, 1995. MR1344217

[23] D. Nualart. Analysis on Wiener space and anticipating stochastic calculus. *Lectures on Probability Theory and Statistics (Saint-Flour, 1995)*, 123–227. *Lecture Notes in Math.* **1690**. Springer, Berlin, 1998. MR1668111

[24] B. L. S. Prakasa Rao. *Statistical Inference for Diffusion Type Processes.* Edward Arnold, London; Oxford University Press, New York, 1999. MR1717690

[25] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*, 3rd edition. Springer-Verlag, Berlin, 1999. MR1725357