LAN property for stochastic differential equations with additive fractional noise and continuous time observation

Eulalia Nualart, Samy Tindel

LAN PROPERTY FOR STOCHASTIC DIFFERENTIAL EQUATIONS WITH ADDITIVE FRACTIONAL NOISE AND CONTINUOUS TIME OBSERVATION

EULALIA NUALART AND SAMY TINDEL

Abstract. We consider a stochastic differential equation with additive fractional noise of Hurst parameter $H > 1/2$, and a non-linear drift depending on an unknown parameter. We show the Local Asymptotic Normality property (LAN) of this parametric model with rate $\sqrt{\tau}$ as $\tau \to \infty$, when the solution is observed continuously on the time interval $[0, \tau]$. The proof uses ergodic properties of the equation and a Poincaré type inequality.

1. Introduction

Let B be a d-dimensional fractional Brownian motion with Hurst parameter $H > 1/2$. Let us recall that B is a centered Gaussian process defined on a complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$. The law of B is thus characterized by its covariance function, which is defined by

$$R_{s;t} \equiv \mathbf{E}\big[B_t^i \, B_s^j\big] = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right) \mathbf{1}_{\{0\}}(i-j), \qquad s, t \in \mathbb{R}. \tag{1}$$

The variance of the increments of B is then given by

$$\mathbf{E}\left[|B_t^i - B_s^i|^2\right] = |t-s|^{2H}, \qquad s, t \in \mathbb{R}, \quad i = 1, \ldots, d, \tag{2}$$

and this implies that almost surely the fBm paths are γ-Hölder continuous for any $\gamma < H$.

In this article, we consider the following $\mathbb{R}^d$-valued stochastic differential equation driven by B:

$$Y_t = y_0 + \int_0^t b(Y_s; \theta)\, ds + \sum_{j=1}^d \sigma_j B_t^j, \qquad t \in [0, \tau]. \tag{3}$$

Here $y_0 \in \mathbb{R}^d$ is a given initial condition, $B = (B^1, \ldots, B^d)$ is the aformentioned fractional Brownian motion (fBm), the unknown parameter θ lies in a certain set Θ which will be specified later on, $\{b(\cdot; \theta),\, \theta \in \Theta\}$ is a known family of drift coefficients with $b(\cdot; \theta) : \mathbb{R}^d \to \mathbb{R}^d$, and $\sigma_1, \ldots, \sigma_d \in \mathbb{R}^d$ are assumed to be known diffusion coefficients.

There has been a wide interest in drift estimation for stochastic equations driven by fractional Brownian motion in the recent past, partly motivated by inverse problems in a biomedical context [20]. However, notice that the existing literature on the topic mainly

Date: December 10, 2015.

2010 *Mathematics Subject Classification.* 62M09; 62F12.

Key words and phrases. fractional Brownian motion, parameter estimation, ergodicity.

First author acknowledges support from the European Union programme FP7-PEOPLE-2012-CIG under grant agreement 333938. Second author is member of the BIGS (Biology, Genetics and Statistics) team at INRIA.

focuses on the case where the dependence $\theta \mapsto b(x;\theta)$ is linear and all coefficients are real-valued. In this situation least squares and maximum likelihood estimators (MLE) for the unknown parameter θ can be computed explicitly, and numerous results are available: the case of continuous observations of Y and where the drift coefficient is linear in both θ and x (that is, fractional Ornstein-Uhlenbeck type process) has been studied e.g. in [2, 16, 18, 23], either in the ergodic ($\theta < 0$) and non-ergodic ($\theta > 0$) case. See also the monograph [30] and the references therein. Results on parameter estimation based on discrete observations of Y in the linear case can be found e.g. in [4, 35]. The case of a dependence of the form $(\theta, x) \mapsto \theta\, b(x)$ is handled e.g. in [21, 33].

However, the case of a general multi-dimensional coefficient $b(\theta, x)$ in equation (3) is also a very natural situation to consider, though we are only aware of the articles [6, 27] giving some positive answers for such a dependence. The current contribution has thus to be thought of as a step in this direction. Indeed, our main aim is to show that the model given by equation (3) satisfies the Local Asymptotic Normality property (LAN) with rate $\sqrt{\tau}$ as $\tau \to \infty$, when the process Y is observed continuously on the time interval $[0, \tau]$.

Before we proceed to a specific statement of our results, let us describe the assumptions we shall work with, which are similar to the ones used in [27]. We start with a standard hypothesis on the parameter set Θ:

Hypothesis 1.1. *The set Θ is compactly embedded in $\mathbb{R}^q$ for a given $q \geq 1$.*

In order to describe the assumptions on our coefficients b, we will use the following notation for partial derivatives:

Notation 1.2. *Let $f : \mathbb{R}^d \times \Theta \to \mathbb{R}$ be a $\mathcal{C}^{p_1,p_2}$ function for $p_1, p_2 \geq 0$. Then for any tuple $(i_1, \dots i_p) \in \{1, \dots, d\}^p$, we set $\partial_x^{i_1 \dots i_p} f$ for $\frac{\partial^p f}{\partial x_{i_1} \dots \partial x_{i_p}}$. Analogously, we use the notation $\partial_\theta^{i_1 \dots i_p} f$ for $\frac{\partial^p f}{\partial \theta_{i_1} \dots \partial \theta_{i_p}}$, where $(i_1, \dots i_p)$ is a tuple in $\{1, \dots, q\}^p$. Moreover, we will write $\partial_x f$ resp. $\partial_\theta f$ for the Jacobi-matrices $(\partial_{x_1} f, \dots, \partial_{x_d} f)$ and $(\partial_{\theta_1} f, \dots, \partial_{\theta_q} f)$. Finally, for the sake of simplicity, we denote by $\langle \cdot, \cdot \rangle$ the Euclidean scalar product in $\mathbb{R}^q$ or $\mathbb{R}^d$, and by $|\cdot|$ the corresponding Euclidean norm.*

Let us now state the linear growth plus inward assumptions on our drift coefficient ensuring ergodic properties for the process Y:

Hypothesis 1.3. (i) *There exists $\alpha > 0$ such that for every $x, y \in \mathbb{R}^d$ and $\theta \in \Theta$ we have*

$$\langle b(x;\theta) - b(y;\theta),\, x - y \rangle \leq -\alpha\, |x - y|^2.$$

(ii) *We have $b \in \mathcal{C}^{2,1}(\mathbb{R}^d \times \Theta; \mathbb{R}^d)$, with $\partial_x b$ and $\partial_{xx}^2 b$ uniformly bounded in (x, θ).*

(iii) *We have $\hat{b} := \partial_\theta b \in \mathcal{C}^{2,0}(\mathbb{R}^d \times \Theta; \mathbb{R}^d \times \mathbb{R}^q)$, with $\partial_x \hat{b}$ and $\partial_{xx}^2 \hat{b}$ uniformly bounded in (x, θ).*

(iv) *There exists $c > 0$ such that for every $x, y \in \mathbb{R}^d$ and $\theta, \theta_0 \in \Theta$, the following Lipschitz and growth conditions hold:*

$$|b(x;\theta)| \leq c\,(1 + |x|)\,, \qquad |\hat{b}(x;\theta)| \leq c\,(1 + |x|)$$
$$|\hat{b}(x;\theta) - \hat{b}(x;\theta_0)| \leq c|\theta - \theta_0|(1 + |x|), \qquad |\hat{b}(x;\theta) - \hat{b}(y;\theta)| \leq c|x - y|.$$

Furthermore, we suppose that the coefficient σ fulfills an invertibility condition of the following form:

Hypothesis 1.4. *The number d of driving fBm equals the dimension of the state space $\mathbb{R}^d$ for Y. In addition, denoting by σ the $d \times d$ matrix with columns $\sigma_1, \ldots, \sigma_d$, we assume that σ is invertible.*

Notice that Hypothesis 1.3 and 1.4 entail the following result (see next section for more details): for a given $\theta \in \Theta$ the solution to equation (3) converges for $t \to \infty$ almost surely to a stationary and ergodic stochastic process $(\overline{Y}_t, t \geq 0)$.

Let us now introduce the concept of LAN property in our context. Towards this aim, for any $\theta \in \Theta$ and $\lambda < H$, we denote by $\mathbf{P}_\theta$ (resp. $\mathbf{P}_\theta^\tau$) the probability laws of the solution to equation (3) in the spaces $\mathcal{C}^\lambda(\mathbb{R}_+; \mathbb{R}^d)$ (resp. $\mathcal{C}^\lambda([0,\tau]; \mathbb{R}^d)$). Moreover, we assume that the process Y is observed continuously in $[0, \tau]$. In this context, the definition of LAN property takes the following form:

Definition 1.5. *We say that the parametric statistical model $\{\mathbf{P}_\theta,\, \theta \in \Theta\}$ satisfies the LAN property at $\theta \in \Theta$ if there exist:*

(i) *A $q \times q$ invertible matrix $\varphi_\tau(\theta)$,*

(ii) *A $q \times q$ positive definite matrix $\Gamma(\theta)$,*

such that for any $u \in \mathbb{R}^q$, the following limit holds true as $\tau \to \infty$:

$$\log \left(\frac{d\mathbf{P}^\tau_{\theta + \varphi_\tau(\theta) u}}{d\mathbf{P}^\tau_\theta} \right) \xrightarrow{\mathcal{L}(\mathbf{P}_\theta)} u^{\mathrm{T}} \mathcal{N}(0, \Gamma(\theta)) - \frac{1}{2} u^{\mathrm{T}} \Gamma(\theta) u,$$

where $\mathcal{N}(0, \Gamma(\theta))$ is a centered q-dimensional Gaussian random variable with covariance matrix $\Gamma(\theta)$.

The LAN property is a fundamental concept in asymptotic theory of statistics, which was developed by Le Cam [24]. The essence of LAN is that the local log likelihood ratio is asymptotically normally distributed, with a locally constant covariance matrix and a mean equal to minus one half the variance. The main application of the LAN property is the following Minimax Theorem. It asserts (if LAN holds true) an asymptotic (minimax) lower bound for the risk with respect to a loss function, for any sequence $\hat{\theta}_\tau$ of estimators of θ.

Theorem 1.6 (Minimax Theorem)**.** [17, Theorem 12.1] *Suppose that the family of probability measures $(\mathbf{P}_\theta)_{\theta \in \Theta}$ satisfies the LAN property at a point θ. Let $(\hat{\theta}_\tau)_{\tau \geq 0}$ be a family of estimators of the parameter θ and $l : \mathbb{R}^q \to [0, +\infty)$ be a loss function such that $l(0) = 0$, $l(x) = l(|x|)$ and $l(|x|) \leq l(|y|)$ if $|x| \leq |y|$. Then we have:*

$$\lim_{\delta \to 0} \liminf_{\tau \to \infty} \sup_{|\theta' - \theta| < \delta} \mathbf{E}_{\theta'} \left[l \left(\varphi_\tau^{-1}(\theta) \left(\hat{\theta}_\tau - \theta' \right) \right) \right] \geq \mathbf{E}_\theta \left[l\left(Z \right) \right], \tag{4}$$

where $\mathcal{L}(Z) = \mathcal{N}(0, \Gamma(\theta)^{-1})$.

Observe that when the loss function l in Theorem 1.6 is the quadratic function, the lower bound in (4) is simply the trace of $\Gamma(\theta)^{-1}$. A sequence of estimators that attains this asymptotic bound is called asymptotic minimax efficient. Therefore, Theorem 1.6 opens the way to a theory which mimics the concept of efficient estimator from the Cramér-Rao lower bound.

As one can see, the LAN property is an important tool in order to quantify the identifiability of a system. The aim of this paper is thus to show the following result for our equation (3).

Theorem 1.7. *Assume Hypothesis 1.3, and recall that $\mathbf{P}_\theta^\tau$ stands for the probability law of the solution to equation (3) in the space $\mathcal{C}^\lambda([0,\tau];\mathbb{R}^d)$. Then, for any $\theta \in \Theta$ and $u \in \mathbb{R}^q$ fixed, as $\tau \to \infty$, we have:*

$$\log \frac{d\mathbf{P}^\tau_{\theta+\frac{u}{\sqrt{\tau}}}}{d\mathbf{P}^\tau_\theta} \xrightarrow{\mathcal{L}(\mathbf{P}_\theta)} u^{\mathrm{T}} \mathcal{N}(0,\Gamma(\theta)) - \frac{1}{2} u^{\mathrm{T}} \Gamma(\theta) u, \tag{5}$$

where the matrix $\Gamma(\theta)$ is defined by:

$$\Gamma(\theta) := \int_{\mathbb{R}_+^2} \frac{\mathbf{E}_\theta[(\hat{b}(\overline{Y}_0;\theta) - \hat{b}(\overline{Y}_{r_1};\theta))^{\mathrm{T}} (\sigma^{-1})^{\mathrm{T}} \sigma^{-1} (\hat{b}(\overline{Y}_0;\theta) - \hat{b}(\overline{Y}_{r_2};\theta))]}{r_1^{1/2+H} r_2^{1/2+H}} \, dr_1 dr_2, \tag{6}$$

In (6), recall that $\overline{Y}$ is the ergodic limit of our process Y and that we have set $\hat{b}$ for the Jacobian matrix $\partial_\theta b$.

Though our result encompasses a wide range of coefficients b, it is worth illustrating it on the simple linear case, that is for the real-valued fractional Ornstein-Uhlenbeck process corresponding to $b(x;\theta) = -\theta\, x$ and $\sigma \equiv \mathbf{1}$. In this case we have:

$$Y_t = -\theta \int_0^t Y_s \, ds + B_t, \qquad t \in [0,\tau], \quad \theta > 0.$$

Let $\hat{\theta}_\tau$ be the MLE of θ from the observation of Y in $[0,\tau]$. Then, it is well-known (see [3, 18]) that

$$\lim_{\tau\to\infty} \hat{\theta}_\tau = \theta \quad \text{a.s.}$$

and that

$$\mathcal{L}(\mathbf{P}_\theta) - \lim_{\tau\to\infty} \sqrt{\tau}(\hat{\theta}_\tau - \theta) = \mathcal{N}(0, 2\theta). \tag{7}$$

In particular, this last result already suggests that the rate of convergence for the LAN property in this case is $\tau^{-1/2}$, as mentioned in [30, p.162]. Our result confirms this intuition, and one can see that the MLE reaches this optimal rate of order $\tau^{-1/2}$. However, observe that the identification of (6) and (7) is far from being clear, leaving open the question of efficiency of the ML estimator in the fractional Ornstein-Uhlenbeck case.

The LAN property for stochastic processes has been widely explored in the literature. To mention a few references close to our contribution, the case of ergodic diffusion processes observed continuously (with an unknown parameter on the drift coefficient) is dealt with in e.g. [22, Proposition 2.2]. The proof follows from Girsanov's theorem, ergodic properties and the central limit theorem for Brownian martingales. The rate of convergence achieved in [22, Proposition 2.2] is $\sqrt{\tau}$ like in our Theorem 1.7, with a function $\Gamma(\theta)$ given by:

$$\Gamma(\theta) = \mathbf{E}_\theta[\hat{b}(\overline{Y};\theta)^{\mathrm{T}} \sigma^{-1}(\overline{Y})^{\mathrm{T}} \sigma^{-1}(\overline{Y}) \hat{b}(\overline{Y};\theta)],$$

which should be compared to our expression (6). The case where the Brownian diffusion process is observed discretely was solved in [11] using an integration by parts formula taken from the Malliavin calculus. Let us also mention that the LAN property is obtained in [7] for some discretely observed fractional noises including fractional Brownian motion. This last result is achieved by a direct expansion of log-likelihood functions, plus a thorough analysis based on properties of Toeplitz matrices.

Finally let us say a few words about the strategy we have followed in order to handle the case of equation (3). In order to compare likelihood functions for different values of the parameter, we take the obvious choice of applying Girsanov's theorem for fractional Brownian motion (following the steps of [26]). We are then left with two technical difficulties in order to get asymptotic results: (i) Handle the singularities popping out of the fractional derivatives in the Girsanov exponent. (ii) Get ergodic results in Hölder type norms for our process Y. Let us also mention that CLTs for martingales, which were an essential tool in the diffusion case, are unavailable in our fBm situation. We sort out this problem by means of Malliavin calculus techniques. Specifically, we shall derive some concentration properties for the Girsanov exponent by means of a Poincaré type inequality (see relation (16) below). This means that some Malliavin derivatives will have to be conveniently upper bounded, which is a non negligible part of our efforts.

Our paper is structured as follows: Section 2 is devoted to preliminary results, including Malliavin calculus background and the derivation of most of the concentration properties for the process Y needed later on. Then we prove Theorem 1.7 in Section 3, following the strategy described above.

Let us close this introduction by giving a set of notations which will prevail until the end of the paper.

Notation 1.8. *We use the following conventions: for 2 quantities a and b, we write $a \lesssim b$ if there exists a universal constant c such that $a \leq c\, b$. In the same way, we write $a \asymp b$ whenever $a \lesssim b$ and $b \lesssim a$. If f is a vector-valued function defined on an interval $[0, \tau]$ and $s, t \in [0, \tau]$, δf_{st} designates the increment $f_t - f_s$.*

2. Auxiliary Results

In this section we first recall some basic facts about stochastic analysis for a fBm B, and also about Young integrals. Then we shall derive some pathwise and probabilistic estimates for equation (3) under the coercive Hypothesis 1.3.

2.1. Stochastic analysis related to B.

The ergodic properties of equation (3) are accurately encoded by the fBm representation given in [15], which goes back to the original paper [25]. We present this construction here for a one-dimensional fBm, for sake of simplicity. Obvious generalizations to the d-dimensional case are left to the reader.

2.1.1. *Wiener space related to B.* Let W be a two-sided Brownian motion, and H a Hurst parameter in $(0, 1)$. We consider the process B defined for $t \in \mathbb{R}$ by:

$$\begin{aligned} B_t &= c_H \int_{\mathbb{R}_-} (-r)^{H-1/2} \left[dW_{t+r} - dW_r \right] \qquad (8) \\ &= c_H \left\{ \int_{-\infty}^0 \left[(-(r-t))^{H-1/2} - (-r)^{H-1/2} \right] dW_r - \int_0^t (-(r-t))^{H-1/2} dW_r \right\}. \end{aligned}$$

Then B is well-defined as a fBm, that is a centered Gaussian process with covariance given by (1).

The process B is closely related to the following fractional derivatives: for $\alpha \in (0,1)$ and $\varphi \in \mathcal{C}_c^\infty(\mathbb{R})$ we set

$$\left[D_+^\alpha \varphi\right]_t = c_\alpha \int_{\mathbb{R}_+} \frac{\varphi_t - \varphi_{t-r}}{r^{1+\alpha}} \, dr, \quad \text{and} \quad \left[I_+^\alpha \varphi\right]_t = \tilde{c}_\alpha \int_{\mathbb{R}_+} \varphi_{t-r} \, r^{\alpha-1} \, dr.$$

With this notation in mind, the following proposition identifies a convenient operator which transform W into B as in (8).

Proposition 2.1. *For $w \in \mathcal{C}_c^\infty(\mathbb{R})$ and $H \in (0,1)$, set*

$$[K_H w]_t = c_H \int_{\mathbb{R}_-} (-r)^{H-1/2} \left[\dot{w}_{t+r} - \dot{w}_r\right] dr.$$

Then the following holds true:

(i) *There exists a constant c_H such that*

$$[K_H w]_t = \begin{cases} -c_H \left([I_+^{H-1/2} w]_t - [I_+^{H-1/2} w]_0 \right), & \text{for } H > \frac{1}{2} \\ -c_H \left([D_+^{1/2-H} w]_t - [D_+^{1/2-H} w]_0 \right), & \text{for } H < \frac{1}{2}. \end{cases}$$

(ii) *For $H > 1/2$, K_H can be extended as an isometry from $L^2(\mathbb{R})$ to $I_+^{H-1/2}(L^2(\mathbb{R}))$.*

(iii) *There exists a constant c_H such that we have $K_H^{-1} = c_H K_{1-H}$.*

The natural Wiener space related to our fractional Brownian motion indexed by $\mathbb{R}$, as represented in (8), is given as follows (see [15] for further details): one can take $\mathcal{B}$ as the underlying space, where

$$\mathcal{B} = \left\{ f \in \mathcal{C}(\mathbb{R}; \mathbb{R}^d); \; \frac{|f_t|}{1+|t|} < \infty \right\}.$$

For this choice of $\mathcal{B}$, equipped with the law $\mathbf{P}$ of our fractional Brownian motion B, there are various ways to express the generic form of an element of the Cameron-Martin space $\bar{\mathcal{H}}$. Among those we will choose the following one: h is an element of $\bar{\mathcal{H}}$ iff there exists an element X_h in the first chaos such that:

$$h_t = \mathbf{E}[B_t X_h], \quad \text{and} \quad \|h\|_{\bar{\mathcal{H}}} = \|X_h\|_{L^2(\Omega)}. \tag{9}$$

Moreover, one can check that $(\mathcal{B}, \bar{\mathcal{H}}, \mathbf{P})$ is an abstract Wiener space. Notice that the space $\bar{\mathcal{H}}$ has to be distinguished from the space $\mathcal{H}$ usually considered (e.g in [28]) for the Malliavin calculus with respect to B. For the Wiener process W, we would have for instance $\mathcal{H} = L^2(\mathbb{R})$, while $\bar{\mathcal{H}}$ has to be identified with the Sobolev space $H^1(\mathbb{R})$.

Let us recall a definition concerning variation norms which can be found e.g in [10]. For $0 \le s < t \le \tau$ and a function $f : [s,t] \to \mathbb{R}^d$, we set:

$$\|h\|_{1-\mathrm{var};[s,t]} = \sup \left\{ \sum_{t_i \in \Pi_{st}} \left| \delta f_{t_i t_{i+1}} \right| ; \, \Pi_{st} \text{ partition of } [s,t] \right\}.$$

With this definition in hand, we can label the following property of elements of $\bar{\mathcal{H}}$ for further use:

$$\|h\|_{1-\mathrm{var};[s,t]} \le (t-s)^H \, \|h\|_{\bar{\mathcal{H}}}. \tag{10}$$

Next we give a representation of the Cameron-Martin space on a finite interval $[0,\tau]$, namely $\bar{\mathcal{H}}([0,\tau])$, which can be found in [8]. To this aim, let us recall the definition of fractional integrals of order $\alpha \in (0,1]$ on a finite interval:

$$[I_{0+}^{\alpha} f]_t = \frac{1}{\Gamma(\alpha)} \int_0^t (t-u)^{\alpha-1} f_u \, du, \quad \text{and} \quad [I_{\tau-}^{\alpha} f]_t = \frac{1}{\Gamma(\alpha)} \int_t^{\tau} (u-t)^{\alpha-1} f_u \, du. \tag{11}$$

With this additional definition in hand and assuming $H > 1/2$, a function h lies into the space $\bar{\mathcal{H}}([0,\tau])$ if it can be written as

$$h = \hat{K}_\tau f, \quad \text{with} \quad \hat{K}_\tau f = I_{0+}^1 [x^\kappa I_{0+}^\kappa (x^{-\kappa} f)] \quad \text{and} \quad f \in L^2([0,\tau]), \tag{12}$$

where we have set $\kappa = H - 1/2$. Related to this representation, we have $\|h\|_{\bar{\mathcal{H}}([0,\tau])} = \|f\|_{L^2([0,\tau])}$ whenever $h = \hat{K}_\tau f$. Notice that this space will play a prominent role in our estimates on finite intervals, associated to another space called $\mathcal{H}([0,\tau])$ which pops out in Malliavin calculus computations (as mentioned above for $\mathcal{H}$ and $\bar{\mathcal{H}}$). Indeed, if $h \in \bar{\mathcal{H}}([0,\tau])$ admits the representation (12), then for a smooth function g on $[0,\tau]$ we have:

$$\int_0^\tau g_s \, dh_s = \int_0^\tau g_s \, s^\kappa [I_{0+}^\kappa (x^{-\kappa} f)]_s \, ds = \int_0^\tau [K_\tau^* g]_s f_s \, ds, \tag{13}$$

where we have set:

$$K_\tau^* g = x^{-\kappa} I_{\tau-}(x^\kappa g),$$

and we recall that $I_{\tau-}$ is defined by (11). and where the integral $\int_0^\tau g_s \, dh_s$ is understood in the Young sense. The closure of smooth functions under the inner product $\langle K_\tau^* g_1, K_\tau^* g_2 \rangle_{L^2([0,\tau])}$ is the space which is called $\mathcal{H}([0,\tau])$. We shall resort to the following formulae (for which we refer e.g to [8]) concerning $\mathcal{H}([0,\tau])$ in the sequel:

$$\|g\|_{\mathcal{H}([0,\tau])}^2 = c_H \int_{[0,\tau]^2} g_{s_1} |s_1 - s_2|^{2H-2} g_{s_2} \, ds_1 ds_2, \qquad \|g\|_{\mathcal{H}([0,\tau])} \le c_H \tau^\kappa \|g\|_{L^2([0,\tau])}. \tag{14}$$

The following relation between norms in $\bar{\mathcal{H}}$ and $W^{1,2}$ will also be needed in the sequel. We include its proof here for sake of completeness, because we haven't been able to track it down in the literature.

Proposition 2.2. *Let $H > 1/2$ and $\tau > 0$. Consider the space $\bar{\mathcal{H}}([0,\tau])$ defined by (12) and the usual Sobolev space $W^{1,2}([0,\tau])$. Then for all $h \in \bar{\mathcal{H}}([0,\tau])$ we have:*

$$\|h\|_{W^{1,2}([0,\tau])} \le c_H \, \tau^{H-\frac{1}{2}} \, \|h\|_{\bar{\mathcal{H}}}. \tag{15}$$

Proof. Start from the representation (12) for a typical element $h \in \bar{\mathcal{H}}([0,\tau])$, and recall that we have set $\kappa = H - 1/2$. Then we have:

$$\|h\|_{W^{1,2}([0,\tau])} = \sup \left\{ \int_0^\tau \ell_s \, \dot{h}_s \, ds \, ; \, \|\ell\|_{L^2([0,\tau])} = 1 \right\}.$$

Consider now a generic element $\ell \in L^2([0,\tau])$ such that $\|\ell\|_{L^2([0,\tau])} = 1$. We resort to (13) in order to get:

$$\left| \int_0^\tau \ell_s \, \dot{h}_s \, ds \right| \le \|K_\tau^* \ell\|_{L^2([0,\tau])} \, \|f\|_{L^2([0,\tau])} = \|K_\tau^* \ell\|_{L^2([0,\tau])} \, \|h\|_{\bar{\mathcal{H}}([0,\tau])} \,,$$

where we recall that f comes from representation (12). Furthermore, invoking our bound (14), we end up with:

$$\left|\int_0^\tau \ell_s\, \dot{h}_s\, ds\right| \le c_H\, \tau^\kappa\, \|\ell\|_{L^2([0,\tau])}\, \|h\|_{\bar{\mathcal{H}}([0,\tau])} = c_H\, \tau^\kappa\, \|h\|_{\bar{\mathcal{H}}([0,\tau])}\,,$$

which proves our claim. □

2.1.2. *Malliavin calculus tools.* We refer to [28] and [34] for basic definitions concerning Malliavin derivatives D and Sobolev spaces $\mathbb{D}^{k,p}$ related to B. Our concentration arguments for the solution to (3) will heavily rely on the following Poincaré type inequality borrowed from [34, page 76].

Proposition 2.3. *Let $F : \mathcal{B} \to \mathbb{R}$ be a functional in $\mathbb{D}^{1,2}$. Then,*

$$\mathbf{E}\left[|F - \mathbf{E}[F]|\right] \le \frac{\pi}{2}\, \mathbf{E}\left[\|DF\|_{\bar{\mathcal{H}}}\right], \tag{16}$$

where we recall that D denotes the Malliavin derivative.

Next we prepare the ground for our Girsanov transform by relating a general drift with respect to a fractional Brownian motion with its counterpart with respect to W (we continue in the one-dimensional setting for the sake of simplicity). This proposition is a slight extension of [15, Lemma 4.2].

Proposition 2.4. *Let $H \in (1/2, 1)$, b^1, b^2 be two paths in $I_+^{H-1/2}(L^2(\mathbb{R}))$, and suppose that $b^j = K_H w^j$ for $j = 1, 2$, where K_H is defined at Proposition 2.1. We also assume that for $t \ge 0$, b^1 and b^2 are linked by the relation:*

$$b_0^2 = b_0^1, \quad \textit{and} \quad b_t^2 = b_t^1 + \int_0^t g_{b,s}\, ds,$$

for a function g_b in $L^2(\mathbb{R})$. Then we also have:

$$w_t^2 = w_t^1 + \int_0^t g_{w,s}\, ds, \quad \textit{with} \quad g_w = D_+^{H-1/2} g_b.$$

Proof. By [15, Lemma 4.2, relation (4.11a)] and [29, relation (4)],

$$g_{w,t} = c_H \frac{d}{dt} \int_{-\infty}^t \frac{g_{b,u}}{(t-u)^{H-1/2}}\, du = c_H \left[D_+^1 I_+^{H+1/2} g_b\right]_t = \left[D_+^{H-1/2} g_b\right]_t,$$

which finishes our proof. □

The previous proposition yields a version of Girsanov's theorem in a fractional Brownian context, inspired by [26] (see also [29]):

Proposition 2.5. *Let B be a fractional Brownian motion and g be an adapted process in $I_+^{H-1/2}(L^2(\mathbb{R}))$. We set $V_t = \int_0^t g_s\, ds$ and $Q_t = B_t + V_t$ for $t \ge 0$. Then Q is a fractional Brownian motion on $[0, \tau]$, under the probability $\tilde{\mathbf{P}}$ defined by:*

$$\frac{d\tilde{\mathbf{P}}}{d\mathbf{P}}_{|_{[0,\tau]}} = e^{-L}, \quad \textit{where} \quad L = \int_0^\tau [D_+^{H-1/2} g]_u\, dW_u + \frac{1}{2}\int_0^\tau [D_+^{H-1/2} g]_u^2\, du, \tag{17}$$

provided $D_+^{H-1/2} g$ satisfies Novikov's conditions on $[0, \tau]$.

2.2. **Generalized Riemann-Stieltjes Integrals.** We are focusing here on the case of a Hurst parameter $H > 1/2$ for sake of simplicity, so that all the stochastic integrals with respect to B should be understood in the Young (or Riemann-Stieltjes) sense. In order to recall the definition of this integral, we set

$$\|f\|_{\infty;[a,b]} = \sup_{t\in[a,b]} |f(t)|, \qquad |f|_{\lambda;[a,b]} = \sup_{s,t\in[a,b]} \frac{|f(t)-f(s)|}{|t-s|^{\lambda}}$$

where $f : \mathbb{R} \to \mathbb{R}^n$ and $\lambda \in (0,1)$. We also recall the classical definition of Hölder-continuous functions:

$$C^{\lambda}([a,b];\mathbb{R}^n) = \left\{ f : [a,b] \to \mathbb{R}^n;\ \|f\|_{\infty;[a,b]} + |f|_{\lambda;[a,b]} < \infty \right\}.$$

Now, let $f \in C^{\lambda}([a,b];\mathbb{R})$ and $g \in C^{\mu}([a,b];\mathbb{R})$ with $\lambda+\mu > 1$. Then it is well known that the Riemann-Stieltjes integral $\int_a^b f_s\, dg_s$ exists, and can be expressed as a limit of Riemann sums.

In the sequel, the only property on Young's integral we shall use is the classical chain rule for changes of variables. Indeed, let $f \in C^{\lambda}([a,b];\mathbb{R})$ with $\lambda > 1/2$ and $F \in C^1(\mathbb{R};\mathbb{R})$. Then we have:

$$F(f_y) - F(f_a) = \int_a^y F'(f_s)\, df_s, \qquad y \in [a,b]. \tag{18}$$

2.3. **Basic properties of solutions to SDEs.** We first state the existence and uniqueness result for equation (3) borrowed from [15, Lemma 3.9].

Proposition 2.6. *Under Hypothesis (1.3), there exists a unique continuous pathwise solution to equation (3) on any arbitrary interval $[0,\tau]$. Moreover the map $Y : (y_0, B) \in \mathbb{R}^d \times \mathcal{C}([0,\tau];\mathbb{R}^d) \to \mathcal{C}([0,\tau];\mathbb{R}^d)$ is locally Lipschitz continuous.*

In addition, exploiting the integrability properties of the stationary fractional Ornstein-Uhlenbeck process, the following uniform estimates hold true. They are shown with more details in [12], and are obtained by comparing Y with the solution to the equation $dU_t = -U_t\, dt + dB_t$ (much in the same way as in the next section 2.4).

Proposition 2.7. *Assume Hypothesis 1.3 holds true and let Y be the solution to equation (3). Then for any $\theta \in \Theta$ and $p \geq 1$ there exist constants $c_p, k_p > 0$ such that*

$$\mathbf{E}\left[|Y_t|^p\right] \leq c_p, \qquad \text{and} \quad \mathbf{E}\left[|\delta Y_{st}|^p\right] \leq k_p |t-s|^{pH},$$

for all $s,t \geq 0$, where we recall that we have set $\delta Y_{st} = Y_t - Y_s$ as mentioned in Notation 1.8.

Let us now state a pathwise bound on the increments of Y, where we use a similar idea as in [19], in order to translate the moment estimates into pathwise bounds.

Proposition 2.8. *Assume Hypothesis 1.3 holds true and let Y be the solution to equation (3). Then for all $\varepsilon \in (0,H)$ there exists a random variable $Z_\varepsilon \in \cap_{p\geq 1} L^p(\Omega)$ such that almost surely we have:*

$$|Y_t| \leq Z_\varepsilon\, (1+t)^{2\varepsilon}, \quad \text{and} \quad |\delta Y_{st}| \leq Z_\varepsilon\, (1+t)^{2\varepsilon}\, |t-s|^{H-\varepsilon}, \tag{19}$$

uniformly for $0 \leq s \leq t$.

Proof. We focus on the bound for δY_{st}, the estimate on $|Y_t|$ being shown in the same manner. Set $\gamma = H - \varepsilon$ and consider $n \geq 1$. Thanks to Garsia's lemma (see e.g. [13]) and recalling our Notation 1.8, for all $s,t \in [0,n]$ we have $|\delta Y_{st}| \lesssim \mathcal{N}_{\gamma,p,n}|t-s|^{\gamma}$, where $p \geq 1$ and $\mathcal{N}_{\gamma,p,n}$ is defined by:

$$\mathcal{N}_{\gamma,p,n} \equiv \left(\int_0^n \int_0^n \frac{|\delta Y_{uv}|^p}{|v-u|^{\gamma p+2}} \, du dv \right)^{\frac{1}{p}}.$$

Furthermore, owing to Proposition 2.7 we immediately get:

$$\mathbf{E}^{1/p} \left[\mathcal{N}_{\gamma,p,n}^p \right] \lesssim n^{\varepsilon},$$

under the constraint $p \geq \varepsilon^{-1}$. Set now $Z_\varepsilon = \sup_{n\in\mathbb{N}} \frac{|\mathcal{N}_{\gamma,p,n}|}{n^{2\varepsilon}}$. Then for all $q > \varepsilon^{-1}$ we have:

$$\mathbf{E}\left[Z_\varepsilon^q\right] = \mathbf{E}\left[\sup_{n\in\mathbb{N}} \frac{|\mathcal{N}_{\gamma,p,n}|^q}{n^{2\varepsilon q}} \right] \leq \sum_{n=1}^{\infty} \frac{\mathbf{E}\left[|\mathcal{N}_{\gamma,p,n}|^q\right]}{n^{2\varepsilon q}} \lesssim \sum_{n=1}^{\infty} \frac{1}{n^{\varepsilon q}} < \infty, \tag{20}$$

namely we have found $Z_\varepsilon \in \cap_{p\geq\varepsilon^{-1}} L^p(\Omega)$. Obviously this also yields $Z_\varepsilon \in \cap_{p\geq 1} L^p(\Omega)$. Finally, replace $n \geq 1$ by $[t]+2$ to obtain our pathwise bound (19). □

Let h be a generic element of $\bar{\mathcal{H}}$ (this space being defined by (9)), and set Y^h for the solution to our equation (3) driven by $B+h$. We shall need an extension of the previous proposition to Y^h, whose proof is left to the reader since it follows along the same lines as Proposition 2.8.

Proposition 2.9. *For $h \in \bar{\mathcal{H}}$ and $\theta \in \Theta$, let Y^h be the solution to:*

$$Y_t^h = y_0 + \int_0^t b(Y_s^h;\theta)\, ds + \sum_{j=1}^d \sigma_j \left(B_t^j + h_t^j\right), \qquad t \in [0,\tau]. \tag{21}$$

Under the same assumptions as for Proposition 2.7, for all $\varepsilon \in (0,H)$ there exists a random variable $Z_\varepsilon \in \cap_{p\geq 1} L^p(\Omega)$ such that almost surely we have:

$$|\delta Y_{st}^h| \leq Z_\varepsilon \,(1+t)^{2\varepsilon}\, (1+\|h\|_{\bar{\mathcal{H}}})\, |t-s|^{H-\varepsilon}, \tag{22}$$

uniformly for $0 \leq s \leq t$.

We next show that our Girsanov type theorem 2.5 can be applied to the solution to (3).

Proposition 2.10. *Assume Hypothesis 1.3 and let Y be the solution to (3), for a given $\theta \in \Theta$. Consider the d-dimensional process Q_t given by:*

$$Q_t = \int_0^t \sigma^{-1} b(Y_s;\theta)\, ds + B_t.$$

Then Q is a d-dimensional fractional Brownian motion under the probability $\tilde{\mathbf{P}}_\theta$ defined by $\frac{d\tilde{\mathbf{P}}_\theta}{d\mathbf{P}_\theta}\big|_{[0,\tau]} = e^{-L}$, with

$$L = \int_0^\tau \langle \sigma^{-1}[D_+^{H-1/2} b(Y;\theta)]_u, dW_u\rangle + \frac{1}{2}\int_0^\tau |\sigma^{-1}[D_+^{H-1/2} b(Y;\theta)]_u|^2\, du. \tag{23}$$

Proof. We first need to show that $D_+^{H-1/2}b(Y;\theta)$ is well defined on $[0,\tau]$, which follows since the function $u\mapsto b(Y_u;\theta)$ is $(H-\varepsilon)$-Hölder, and thus $(H-1/2)$-Hölder.

Second, we need to verify Novikov's condition (see [9, Theorem 1.1]), that is, that there exists $\lambda>0$ such that

$$\sup_{t\in[0,\tau]}\mathbf{E}_\theta\left[\exp\left(\lambda\int_0^t|\sigma^{-1}[D_+^{H-1/2}b(Y;\theta)]_s|^2ds\right)\right]<\infty. \tag{24}$$

Towards this aim, we extend the definition of Y to $\mathbb{R}_-$ by setting $Y_t=y_0$ for all $t\le 0$. Then,

$$|\sigma^{-1}[D_+^{H-1/2}b(Y)]_s|=c_H\left|\int_{\mathbb{R}_+}\frac{\sigma^{-1}(b(Y_s;\theta)-b(Y_{s-r};\theta))}{r^{H+1/2}}\,dr\right|\le c_H(A_1+A_2),$$

where

$$A_1=\left|\int_0^s\frac{\sigma^{-1}(b(Y_s;\theta)-b(Y_{s-r};\theta))}{r^{H+1/2}}\,dr\right|\quad\text{and}\quad A_2=\left|\int_s^\infty\frac{\sigma^{-1}(b(Y_s;\theta)-b(y_0;\theta))}{r^{H+1/2}}\,dr\right|.$$

Using the mean value theorem and the uniform boundedness of $\partial_x b$, the second term is easily bounded as:

$$A_2\le c_H\frac{|Y_s-y_0|}{s^{H-1/2}}.$$

Similarly, we get that for all $0<\varepsilon<\frac{1}{2}$,

$$A_1\le c\sup_{0<r\le s}\frac{|Y_s-Y_{s-r}|}{r^{H-\varepsilon}}\int_0^s\frac{1}{r^{\frac{1}{2}+\varepsilon}}dr=c_Hs^{\frac{1}{2}-\varepsilon}\sup_{0<r\le s}\frac{|Y_s-Y_{s-r}|}{r^{H-\varepsilon}}$$

Then, proceeding as in [29] (where Fernique's theorem is invoked), we get the desired result. □

2.4. **Contraction properties of the SDE.** The inward Hypothesis 1.3 for the function b yields some contraction properties for the solution to equation (3), which will be useful for our purposes. We summarize some of them in this section.

Let us start by giving a result quantifying the way two solutions starting from different initial conditions are getting close as t goes to ∞.

Proposition 2.11. *Assume Hypothesis (1.3). Let X be the solution to equation (3) starting from initial condition x, and Y the solution to the same equation starting from initial condition y. Then we have:*

$$\left|X_t-Y_t\right|\le\left|x-y\right|e^{-\alpha t}, \tag{25}$$

where the constant α is introduced in Hypothesis 1.3.

Proof. Recall that we have assumed:

$$\langle b(x;\theta)-b(y;\theta),\,x-y\rangle\le-\alpha|x-y|^2.$$

Now it is readily checked that the difference $X-Y$ satisfies

$$X_t-Y_t=(x-y)+\int_0^t\left[b(X_s;\theta)-b(Y_s;\theta)\right]\,ds,\qquad t\ge 0.$$

Applying the change of variable formula in the Young setting (18), we thus get:

$$
\begin{aligned}
|X_t - Y_t|^2 &= |x-y|^2 + 2\int_0^t \langle b(X_s;\theta) - b(Y_s;\theta),\, X_s - Y_s\rangle\, ds \\
&\le |x-y|^2 - 2\alpha \int_0^t |X_s - Y_s|^2\, ds.
\end{aligned}
$$

Our claim thus follows by a direct application of Gronwall's lemma. $\square$

The previous proposition can be enhanced in order to get a bound on increments for $X - Y$.

Proposition 2.12. *Under the same assumptions as in Proposition 2.11, the following inequality holds true for* $0 \le u \le v < \infty$*:*

$$
\left|\delta X_{uv} - \delta Y_{uv}\right| \le c\left|x-y\right| e^{-\alpha u}|u-v|, \tag{26}
$$

where the constant α *is introduced in Hypothesis 1.3.*

Proof. The same computations as in the proof of Proposition 2.11 show that

$$
\delta X_{uv} - \delta Y_{uv} = \int_u^v \left[b(X_s;\theta) - b(Y_s;\theta)\right]\, ds.
$$

We then use the mean value theorem and the uniform bound on $\partial_x b$, together with inequality (25), which yields our result. $\square$

We now study the contraction properties of the map $h \mapsto Y^h$ for $h \in \bar{\mathcal{H}}$.

Proposition 2.13. *Assume Hypothesis 1.3. Let* h *be an element of* $\bar{\mathcal{H}}$ *and* Y, Y^h *be solutions to the system:*

$$
Y_t = y_0 + \int_0^t b(Y_s;\theta)\, ds + \sum_{j=1}^d \sigma_j B_t^j, \quad \text{and} \quad Y_t^h = y_0 + \int_0^t b(Y_s^h;\theta)\, ds + \sum_{j=1}^d \sigma_j \left(B_t^j + h_t^j\right). \tag{27}
$$

Then the difference $Y^h - Y$ *satisfies:*

$$
|Y_t^h - Y_t| \le c \exp\left(-\frac{\alpha\, t}{2}\right) \|h\|_{\bar{\mathcal{H}}}, \tag{28}
$$

uniformly in $t \in \mathbb{R}_+$.

Proof. Consider $s \le t$ and proceed again as for the proof of Proposition 2.11 in order to get:

$$
|Y_s^h - Y_s|^2 \le -2\alpha \int_0^s |Y_r^h - Y_r|^2\, dr + \sum_{j=1}^d \int_0^t \left\langle Y_r^h - Y_r,\, \sigma_j\right\rangle dh_r^j,
$$

and thus, applying Cauchy-Schwarz inequality we end up with:

$$
|Y_s^h - Y_s|^2 \le -2\alpha \int_0^s |Y_r^h - Y_r|^2\, dr + \sum_{j=1}^d |\sigma_j|\, \|h^j\|_{W^{1,2}([0,t])} \left(\int_0^s |Y_r^h - Y_r|^2\, ds\right)^{1/2}.
$$

Now recall that $ab \le \frac{a^2}{2\varepsilon^2} + \frac{\varepsilon^2 b^2}{2}$ for $a, b, \varepsilon > 0$. Applying this inequality to $\varepsilon^2 = 2\alpha$ and resorting to Proposition 2.2, we obtain:

$$|Y_s^h - Y_s|^2 \le -\alpha \int_0^s |Y_r^h - Y_r|^2 \, dr + \sum_{j=1}^d |\sigma_j|^2 \frac{\|h\|_{\bar{\mathcal{H}}}^2 \, t^{2H-1}}{4\alpha}.$$

The last inequality being valid uniformly in $s \le t$, our claim is thus easily deduced by a direct application of Gronwall's lemma plus the fact that $t^{2H-1} \exp(-\alpha t/2)$ is bounded over the time interval $\mathbb{R}_+$. □

We now push forward our analysis to include the increments of processes.

Proposition 2.14. *Assume Hypothesis 1.3. Let h be an element of $\bar{\mathcal{H}}$ and Y^h, Y be solutions to the system (27). Then for $0 \le u \le v$, the difference $Y^h - Y$ satisfies:*

$$|\delta Y_{uv}^h - \delta Y_{uv}| \le c_1 \exp\left(-\frac{\alpha\, u}{2}\right) \|h\|_{\bar{\mathcal{H}}} \, (v-u)^{H/2}.$$

Proof. Like in the proof of Proposition 2.13, we start from the inequality:

$$|\delta Y_{uv}^h - \delta Y_{uv}|^2 \le -2\alpha \int_u^v |Y_s^h - Y_s|^2 \, ds + \sum_{j=1}^d \int_u^v \left\langle Y_s^h - Y_s, \, \sigma_j \right\rangle dh_s^j,$$

and since the first term on the right hand side above is negative, we get:

$$|\delta Y_{uv}^h - \delta Y_{uv}|^2 \le \sum_{j=1}^d \int_u^v \left\langle Y_s^h - Y_s, \, \sigma_j \right\rangle dh_s^j \le \sup_{s \in [u,v]} |Y_s^h - Y_s| \sum_{j=1}^d |\sigma_j| \|h^j\|_{1-\text{var},[u,v]}.$$

We now plug relations (10) and (28) in order to get:

$$|\delta Y_{uv}^h - \delta Y_{uv}|^2 \le c_1 \exp\left(-\alpha\, u\right) \|h\|_{\bar{\mathcal{H}}}^2 (v-u)^H.$$

□

As a consequence of Propositions 2.13 and 2.14, the following bounds for the $\bar{\mathcal{H}}$-norm of the Malliavin derivative of the process and its increments hold.

Corollary 2.15. *Assume Hypothesis 1.3. Let Y_t be the solution to the system (3). Then, for all $t > 0$, Y_t belongs to $\mathbb{D}^{1,2}$, and the Malliavin derivative satisfies that*

$$\|DY_t\|_{\bar{\mathcal{H}}} \le c \exp\left(-\frac{\alpha\, t}{2}\right), \tag{29}$$

uniformly in $t \in \mathbb{R}_+$. Moreover, for $0 \le u \le v$,

$$\|D(\delta Y_{uv})\|_{\bar{\mathcal{H}}} \le c_1 \exp\left(-\frac{\alpha\, u}{2}\right) (v-u)^{H/2}, \tag{30}$$

uniformly in u and v.

Proof. The fact that Y_t belongs to $\mathbb{D}^{1,2}$ follows from (28) and [34, Lemma 2, page 72]. In order to check (29), we first approximate Y_t by a sequence of cylindrical random variables, use the definition of the Malliavin derivative for cylindrical random variables (see e.g. [34, Section 2.2]), and apply the inequality of Proposition 2.13. Then passing to the limit by closability and taking $\bar{\mathcal{H}}$ norms, we obtain (29). Finally, (30) is proved similarly using Proposition 2.14. $\square$

2.5. **Ergodic Properties of the SDE.** We recall here some basic facts about the limiting behavior of equation (3), mainly taken from [12]. We still work on the Wiener space $(\mathcal{B}, \bar{\mathcal{H}}, \mathbf{P})$ introduced in Section 2.1.1, and seen as the canonical probability space. Together with the shift operators $\theta_t : \Omega \to \Omega$ defined by

$$\theta_t \omega(\cdot) = \omega(\cdot + t) - \omega(t), \qquad t \in \mathbb{R}, \quad \omega \in \Omega,$$

our probability space defines an ergodic metric dynamical system, see e.g. [14]. In particular, the measure $\mathbf{P}$ is invariant to the shift operators θ_t, i.e. the shifted process $(B_s(\theta_t \cdot))_{s\in\mathbb{R}}$ is still an m-dimensional fractional Brownian motion and for any integrable random variable $F : \Omega \to \mathbb{R}$ we have

$$\lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau F(\theta_t(\omega))\, dt = \mathbf{E}[F],$$

for $\mathbf{P}$-almost all $\omega \in \Omega$.

Under our coercive hypothesis on the drift coefficient of equation (3), the following limit theorem is borrowed from [12, Section 4]. Notice that its proof is based on contraction properties for the stochastic equation.

Theorem 2.16. *Let Hypothesis 1.3 hold, and consider the unique solution Y to equation (3) as given in Proposition 2.6. Then there exists a random variable $\overline{Y} : \Omega \to \mathbb{R}^d$ such that*

$$\lim_{t\to\infty} \left|Y_t(\omega) - \overline{Y}(\theta_t \omega)\right| = 0$$

for $\mathbf{P}$-almost all $\omega \in \Omega$. Moreover, we have $\mathbf{E}[|\overline{Y}|^p] < \infty$ for all $p \geq 1$.

It is worth observing that the convergence of Y towards $\overline{Y}$ can also be quantified in Hölder norm, like in Section 2.4:

Proposition 2.17. *Let $\alpha \in (0, H)$. There exists a random variable Z admitting moments of any order and a constant $c > 0$ such that for all $0 \leq s \leq t$ we have:*

$$\left|Y_t - \overline{Y}_t\right| \leq Z\, e^{-cs} \quad \textit{and} \quad \left|\delta\left[Y - \overline{Y}\right]_{st}\right| \leq Z\, e^{-cs}(t-s)^\alpha.$$

Proof. The proof is done along the same lines as for relation (26), and details are left to the patient reader for sake of conciseness. $\square$

We now recall, as in [27], that the integrability of $\overline{Y}$ implies the ergodicity of equation (3):

Proposition 2.18. *Assume Hypothesis 1.3 holds true. Then for any $\theta \in \Theta$ and any $f \in \mathcal{C}^1(\mathbb{R}^d; \mathbb{R})$ such that*

$$|f(x)| + |\partial_x f(x)| \leq c\left(1 + |x|^N\right), \qquad x \in \mathbb{R}^d,$$

for some $c > 0$, $N \in \mathbb{N}$, we have

$$\lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau f(Y_t)\, dt = \mathbf{E}[f(\overline{Y})], \qquad \mathbf{P}\text{-}a.s. \tag{31}$$

3. Proof of Theorem 1.7

We can now gather the information we have obtained on the system (3) in order to complete the proof of Theorem 1.7. This will be divided in different steps.

Step 1. The first step of the proof consists in applying Girsanov's theorem. Namely, fix $\theta \in \Theta$, and set $\theta_\tau = \theta + \tau^{-1/2} u$. Then Proposition 2.10 implies that

$$\begin{aligned}\log\left(\frac{d\mathbf{P}^\tau_{\theta_\tau}}{d\mathbf{P}^\tau_\theta}\right) = &-\int_0^\tau \langle \sigma^{-1}([D_+^{H-1/2} b(Y;\theta_\tau)]_t - [D_+^{H-1/2} b(Y;\theta)]_t), dW_t\rangle \\ &- \frac{1}{2}\int_0^\tau |\sigma^{-1}([D_+^{H-1/2} b(Y;\theta_\tau)]_t - [D_+^{H-1/2} b(Y;\theta)]_t)|^2 dt.\end{aligned} \tag{32}$$

Step 2. We next linearize relation (32). To this aim, recall that $\hat{b}$ is defined at Hypothesis 1.3 as $\hat{b} = \partial_\theta b$. We add and substract the d-dimensional vector

$$[D_+^{H-1/2} \hat{b}(Y;\theta)]_t(\theta_\tau - \theta) = \frac{1}{\sqrt{\tau}}[D_+^{H-1/2} \hat{b}(Y;\theta)]_t u,$$

where we abbreviate $\langle [D_+^{H-1/2}\hat{b}(Y;\theta)]_t,\, u\rangle_{\mathbb{R}^q}$ into $[D_+^{H-1/2}\hat{b}(Y;\theta)]_t u$ for notational sake. This easily yields:

$$\log\left(\frac{d\mathbf{P}^\tau_{\theta_\tau}}{d\mathbf{P}^\tau_\theta}\right) = I_1 - I_2 - \frac{1}{2} I_3 - I_4, \tag{33}$$

where

$$\begin{aligned}
I_1 &= \frac{1}{\sqrt{\tau}}\int_0^\tau \langle \sigma^{-1}[D_+^{H-1/2}\hat{b}(Y;\theta)]_t u, dW_t\rangle - \frac{1}{2\tau}\int_0^\tau |\sigma^{-1}[D_+^{H-1/2}\hat{b}(Y;\theta)]_t u|^2 dt \\
I_2 &= \int_0^\tau \langle \sigma^{-1}([D_+^{H-1/2} b(Y;\theta_\tau)]_t - [D_+^{H-1/2} b(Y;\theta)]_t - [D_+^{H-1/2}\hat{b}(Y;\theta)]_t(\theta_\tau - \theta)), dW_t\rangle \\
I_3 &= \int_0^\tau |\sigma^{-1}([D_+^{H-1/2} b(Y;\theta_\tau)]_t - [D_+^{H-1/2} b(Y;\theta)]_t - [D_+^{H-1/2}\hat{b}(Y;\theta)]_t(\theta_\tau - \theta))|^2 dt
\end{aligned}$$

and

$$\begin{aligned}
I_4 = \int_0^\tau \langle \sigma^{-1}(&[D_+^{H-1/2} b(Y;\theta_\tau)]_t - [D_+^{H-1/2} b(Y;\theta)]_t \\
&- [D_+^{H-1/2}\hat{b}(Y;\theta)]_t(\theta_\tau - \theta)), \sigma^{-1}[D_+^{H-1/2}\hat{b}(Y;\theta)]_t(\theta_\tau - \theta)\rangle dt.
\end{aligned}$$

Step 3. In this step we set the ground for the identification of the main contribution to our log-likelihood. That is, we wish to show that as $\tau \to \infty$ we have:

$$\frac{1}{\tau}\int_0^\tau |\sigma^{-1}[D_+^{H-1/2}\hat{b}(Y;\theta)]_t u|^2 dt \xrightarrow{\mathbf{P}_\theta} u^{\mathrm{T}}\Gamma(\theta) u, \tag{34}$$

where $\Gamma(\theta)$ is given by (6). Observe that (34) together with the multivariate central limit theorem for Brownian martingales (cf. [22, Theorem 1.21]), imply that I_1 is the term that contributes to the limit in (5). In Steps 7 and 8 below we will show that I_2, I_3 and I_4 are negligible terms with respect to this main contribution.

In order to prove (34), according to the expression of $D_+^{H-1/2}$, we thus have to find an equivalent for the quantity

$$J_\tau(Y) = \int_0^\tau \left|\sigma^{-1} N_t(Y)\right|^2 dt, \quad \text{with} \quad N_t(Y) = \int_{\mathbb{R}_+} \frac{(\hat{b}(Y_t;\theta) - \hat{b}(Y_{t-r};\theta))u}{r^{H+1/2}}\, dr. \tag{35}$$

Furthermore, in order to define $N_t(Y)$ properly, we have to extend the definition of Y to $\mathbb{R}_-$ by setting $Y_t = y_0$ for all $t \le 0$. Next we decompose $N_t(Y)$ into $N_{1,t}(Y) + N_{2,t}(Y)$, where:

$$N_{1,t}(Y) = \int_0^t \frac{(\hat{b}(Y_t;\theta) - \hat{b}(Y_{t-r};\theta))u}{r^{H+1/2}}\, dr, \quad N_{2,t}(Y) = \int_t^\infty \frac{(\hat{b}(Y_t;\theta) - \hat{b}(y_0;\theta))u}{r^{H+1/2}}\, dr. \tag{36}$$

We shall bound those terms separately.

Step 4. This step is devoted to reduce our computations to an evaluation for the expected value. Specifically, we shall prove that:

$$\lim_{\tau\to\infty} \frac{1}{\tau} \mathbf{E}_\theta\left[|J_\tau - \mathbf{E}_\theta[J_\tau]|\right] = 0, \tag{37}$$

where we write J_τ instead of $J_\tau(Y)$ in order to alleviate notations for the remainder of this step. Observe that owing to the Poincaré type inequality (16), relation (37) can be reduced to show that:

$$\lim_{\tau\to\infty} \frac{\mathbf{E}_\theta\left[\|DJ_\tau\|_{\bar{\mathcal{H}}}\right]}{\tau} = 0, \tag{38}$$

which is what we proceed to do now.

Indeed, by the chain rule we have:

$$DJ_\tau = 2\int_0^\tau \langle \sigma^{-1} N_t, D\left(\sigma^{-1} N_t\right)\rangle\, dt,$$

where we have used the notation N_t for $N_t(Y)$. Therefore, by Cauchy-Schwarz' inequality, we obtain:

$$\|DJ_\tau\|_{\bar{\mathcal{H}}} \lesssim \int_0^\tau |\sigma^{-1} N_t|\, \|D\left(\sigma^{-1} N_t\right)\|_{\bar{\mathcal{H}}}\, dt.$$

Thus, another use of Cauchy-Schwarz' inequality yields:

$$\frac{\mathbf{E}_\theta\left[\|DJ_\tau\|_{\bar{\mathcal{H}}}\right]}{\tau} \lesssim \frac{\mathbf{E}_\theta^{1/2}\left[J_\tau\right]\, \mathbf{E}_\theta^{1/2}\left[\|\,\|D\left(\sigma^{-1} N_\cdot\right)\|_{\bar{\mathcal{H}}}\|^2_{L^2([0,\tau])}\right]}{\tau}.$$

We shall prove (see the forthcoming inequality (46)) that $\mathbf{E}_\theta[J_\tau]$ is of order τ. Hence:

$$\begin{aligned}
\frac{\mathbf{E}_\theta\left[\|DJ_\tau\|_{\bar{\mathcal{H}}}\right]}{\tau} &\lesssim \frac{\mathbf{E}_\theta^{1/2}\left[\|\,\|\sigma^{-1} D\left(N_\cdot\right)\|_{\bar{\mathcal{H}}}\|^2_{L^2([0,\tau])}\right]}{\tau^{1/2}} \\
&\lesssim \frac{\mathbf{E}_\theta^{1/2}\left[\|\,\|D\left(N_\cdot\right)\|_{\bar{\mathcal{H}}}\|^2_{L^2([0,\tau])}\right]}{\tau^{1/2}}.
\end{aligned} \tag{39}$$

Let us now compute DN_t. We can decompose N_t as in (36) and recall that y_0 is considered as a constant, in order to get $DN_t = DN_{1,t} + DN_{2,t}$, with:

$$\begin{aligned} DN_{1,t} &= \int_0^t \frac{(\langle \partial_x \hat{b}(Y_t;\theta),\, DY_t\rangle - \langle \partial_x \hat{b}(Y_{t-r};\theta), DY_{t-r}\rangle)u}{r^{H+1/2}}\, dr \\ DN_{2,t} &= \int_t^\infty \frac{\langle \partial_x \hat{b}(Y_t;\theta),\, DY_t\rangle u}{r^{H+1/2}}\, dr. \end{aligned} \tag{40}$$

We treat again those two terms separately.

The term $DN_{2,t}$ in (40) can be explicitly computed as $t^{1/2-H}\langle \partial_x \hat{b}(Y_t;\theta), DY_t\rangle u$. Thus, according to Hypothesis 1.3(iii), we get:

$$\mathbf{E}_\theta^{1/2}\left[\| \, \|D\left(N_{2,\cdot}\right)\|_{\bar{\mathcal{H}}}\|_{L^2([0,\tau])}^2\right] \lesssim \mathbf{E}_\theta^{1/2}\left[\int_0^\tau \|DY_t\|_{\bar{\mathcal{H}}}^2\, t^{1-2H} dt\right].$$

By Corollary 2.15, $\|DY_t\|_{\bar{\mathcal{H}}}$ is uniformly bounded in $t > 0$. This yields:

$$\frac{\mathbf{E}_\theta^{1/2}\left[\| \, \|D\left(N_{2,\cdot}\right)\|_{\bar{\mathcal{H}}}\|_{L^2([0,\tau])}^2\right]}{\tau^{1/2}} \lesssim \frac{\tau^{1-H}}{\tau^{1/2}} = \frac{1}{\tau^{H-1/2}},$$

which tends to zero as $\tau \to \infty$ since $H > \frac{1}{2}$.

We next treat the term $DN_{1,t}$ in (40). Observe that it can be expressed as:

$$\int_0^t \frac{(\langle \partial_x \hat{b}(Y_t;\theta), DY_t - DY_{t-r}\rangle - \langle \partial_x \hat{b}(Y_{t-r};\theta) - \partial_x \hat{b}(Y_t;\theta), DY_{t-r}\rangle)u}{r^{H+1/2}} dr.$$

Therefore, by the mean value theorem and Hypothesis 1.3(iii), we get that

$$\|DN_{1,t}\|_{\bar{\mathcal{H}}} \lesssim \int_0^t \frac{\|DY_t - DY_{t-r}\|_{\bar{\mathcal{H}}}}{r^{H+1/2}} dr + \int_0^t \frac{|Y_{t-r} - Y_t| \, \|DY_{t-r}\|_{\bar{\mathcal{H}}}}{r^{H+1/2}} dr.$$

We now appeal to Proposition 2.8 and Corollary 2.15, to obtain that for all $\varepsilon \in (0, H)$,

$$\begin{aligned} &\| \, \|D\left(N_{1,\cdot}\right)\|_{\bar{\mathcal{H}}}\|_{L^2([0,\tau])}^2 \\ &\lesssim \int_0^\tau dt \left|\int_0^t \frac{\exp\left(-c(t-r)\right)}{r^{(H+1)/2}}\, dr\right|^2 + Z_\varepsilon^2 \int_0^\tau dt\, (1+t)^{4\varepsilon} \left|\int_0^t \frac{\exp\left(-c(t-r)\right)}{r^{1/2+\varepsilon}}\, dr\right|^2, \end{aligned} \tag{41}$$

where Z_ε is a random variable with finite moments of all orders. The integrals in the right hand side above can be bounded as follows: for $\beta \in (0,1)$ and a generic constant $c > 0$ which can change from line to line, observe that:

$$\begin{aligned} I_{\beta,t} \equiv \int_0^t \frac{\exp\left(-c(t-r)\right)}{r^\beta}\, dr &= \int_0^{t/2} \frac{\exp\left(-c(t-r)\right)}{r^\beta}\, dr + \int_{t/2}^t \frac{\exp\left(-c(t-r)\right)}{r^\beta}\, dr \\ &\lesssim \exp\left(-c\alpha t\right) t^{1-\beta} + t^{-\beta} \lesssim t^{-\beta}, \end{aligned}$$

uniformly in t. This bound is useful for $t \geq 1$. When $t \leq 1$, we can bound $I_{\beta,t}$ by a constant c. Therefore, $I_{\beta,t}$ can be uniformly bounded by $c(1+t)^{-\beta}$. Reporting this information into (41), we obtain that for $\tau \geq 1$,

$$\| \, \|DN_{1,\cdot}\|_{\bar{\mathcal{H}}}\|_{L^2([0,\tau])}^2 \lesssim \int_0^\tau \frac{1}{(1+t)^{H+1}}\, dt + Z_\varepsilon^2 \int_0^\tau \frac{(1+t)^{4\varepsilon}}{(1+t)^{1+2\varepsilon}}\, dt \lesssim Z_\varepsilon^2 \tau^{2\varepsilon}.$$

Thus, for $\tau \geq 1$,

$$\frac{\mathbf{E}_\theta^{1/2}\left[\| \, \|DN_{1,\cdot}\|_{\bar{\mathcal{H}}}\|^2_{L^2([0,\tau])}\right]}{\tau^{1/2}} \lesssim \tau^{\varepsilon-1/2},$$

which tends to zero as $\tau \to \infty$ as long as we choose $\varepsilon < \frac{1}{2}$. Gathering our bounds on DN_1, DN_2 and plugging them into (39), the proof of (38) is now completed. As a consequence of this inequality together with the Poincaré type bound (16), we get that:

$$\lim_{\tau\to\infty} \mathbf{E}_\theta\left[\left|\frac{J_\tau - \mathbf{E}_\theta[J_\tau]}{\tau}\right|\right] = 0.$$

This allows us to be reduced to an evaluation of $\mathbf{E}_\theta[J_\tau]$ only in the sequel.

Step 5: We now show that one can approximate the quantities related to Y by quantities related to the stationary solution $\overline{Y}$. Namely, we denote by $J_\tau(\overline{Y})$ the same quantity as $J_\tau(Y)$ (see (35)), with Y replaced by $\overline{Y}$, and we set $N_t(\overline{Y})$, $N_{1,t}(\overline{Y})$, $N_{2,t}(\overline{Y})$ accordingly. Observe that, with (34) in mind, we have to compare $N_t(Y)$ with $N_t(\overline{Y})$ up to some $\mathrm{O}(1)$ terms. We will achieve this aim thanks to our decomposition (36).

Step 5a. The term $N_{2,t}(Y)$ is easily estimated: it is readily checked that

$$N_{2,t}(Y) = \frac{c_H(\hat{b}(Y_t;\theta) - \hat{b}(y_0;\theta))u}{t^{H-1/2}}.$$

Invoking Proposition 2.7 as well as the linear growth for $\hat{b}$, we get $\lim_{t\to\infty} N_{2,t}(Y) = 0$, almost surely and in any $L^p(\Omega)$ for $p \geq 1$.

Step 5b. Let us compare $N_{1,t}(\overline{Y})$ and $N_{1,t}(Y)$. One can write:

$$N_{1,t}(\overline{Y}) - N_{1,t}(Y) = \int_0^t \frac{\delta\big[\hat{b}(Y;\theta) - \hat{b}(\overline{Y};\theta)\big]_{t-r,t} u}{r^{H+1/2}}\, dr \equiv N_{11,t} + N_{12,t},$$

where we have set:

$$N_{11,t} = \int_0^{t/2} \frac{\delta\big[\hat{b}(Y;\theta) - \hat{b}(\overline{Y};\theta)\big]_{t-r,t} u}{r^{H+1/2}}\, dr, \qquad N_{12,t} = \int_{t/2}^{t} \frac{\delta\big[\hat{b}(Y;\theta) - \hat{b}(\overline{Y};\theta)\big]_{t-r,t} u}{r^{H+1/2}}\, dr.$$

We now show that those two terms give raise to a negligible (that is $\mathrm{o}(1)$ in terms of t) contribution.

The term $N_{11,t}$ involves some increments of $Y - \overline{Y}$, and should thus be handled thanks to Proposition 2.17. However, due to the composition with the nonlinearity $\hat{b}$, we have to resort to a special treatment concerning rectangular increments. To this aim, let us define a path $a : [0,1]^2 \to \mathbb{R}$ of the following form:

$$a(\lambda,\mu) = Y_{t-r} + \lambda(Y_t - Y_{t-r}) + \mu(\overline{Y}_{t-r} - Y_{t-r}) + \lambda\mu\,\Delta\overline{Y}_{t,r},$$

where we have set $\Delta\overline{Y}_{t-r,t} = \overline{Y}_t - \overline{Y}_{t-r} - Y_t + Y_{t-r}$. We also write A for the composition of a with the non linearity $\hat{b}$, that is $A(\lambda,\mu) = \hat{b}(a(\lambda,\mu))$. It is then readily checked that:

$$\begin{aligned}\hat{b}(\overline{Y}_t) - \hat{b}(\overline{Y}_{t-r}) - \hat{b}(Y_t) + \hat{b}(Y_{t-r}) &= A(1,1) - A(1,0) - A(0,1) + A(0,0) \\ &= \int_{[0,1]^2} \partial^2_{\lambda\mu} A(\lambda,\mu)\, d\lambda d\mu. \end{aligned} \tag{42}$$

Furthermore, thanks to the fact that the first and second order derivative of $\hat{b}$ are bounded (see Hypothesis 1.3), we get that:

$$\begin{aligned}
\left|\partial^2_{\lambda\mu} A(\lambda,\mu)\right| &\lesssim |\partial_\lambda a(\lambda,\mu)\,\partial_\mu a(\lambda,\mu)| + \left|\partial^2_{\lambda\mu} a(\lambda,\mu)\right| \\
&\lesssim \left(|\delta Y_{t-r,t}| + \left|\Delta\overline{Y}_{t-r,t}\right|\right)\left(\left|\overline{Y}_t - Y_t\right| + \left|\Delta\overline{Y}_{t-r,t}\right|\right) + \left|\Delta\overline{Y}_{t-r,t}\right| \\
&\lesssim \left(|\delta Y_{t-r,t}| + \left|\Delta\overline{Y}_{t-r,t}\right|\right)\left(\left|\overline{Y}_t - Y_t\right| + \left|\overline{Y}_{t-r} - Y_{t-r}\right|\right) + \left|\Delta\overline{Y}_{t-r,t}\right|
\end{aligned}$$

We now bound $\delta Y_{t-r,t}$ thanks to Proposition 2.8 and $\overline{Y}_t - Y_t$, $\overline{Y}_{t-r} - Y_{t-r}$ and $\Delta\overline{Y}_{t-r,t}$ according to Proposition 2.17. Gathering all those estimates we get that for all $\alpha < H$ and $\varepsilon < H$, the following holds true:

$$\left|\partial^2_{\lambda\mu} A(\lambda,\mu)\right| \lesssim \left(Z\,(1+t)^{2\varepsilon}\, r^{H-\varepsilon} + Z\, e^{-c(t-r)} r^{\alpha}\right) Z\, e^{-c(t-r)} + Z e^{-c(t-r)} r^{\alpha}. \tag{43}$$

Therefore, reporting (43) in (42) and then in the definition of $N_{11,t}$, and taking $\alpha > H - \frac{1}{2}$ and $\varepsilon < \frac{1}{2}$, we get that:

$$\begin{aligned}
N_{11,t} &\lesssim Z \int_0^{t/2} \frac{e^{-c(t-r)}}{r^{H+1/2-\alpha}}\, dr + Z(1+t)^{2\varepsilon} \int_0^{t/2} \frac{e^{-c(t-r)}}{r^{1/2+\varepsilon}}\, dr \\
&\lesssim Z e^{-ct} t^{1/2-H+\alpha} + Z(1+t)^{2\varepsilon} e^{-ct} t^{1/2-\varepsilon}.
\end{aligned}$$

It is now readily checked that $\lim_{t\to\infty} N_{11,t} = 0$ almost surely with uniform $L^p(\Omega)$ bounds, and thus yields a negligible contribution.

The term $N_{12,t}$ is bounded in the same way for $t \geq 1$: choosing $\alpha > H - 1/2$ and $\varepsilon < \frac{1}{2}$, we have:

$$\begin{aligned}
N_{12,t} &\lesssim Z \int_{t/2}^{t} \frac{e^{-c(t-r)}}{r^{H+1/2-\alpha}}\, dr + Z(1+t)^{2\varepsilon} \int_{t/2}^{t} \frac{e^{-c(t-r)}}{r^{1/2+\varepsilon}}\, dr \\
&\lesssim \frac{Z}{t^{H+1/2-\alpha}} \int_{t/2}^{t} e^{-c(t-r)}\, dr + \frac{Z(1+t)^{2\varepsilon}}{t^{1/2+\varepsilon}} \int_{t/2}^{t} e^{-c(t-r)}\, dr \lesssim \frac{Z}{t^{H+1/2-\alpha}} + \frac{Z}{t^{1/2-\varepsilon}},
\end{aligned}$$

which gives another negligible contribution.

Step 5c. We should also bound the term $N_{2,t}(\overline{Y})$. Notice that $\overline{Y}$ is not constant on $\mathbb{R}_-$, so that:

$$N_{2,t}(\overline{Y}) = \int_t^\infty \frac{(\hat{b}(\overline{Y}_t;\theta) - \hat{b}(\overline{Y}_{t-r};\theta))u}{r^{H+1/2}}\, dr.$$

We shall simply bound this quantity as:

$$|N_{2,t}(\overline{Y})| \leq |u| \int_t^\infty \frac{|\hat{b}(\overline{Y}_t;\theta)| + |\hat{b}(\overline{Y}_{t-r};\theta)|}{r^{H+1/2}}\, dr. \tag{44}$$

In addition, Proposition 2.8 can be shown for $\overline{Y}$ exactly in the same way as for Y. Therefore, for any $\varepsilon > 0$ we get:

$$|\hat{b}(\overline{Y}_t;\theta)| + |\hat{b}(\overline{Y}_{t-r};\theta)| \leq Z_\varepsilon \left(1 + |t|^{2\varepsilon} + |r|^{2\varepsilon}\right),$$

where Z_ε is a random variable admitting moments of all order. Plugging this bound into (44), we get a bound of the form $|N_{2,t}(\overline{Y})| \leq Z_\varepsilon t^{-(H-1/2-2\varepsilon)}$ for an arbitrarily small constant ε.

Summarizing Steps *5a* to *5c*, we have found that $N_t(Y) - N_t(\overline{Y})$ is of order $t^{-\eta} Z$ with $\eta > 0$ and $Z \in \cap_{p\geq 1} L^p(\Omega)$. Hence $J_\tau(Y) - J_\tau(\overline{Y})$ is of order $\tau^{1-2\eta}$, which is a negligible term on the scale τ. This allows us to consider the limiting behavior of $J_\tau(\overline{Y})$ instead of $J_\tau(Y)$, which is what we do in the sequel of the proof.

Step 6. We are now reduced to the analysis of the quantity $\mathbf{E}_\theta[J_\tau(\overline{Y})]$, which is equal to $u^{\mathrm{T}}\Psi u$ where:

$$\Psi = \int_0^\tau dt \int_{\mathbb{R}_+^2} \frac{\mathbf{E}_\theta[(\hat{b}(\overline{Y}_t;\theta) - \hat{b}(\overline{Y}_{t-r_1};\theta))^{\mathrm{T}}(\sigma^{-1})^{\mathrm{T}}\sigma^{-1}(\hat{b}(\overline{Y}_t;\theta) - \hat{b}(\overline{Y}_{t-r_2};\theta))]}{r_1^{1/2+H} r_2^{1/2+H}} \, dr_1 dr_2. \quad (45)$$

By stationarity of $\overline{Y}$, the expected value does not depend on t:

$$\begin{aligned}&\mathbf{E}_\theta[(\hat{b}(\overline{Y}_t;\theta) - \hat{b}(\overline{Y}_{t-r_1};\theta))^{\mathrm{T}}(\sigma^{-1})^{\mathrm{T}}\sigma^{-1}(\hat{b}(\overline{Y}_t;\theta) - \hat{b}(\overline{Y}_{t-r_2};\theta))] \\ &\quad = \mathbf{E}_\theta[(\hat{b}(\overline{Y}_0;\theta) - \hat{b}(\overline{Y}_{r_1};\theta))^{\mathrm{T}}(\sigma^{-1})^{\mathrm{T}}\sigma^{-1}(\hat{b}(\overline{Y}_0;\theta) - \hat{b}(\overline{Y}_{r_2};\theta))].\end{aligned}$$

Furthermore, we have already observed that Proposition 2.8 is also satisfied by $\overline{Y}$. Hence, thanks to the fact that $\hat{b}$ is a Lipschitz function in x, the following bound holds true:

$$|\mathbf{E}_\theta[(\hat{b}(\overline{Y}_0;\theta) - \hat{b}(\overline{Y}_{r_1};\theta))^{\mathrm{T}}(\sigma^{-1})^{\mathrm{T}}\sigma^{-1}(\hat{b}(\overline{Y}_0;\theta) - \hat{b}(\overline{Y}_{r_2};\theta))]| \lesssim \left(r_1^H \wedge 1\right)\left(r_2^H \wedge 1\right).$$

Plugging this information into (45), we obtain that Ψ is a convergent integral and:

$$\mathbf{E}_\theta[J_\tau(\overline{Y})] = \tau u^{\mathrm{T}}\Gamma(\theta) u, \quad (46)$$

where $\Gamma(\theta)$ is the convergent integral given by (6). In particular, we have:

$$\lim_{\tau\to\infty} \frac{\mathbf{E}_\theta[J_\tau(\overline{Y})]}{\tau} = u^{\mathrm{T}}\Gamma(\theta)u.$$

Summarizing our considerations up to now, we have proved (34). Hence I_1 in (33) converges to the term $u^{\mathrm{T}}\mathcal{N}(0,\Gamma(\theta)) - \frac{1}{2}u^{\mathrm{T}}\Gamma(\theta)u$ in (5).

Step 7. We next show that the term I_3 in (33) converges to zero in $\mathbf{P}_\theta$-probability as $\tau\to\infty$. For this, set $g_t(\theta) = \sigma^{-1}[D_+^{H-1/2} b(Y;\theta)]_t$. One can recast I_3 into:

$$I_3 = \int_0^\tau \left|g_t(\theta_\tau) - g_t(\theta) - \partial_\theta g_t(\theta)(\theta_\tau - \theta)\right|^2 \, dt,$$

and we also recall that $\theta_\tau - \theta = \tau^{-1/2}u$. In addition, a simple application of Taylor's expansion for multivariate function yields the existence of a $\lambda \in (0,1)$ such that:

$$g_t(\theta_\tau) - g_t(\theta) = \partial_\theta g_t(\xi_\lambda)(\theta_\tau - \theta), \quad \text{where} \quad \xi_\lambda = \theta + \lambda(\theta_\tau - \theta).$$

Notice that under our standing Hypothesis 1.3 we have $\partial_\theta g_t(\xi_\lambda) = \sigma^{-1}[D_+^{H-1/2}\hat{b}(Y;\xi_\lambda)]_t$. We thus get:

$$I_3 = \frac{1}{\tau}\int_0^\tau |\sigma^{-1}M_t(Y)u|^2 \, dt, \quad \text{where} \quad M_t(Y) = \int_{\mathbb{R}_+} \frac{\delta[\hat{b}(Y;\xi_\lambda) - \hat{b}(Y;\theta)]_{t-r,t}}{r^{1/2+H}} \, dr.$$

Next we decompose $M_t(Y)$ into $M_{1,t}(Y) + M_{2,t}(Y)$, where

$$M_{1,t}(Y) = \int_0^t \frac{\delta[\hat{b}(Y;\xi_\lambda) - \hat{b}(Y;\theta)]_{t-r,t}}{r^{1/2+H}} \, dr, \qquad M_{2,t}(Y) = \int_t^\infty \frac{\delta[\hat{b}(Y;\xi_\lambda) - \hat{b}(Y;\theta)]_{t-r,t}}{r^{1/2+H}} \, dr.$$

Recall that we have extended the definition of Y to $\mathbb{R}_-$ by setting $Y_t = y_0$ for all $t \leq 0$. Thus, using Hypothesis 1.3, we obtain that

$$\tau^{-1}\int_0^\tau |\sigma^{-1}M_{2,t}(Y)u|^2\,dt \lesssim \tau^{-2}\int_0^\tau \frac{(1+|Y_t|^2)}{t^{1-2H}}dt.$$

Thus, by Proposition 2.7, the $L^1(\Omega)$-norm of this term is bounded by $c_H\tau^{-2H}$. Hence, this term converges in $\mathbf{P}_\theta$-probability to zero as $\tau\to\infty$.

On the other hand, fix $\alpha\in(\frac{1}{1+2H},\frac{1}{2H})$, and write

$$M_{1,t}(Y) = \int_0^t \frac{(\delta[\hat{b}(Y;\xi_\lambda)-\hat{b}(Y;\theta)]_{t-r,t})^\alpha(\delta[\hat{b}(Y;\xi_\lambda)-\hat{b}(Y;\theta)]_{t-r,t})^{1-\alpha}}{r^{1/2+H}}\,dr.$$

Then, appealing to Hypothesis 1.3, we get that

$$|M_{1,t}(Y)| \lesssim \frac{1}{\tau^{\alpha/2}}\int_0^t \frac{(1+|Y_t|^\alpha+|Y_{t-r}|^\alpha)\,|Y_t-Y_{t-r}|^{1-\alpha}}{r^{H(1-\alpha)}r^{1/2+\alpha H}}\,dr$$

Therefore,

$$|M_{1,t}(Y)|^2$$
$$\lesssim \frac{1}{\tau^\alpha}\int_0^t\int_0^t \frac{(1+|Y_t|^\alpha+|Y_{t-r_1}|^\alpha)\,|Y_t-Y_{t-r_1}|^{1-\alpha}}{r_1^{H(1-\alpha)}r_1^{1/2+\alpha H}}\frac{(1+|Y_t|^\alpha+|Y_{t-r_2}|^\alpha)\,|Y_t-Y_{t-r_2}|^{1-\alpha}}{r_2^{H(1-\alpha)}r_2^{1/2+\alpha H}}\,dr_1dr_2.$$

Thus, again by Proposition 2.7, the $L^1(\Omega)$-norm of the term $\frac{1}{\tau}\int_0^\tau|\sigma^{-1}M_{1,t}(Y)u|^2dt$ is bounded by

$$\frac{c_H}{\tau^{1+\alpha}}\int_0^\tau\int_0^t\int_0^t \frac{1}{r_1^{1/2+\alpha H}}\frac{1}{r_2^{1/2+\alpha H}}\,dr_1dr_2dt = \frac{c_H}{\tau^{1+\alpha}}\int_0^\tau t^{1-2H\alpha}dt = \frac{c_H}{\tau^{\alpha(2H+1)-1}},$$

which converges to zero as $\tau\to\infty$ since $\alpha>\frac{1}{2H+1}$.

Step 8. We finally show that I_2 and I_4 are also negligible terms. Since I_3 is the quadratic variation of the martingale I_2, this implies that I_2 converges to zero in $\mathbf{P}_\theta$-probability as $\tau\to\infty$. On the other hand, applying Cauchy-Schwarz inequality to I_4, we get that

$$|I_4| \leq \left(\frac{1}{\tau}\int_0^\tau |\sigma^{-1}[D_+^{H-1/2}\hat{b}(Y;\theta)]_t u|^2dt\right)^{1/2}\times I_3^{1/2},$$

Thus, by the results in Steps 3 and 4, we obtain that I_4 converges to zero in $\mathbf{P}_\theta$-probability as $\tau\to\infty$, which completes the proof of the theorem.

References

[1] L. Arnold (1997): *Random Dynamical Systems.* Berlin, Springer.

[2] R. Belfadli, K. Es-Sebaiy, Y. Ouknine (2011): Parameter Estimation for Fractional Ornstein-Uhlenbeck Processes: Non-ergodic Case. In: *Frontiers in Science and Engineering.* Edited by Hassan II Academy of Science and Technology. **1**, 1–16.

[3] B. Bercu, L. Coutin, N. Savy (2008): Sharp Large Deviations for the Fractional Ornstein-Uhlenbeck Process. *Theory Probab. Appl.* **55**, 575–610.

[4] A. Brouste, S. M. Iacus (2013): Parameter estimation for the discretely observed fractional Ornstein—Uhlenbeck process and the Yuima R package. *Journal Computational Statistics* **28**, 1529-1547.

[5] H. Crauel, A. Debussche, F. Flandoli (1997): Random attractors. *J. Dyn. Diff. Eqns.* **9**, 307–341.

[6] A. Chronopoulou, S. Tindel (2013): On inference for fractional differential equations. *Stat. Infer. Stoch. Process.* **16**, no. 1, 29–61.

[7] S. Cohen, F. Gamboa, C. Lacaux, J.-M. Loubes (2013): LAN property for some fractional type Brownian motion. *ALEA* **10**, 91–106.

[8] L. Deucreusefond (2003): Stochastic integration with respect to fractional Brownian motion. *Theory and applications of long-range dependence*, 203-226, Birkhäuser.

[9] A. Friedman (1975): *Stochastic Differential Equations and Applications.* Academic Press, New York.

[10] P. Friz, N. Victoir (2010): *Multidimensional dimensional processes seen as rough paths.* Cambridge University Press.

[11] E. Gobet (2002): LAN property for ergodic diffusions with discrete observations, *Ann. I. H. Poincaré*, **38**, 711-737.

[12] M. Garrido-Atienza, P. Kloeden, A. Neuenkirch (2009): Discretization of stationary solutions of stochastic systems driven by fractional Brownian motion. *Appl. Math. Optim.* **60**, no. 2, 151–172.

[13] A.M. Garcia, E. Rodemich, H. Rumsey Jr. (1978): A real variable lemma and the continuity of paths of some Gaussian processes. *Indiana Math. J.* **20**, 565–578.

[14] M. Garrido-Atienza, B. Schmalfuss (2011): Ergodicity of the infinite-dimensional fractional Brownian motion. *J. Dynam. Differential Equations.* **23**, no. 3, 671–681.

[15] M. Hairer (2005): Ergodicity of stochastic differential equations driven by fractional Brownian motion. *Ann. Probab.* **33**, no. 2, 703–758.

[16] Y. Hu, D. Nualart (2010): Parameter estimation for fractional Ornstein-Uhlenbeck processes. *Stat. Prob. Lett.* **80**, 1030–1038.

[17] I. A. Ibragimov and R. Z. Has'minskii, R. Z. (1981): *Statistical estimation - Asymptotic theory*, volume 16 of *Applications of Mathematics*, Springer-Verlag, New York. Translated from the Russian by Samuel Kotz.

[18] M. Kleptsyna, A. Le Breton (2002): Statistical analysis of the fractional Ornstein-Uhlenbeck type process. *Stat. Inference Stoch. Process.* **5**, no. 3, 229–248.

[19] P. Kloeden, A. Neuenkirch (2007): The pathwise convergence of approximation schemes for stochastic differential equations, *LMS J. Comp. Math.* **10** , 235–253.

[20] S. Kou, X. Sunney-Xie (2004): Generalized Langevin equation with fractional Gaussian noise: subdiffusion within a single protein molecule. *Phys. Rev. Lett.* **93**, no. 18.

[21] Y. Kozachenko, A. Melnikov, Y. Mishura (2004): On drift parameter estimation in models with fractional Brownian motion. *Statistics* **49**, no. 1, 35-62.

[22] Y. A. Kutoyants, Y.A. (2004): *Statistical Inference for Ergodic Diffusion Processes*, Springer-Verlag.

[23] A. Le Breton (1998): Filtering and parameter estimation in a simple linear system driven by a fractional Brownian motion. *Stat. Probab. Lett.* **38**, no. 3, 263–274.

[24] L. Le Cam (1960): Locally asymptotically normal families of distributions, *Univ. California, Publ. Statist*, **3**, 37-98.

[25] B. Mandelbrot, J. Van Ness (1968): Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **10**, 422–437.

[26] S. Moret, D. Nualart (2002): Onsager-Machlup functional for the fractional Brownian motion. *Probab. Theory Relat. Fields* **124**, 227–260.

[27] A. Neuenkirch, S. Tindel (2014): A least square-type procedure for parameter estimation in stochastic differential equations with additive fractional noise. *Stat. Infer. Stoch. Process.* **17**, no. 1, 99–120.

[28] D. Nualart (2006): *The Malliavin Calculus and Related Topics.* Probability and its Applications. Springer-Verlag, 2nd Edition.

[29] D. Nualart, Y. Ouknine (2002): Regularization of differential equations by fractional noise. *Stochastic Processes and Their Applications* **102**, 103–116.

[30] B. L. S. Prakasa Rao (2010): Statistical inference for fractional diffusion processes. Wiley Series in Probability and Statistics, Chichester, John Wiley & Sons.

[31] S. Riedel (2014): Talagrand's transportation-cost inequality and applications to (rough) path spaces. Preprint.

[32] A.N. Shiryaev (1995): *Probability.* 2nd ed. Graduate Texts in Mathematics. New York, Springer.

[33] C. Tudor, F. Viens (2007): Statistical aspects of the fractional stochastic calculus. *Ann. Statist.* **35**, no. 3, 1183–1212.

[34] A.S. Üstünel (1995): *An introduction to analysis on Wiener space.* Lecture Notes in Mathematics **1610**. Springer-Verlag.

[35] W. Xiao, W. Zhang, W. Xu. (2011): Parameter estimation for fractional Ornstein-Uhlenbeck processes at discrete observation. *Applied Mathematical Modelling* **35**, 4196–4207.

(Eulalia Nualart) Department of Economics and Business, Universitat Pompeu Fabra and Barcelona Graduate School of Economics, Ramón Trias Fargas 25-27, 08005 Barcelona, Spain

E-mail address: eulalia@nualart.es

URL: http://nualart.es

(Samy Tindel) Department of Mathematics, Purdue University, 150 N. University Street, W. Lafayette, IN 47907, USA.

E-mail address: stindel@purdue.edu

URL: https://www.math.purdue.edu/~stindel/