A LEAST SQUARE-TYPE PROCEDURE FOR PARAMETER ESTIMATION IN STOCHASTIC DIFFERENTIAL EQUATIONS WITH ADDITIVE FRACTIONAL NOISE

ANDREAS NEUENKIRCH AND SAMY TINDEL

Abstract. We study a least square-type estimator for an unknown parameter in the drift coefficient of a stochastic differential equation with additive fractional noise of Hurst parameter $H > 1/2$. The estimator is based on discrete time observations of the stochastic differential equation, and using tools from ergodic theory and stochastic analysis we derive its strong consistency.

1. Introduction

In this article, we will consider the following $\mathbb{R}^d$-valued stochastic differential equation (SDE)

$$Y_t = y_0 + \int_0^t b(Y_s; \vartheta_0)\, ds + \sum_{j=1}^m \sigma_j B_t^j, \qquad t \in [0, T]. \tag{1}$$

Here $y_0 \in \mathbb{R}^d$ is a given initial condition, $B = (B^1, \dots, B^m)$ is an m-dimensional fractional Brownian motion (fBm) with Hurst parameter $H \in (0, 1)$, the unknown parameter ϑ_0 lies in a certain set Θ which will be specified later on, $\{b(\cdot; \vartheta), \vartheta \in \Theta\}$ is a family of drift coefficients with $b(\cdot; \vartheta) : \mathbb{R}^d \to \mathbb{R}$, and $\sigma_1, \dots, \sigma_m \in \mathbb{R}^d$ are assumed to be known diffusion coefficients.

Let us recall that B is a centred Gaussian process defined on a complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Its law is thus characterized by its covariance function, which is defined by

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

The variance of the increments of B is then given by

$$\mathbf{E}\,|B_t^i - B_s^i|^2 = |t-s|^{2H}, \qquad s, t \in \mathbb{R}, \quad i = 1, \dots, m,$$

and this implies that almost surely the fBm paths are γ-Hölder continuous for any $\gamma < H$. Furthermore, for $H = 1/2$, fBm coincides with the usual Brownian motion, converting the family $\{B^H, H \in (0, 1)\}$ into the most natural generalization of this classical process. In the current paper we assume that the Hurst coefficient satisfies $H > 1/2$ and we focus on the estimation of the parameter $\vartheta_0 \in \Theta$. Note that the Hurst parameter and the diffusion coefficients can be estimated via the quadratic variation of Y, see e.g. [1, 4, 15].

Date: November 8, 2011.

2010 *Mathematics Subject Classification.* 60G15; 62M09; 62F12.

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

S. Tindel is member of the BIGS (Biology, Genetics and Statistics) team at INRIA.

Estimators for the unknown parameter in equation (1) based on continuous observation of Y have been studied e.g. in [2, 14, 17, 18, 20, 21, 22]. Estimators based on discrete time data, which are important for practical applications, are then obtained via discretization. However, to the best of our knowledge no genuine estimators based on discrete time data have been analyzed yet.

We propose here a least square estimator for ϑ_0 based on discrete observations of the process Y at times $\{t_k;\, 0 \le k \le n\}$. For simplicity, we shall take equally spaced observation times with $t_{k+1} - t_k = \kappa\, n^{-\alpha} := \alpha_n$ with given $\alpha \in (0,1), \kappa > 0$. We call our method least square-type procedure, insofar as we consider a quadratic statistics of the form

$$Q_n(\vartheta) = \frac{1}{n\alpha_n^2} \sum_{k=0}^{n-1} \left(\left| \delta Y_{t_k t_{k+1}} - b(Y_{t_k};\vartheta)\alpha_n \right|^2 - \|\sigma\|^2 \alpha_n^{2H} \right), \tag{2}$$

where $\delta Y_{u_1 u_2} := Y_{u_2} - Y_{u_1}$ for any $0 \le u_1 \le u_2 \le T$ and $\|\sigma\|^2 = \sum_{j=1}^m |\sigma_j|^2$.

Let us now describe the assumptions under which we shall work, starting from a standard hypothesis on the parameter set Θ:

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

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 $p_1, p_2 \ge 1$. Then for any tuple $(i_1, \dots i_{p_1}) \in \{1, \dots, d\}^{p_1}$, we set $\partial_{i_1 \dots i_{p_1}} f$ for $\frac{\partial^{p_1} f}{\partial x_{i_1} \dots \partial x_{i_{p_1}}}$. Moreover, we will write $\partial_x f$ resp. $\partial_\vartheta f$ for the Jacobi-matrices $(\partial_{x_1} f, \dots, \partial_{x_d} f)$ and $(\partial_{\vartheta_1} f, \dots, \partial_{\vartheta_q} f)$.*

With this notation in mind, our drift coefficients and their derivatives will satisfy a polynomial growth condition, plus an inward condition which is traditional for estimation procedures in the Brownian diffusion case (see e.g [5, 16]):

Hypothesis 1.3. *We have $b \in \mathcal{C}^{1,1}(\mathbb{R}^d \times \Theta; \mathbb{R}^d)$ and there exist constants $c_1, c_2 > 0$ and $N \in \mathbb{N}$ such that:*
(i) *For every $x, y \in \mathbb{R}^d$ and $\vartheta \in \Theta$ we have*

$$\langle b(x;\vartheta) - b(y;\vartheta),\, x - y \rangle \le -c_1 |x-y|^2$$

(ii) *For every $x \in \mathbb{R}^d$ and $\vartheta \in \Theta$ the following growth bounds are satisfied:*

$$|b(x;\vartheta)| \le c_2 \left(1 + |x|^N\right), \quad |\partial_x b(x;\vartheta)| \le c_2 \left(1 + |x|^N\right), \quad |\partial_\vartheta b(x;\vartheta)| \le c_2 \left(1 + |x|^N\right).$$

As a consequence of the above assumptions on the drift coefficient and the initial condition, for given $\vartheta_0 \in \Theta$ the solution of equation (1) converges for $t \to \infty$ to a stationary and ergodic stochastic process $(\overline{Y}_t, t \ge 0)$, see the next section.

Finally, we also assume that our-drift coefficient is of gradient-type, i.e.:

Hypothesis 1.4. *There exists a function $U \in \mathcal{C}^{2,1}(\mathbb{R}^d \times \Theta; \mathbb{R})$ such that*

$$\partial_x U(x;\vartheta) = b(x;\vartheta), \qquad x \in \mathbb{R}^d,\ \vartheta \in \Theta.$$

With those assumptions in mind, we obtain the following convergence result:

Theorem 1.5. *Assume that the Hypotheses 1.1, 1.3 and 1.4 are satisfied for equation (1) and that we moreover have $H > 1/2$. Let $Q_n(\vartheta)$ be defined by (2). Then we have*

$$\sup_{\vartheta\in\Theta} \left| Q_n(\vartheta) - \left(\mathbf{E}\,|b(\overline{Y}_0;\vartheta_0)|^2 - \mathbf{E}|b(\overline{Y}_0;\vartheta)|^2\right)\right| \to 0 \tag{3}$$

in the $\mathbf{P}$*-almost sure sense.*

This convergence is in contrast to the case $H = 1/2$, i.e. to the case of SDEs with additive Brownian noise. There it holds

$$\sup_{\vartheta\in\Theta} \left| Q_n(\vartheta) - \left(\mathbf{E}\,|b(\overline{Y}_0;\vartheta_0) - b(\overline{Y}_0;\vartheta)|^2\right)\right| \to 0 \tag{4}$$

in the $\mathbf{P}$-almost sure sense, see Remark 3.6, and usually the consistent least squares estimator

$$\operatorname{argmin}_{\vartheta\in\Theta} \sum_{k=0}^{n-1} \left|\delta Y_{t_k t_{k+1}} - b(Y_{t_k};\vartheta)\alpha_n\right|^2,$$

is considered, see e.g. [5, 16]. The difference in the limits (3) and (4) is due to the higher smoothness and long-range dependence of fractional Brownian motion for $H > 1/2$. Our estimator can thus be seen as a "zero squares" estimator instead of a classical least square estimator. In order to show its convergence, we shall work under the following natural identifiability assumption:

Hypothesis 1.6. *For any $\vartheta_0 \in \Theta$, we have*

$$\mathbf{E}\,|b(\overline{Y}_0;\vartheta_0)|^2 = \mathbf{E}|b(\overline{Y}_0;\vartheta)|^2 \quad \textit{iff} \quad \vartheta = \vartheta_0.$$

With this additional Hypothesis, the main result of the current article is the consistency of the zero squares estimator based on the statistics Q_n:

Theorem 1.7. *Assume that the Hypotheses 1.1, 1.3, 1.4 and 1.6 are satisfied for equation (1) and let $H > 1/2$. Let $Q_n(\vartheta)$ be defined by (2), and let $\widehat{\vartheta}_n = \operatorname{argmin}_{\vartheta\in\Theta} |Q_n(\vartheta)|$. Then for any $\vartheta_0 \in \Theta$, we have $\lim_{n\to\infty} \widehat{\vartheta}_n = \vartheta_0$ in the* $\mathbf{P}$*-almost sure sense.*

Note that minimizing $|Q_n(\vartheta)|$ is of course equivalent to finding the zero of $Q_n(\vartheta)$. Let us shortly compare Theorem 1.7 with the existing literature on estimation procedures for fBm driven equations:

(i) Most of the previous results, see e.g. [2, 14, 17, 21], deal with the one-dimensional fractional Ornstein-Uhlenbeck process in a continuous observation setting. In particular, for this process simple continuous time least-square estimators are obtained in [2, 14], for which also covergence rates and asymptotic error distributions are derived. Compared to these results our estimation procedure covers a broad class of ergodic multi-dimensional equations and relies on discrete data only.

(ii) A general estimation procedure based on moment matching is established in [20]. However, the main assumption in [20] is that many independent observations of sample paths over a short time interval are available, which is not the case in many practical situations where rather one sample path is discretely observed for a long time period. Let us also mention the article [3], in which a general discrete data maximum likelihood type procedure has been designed for parameter estimation in both the drift and diffusion coefficients, however without proof of consistency.

(iii) Our current work probably compares best with the maximum likelihood estimator analyzed in [22]. The latter pioneering reference focused on one-dimensional SDEs of the form

$$dY_t = \vartheta_0 h(Y_t)\, dt + dB_t$$

with $h : \mathbb{R} \to \mathbb{R}$ satisfying suitable regularity assumptions. Strong consistency is obtained for the continuous time estimator and also for a discretized version of the estimator. However, the discretized estimator involves rather complicated operators related to the kernel functions arising in the Wiener-integral representation of fBm, which are avoided in our approach. Moreover, in contrast to [22] the consistency proof for our estimator does not rely on Malliavin calculus methods.

So, in view of the existing results in the literature, Theorem 1.7 can be seen a step towards simple and implementable parameter estimation procedures for SDEs driven by fBm.

Finally, let us comment on the assumptions we have imposed on the drift coefficient and on the Hurst parameter:

(a) The hypothesis of Theorem 1.5 are standard for the case $H = 1/2$, except Hypothesis 1.4 which restricts us to gradient-type drift coefficients. We require this condition to show an ergodic-type result for weighted sums of the increments of fBm, see Lemma 3.2. However, this Hypothesis 1.4 is also implicitly present in the additional condition of [16, Theorem 1].

(b) It can easily be shown that whenever ϑ is a one-dimensional coefficient (namely for $q = 1$), Hypothesis 1.6 is satisfied if the drift coefficient is of the form $b(x; \vartheta) = \vartheta h(x)$ for some $h : \mathbb{R}^d \to \mathbb{R}^d$ and if the stationary solution is non-degenerate, i.e. we have $\mathbf{E}|\overline{Y}_0|^2 \neq 0$. The latter conditions hold in particular in the case of the ergodic fractional Ornstein-Uhlenbeck process. It would be nice to obtain criteria for richer classes of examples, but this would rely on differentiability and non-degeneracy properties of the map $\vartheta \mapsto \mathbf{E}|b(\overline{Y}_0; \vartheta)|^2$ (see [11] in the Markovian case). We wish to investigate this question in future works.

(c) Even if the noise enters additively in our equation, we still need the assumption $H > 1/2$ in order to prove Theorem 1.7. Indeed, this hypothesis ensures the convergence of some deterministic and stochastic Riemann sums in the computations below (see Remark 3.3 for further details). Whether an adaptation of the proposed zero squares estimator is also convergent in the case $H < 1/2$ remains an open problem.

Let us finish this introduction with the simplest example of an equation which satisfies the above assumptions: namely the one-dimensional fractional Ornstein-Uhlenbeck process given by

$$dY_t = \vartheta_0 Y_t\, dt + dB_t, \qquad Y_0 = y_0 \in \mathbb{R}$$

with $\vartheta_0 < 0$. The solution of this SDE reads as

$$Y_t = y_0 \exp(\vartheta_0 t) + \exp(\vartheta_0 t) \int_0^t \exp(-\vartheta_0 s)\, dB_s.$$

For $t \to \infty$ this process converges to the stationary fractional Ornstein-Uhlenbeck process

$$\exp(\vartheta_0 t) \int_{-\infty}^t \exp(-\vartheta_0 s)\, dB_s, \qquad t \geq 0,$$

see e.g. [7]. Here straightforward computations yield the explicit estimator

$$\widehat{\vartheta}_n = \frac{\sum_{k=0}^{n-1} Y_{t_k} \delta Y_{t_k t_{k+1}}}{\sum_{k=0}^{n-1} Y_{t_k}^2 \alpha_n} - \sqrt{\left(\frac{\sum_{k=0}^{n-1} Y_{t_k} \delta Y_{t_k t_{k+1}}}{\sum_{k=0}^{n-1} Y_{t_k}^2 \alpha_n} \right)^2 - \frac{\sum_{k=0}^{n-1} \left(|\delta Y_{t_k t_{k+1}}|^2 - \alpha_n^{2H} \right)}{\sum_{k=0}^{n-1} Y_{t_k}^2 \alpha_n^2}}.$$

Notice that even in the case of the fractional Ornstein-Uhlenbeck process, we could neither proof the consistency nor show the inconsistency of our estimator for $H < 1/2$.

The remainder of this paper is structured as follows: In Section 2 we give some auxiliary results on stochastic calculus for fractional Brownian motion. Section 3 is then devoted to the proof of our main theorems.

2. Auxiliary Results

2.1. **Ergodic Properties of the SDE.** To deduce the ergodic properties of SDE (1) we will work without loss of generality on the canonical probability space $(\Omega, \mathcal{F}, \mathbf{P})$, i.e. $\Omega = C_0(\mathbb{R}, \mathbb{R}^m)$ equipped with the compact open topology, $\mathcal{F}$ is the corresponding Borel-σ-algebra and $\mathbf{P}$ is the distribution of the fractional Brownian motion B, which is consequently given here by the canonical process $B_t(\omega) = \omega(t)$, $t \in \mathbb{R}$. Together with the shift operators $\theta_t : \Omega \to \Omega$ defined by

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

the canonical probability space defines an ergodic metric dynamical system, see e.g. [9]. 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

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

for $\mathbf{P}$-almost all $\omega \in \Omega$. Owing to the results in Section 4 of [7] we have the following:

Theorem 2.1. *Let Hypothesis 1.3 hold. Then for any $\vartheta_0 \in \Theta$ the following holds:*
(i) *Equation (1) admits a unique solution Y in $C^\lambda(\mathbb{R}_+; \mathbb{R}^d)$ for all $\lambda < H$.*
(ii) *There exists a random variable $\overline{Y} : \Omega \to \mathbb{R}^d$ such that*

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

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

Note that the law of $\overline{Y}$ must coincide with the attracting invariant measure for (1) given in [10], see also [12, 13]. Moreover, proceeding as in [7] we have:

Proposition 2.2. *Assume Hypothesis 1.3 holds true. Then for any $\vartheta_0 \in \Theta$ and $p \geq 1$ there exist constants $c_p, k_p > 0$ such that*

$$\mathbf{E}\, |Y_t|^p \leq c_p, \qquad \mathbf{E}\, |Y_t - Y_s|^p \leq k_p |t - s|^{pH}, \qquad \textit{for all} \quad s, t \geq 0.$$

The integrability of $\overline{Y}$ now implies the ergodicity of equation (1):

Proposition 2.3. *Assume Hypothesis 1.3 holds true. Then for any $\vartheta_0 \in \Theta$ and any $f \in \mathcal{C}^1(\mathbb{R}^d;\mathbb{R})$ such that*

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

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

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

Proof. Since the shift operator is ergodic and f has polynomial growth, we have

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T f(\overline{Y}(\theta_t))\,dt = \mathbf{E} f(\overline{Y}) \qquad \mathbf{P}\text{-}a.s.$$

Moreover, since

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

by Theorem 2.1 and f is polynomially Lipschitz, the assertion easily follows. □

2.2. **Generalized Riemann-Stieltjes Integrals.** 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 $\lambda \in (0,1)$.

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 $\int_a^b f(x)\,dg(x)$ exists, see e.g. [24]. Also, the classical chain rule for the change of variables remains valid, see e.g. [25]: 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

$$F(f(y)) - F(f(a)) = \int_a^y F'(f(x))\,df(x), \qquad y \in [a,b]. \tag{5}$$

Moreover, one has a density type formula: let $f, h \in C^\lambda([a,b];\mathbb{R})$ and $g \in C^\mu([a,b];\mathbb{R})$ with $\lambda + \mu > 1$. Then for

$$\varphi : [a,b] \to \mathbb{R}, \quad \varphi(y) = \int_a^y f(x)\,dg(x), \qquad y \in [a,b],$$

we have

$$\int_a^b h(x)\,d\varphi(x) = \int_a^b h(x)f(x)\,dg(x). \tag{6}$$

For later use, we also note the following estimate, which can be found e.g. in [24].

Proposition 2.4. *Let f, g as above. There exists a constant $c_{\lambda,\mu}$ (independent of a, b) such that*

$$\left|\int_a^b (f(s) - f(a))dg(s)\right| \le c_{\lambda,\mu} |f|_{\lambda;[a,b]} |g|_{\mu;[a,b]} |b-a|^{\lambda+\mu}$$

holds for all $a, b \in [0,\infty)$.

2.3. **The Garcia-Rademich-Rumsey Lemma.** We will use the following variant of the Garcia-Rademich-Rumsey Lemma [8]:

Lemma 2.5. *Let $q > 1$, $\alpha \in (1/q, 1)$ and $f : [0, \infty) \to \mathbb{R}$ be a continuous function. Then there exists a constant $c_{\alpha,q} > 0$, depending only on α, q, such that*

$$|f|^q_{[s,t];\alpha-1/q} \leq c_{\alpha,q} \int_s^t \int_s^t \frac{|f(u) - f(v)|^q}{|u - v|^{1+q\alpha}} \, du \, dv.$$

2.4. **A Lemma on Pathwise Convergence Rates.** The following Lemma (see e.g. [19]), which is a direct consequence of the Borel-Cantelli Lemma, allows us to turn convergence rates in the p-th mean into pathwise convergence rates.

Lemma 2.6. *Let $\alpha > 0$, $p_0 \in \mathbb{N}$ and $c_p \in [0, \infty)$ for $p \geq p_0$. In addition, let Z_n, $n \in \mathbb{N}$, be a sequence of random variables such that*

$$(\mathbf{E}|Z_n|^p)^{1/p} \leq c_p \cdot n^{-\alpha}$$

for all $p \geq p_0$ and all $n \in \mathbb{N}$. Then for all $\varepsilon > 0$ there exists a random variable η_ε such that

$$|Z_n| \leq \eta_\varepsilon \cdot n^{-\alpha+\varepsilon} \qquad a.s.$$

for all $n \in \mathbb{N}$. Moreover, $\mathbf{E}|\eta_\varepsilon|^p < \infty$ for all $p \geq 1$.

2.5. **Quadratic Variations of Fractional Brownian Motion.** The following result for the behavior of the quadratic variations of a one-dimensional fractional Brownian motion β with Hurst parameter H is well known, see e.g [23]. Indeed, for $H < 3/4$ we have

$$\mathbf{E}\Big|\frac{1}{n}\sum_{k=0}^{n-1}\big[|\delta_{kk+1}\beta|^2 - 1\big]\Big|^2 \leq c_H \cdot \frac{1}{n}, \tag{7}$$

while for $H = \frac{3}{4}$, $n > 1$, it holds

$$\mathbf{E}\Big|\frac{1}{n}\sum_{k=0}^{n-1}\big[|\delta_{kk+1}\beta|^2 - 1\big]\Big|^2 \leq c_{3/4} \cdot \frac{\log(n)}{n}. \tag{8}$$

Finally, if $H \in (\frac{3}{4}, 1)$ then

$$\mathbf{E}\Big|\frac{1}{n}\sum_{k=0}^{n-1}\big[|\delta_{kk+1}\beta|^2 - 1\big]\Big|^2 \leq c_H \cdot \frac{1}{n^{4-4H}}. \tag{9}$$

Here, $c_H > 0$ denotes a constant depending only on H.

3. Proof of Theorems 1.5 and 1.7

We will denote constants, whose particular value is not important (and which do not depend on ϑ or n) by c, regardless of their value.

Recall that

$$Q_n(\vartheta) = \frac{1}{n\alpha_n^2} \sum_{k=0}^{n-1} \left(\left| \delta Y_{t_k t_{k+1}} - b(Y_{t_k}; \vartheta)\alpha_n \right|^2 - \|\sigma\|^2 \alpha_n^{2H} \right).$$

For $t \geq 0$, setting

$$F_t = \sum_{j=1}^{m} \sigma_j B_t^{(j)}, \quad \text{and} \quad r_k = \int_{t_k}^{t_{k+1}} \left(b(Y_u; \vartheta_0) - b(Y_{t_k}; \vartheta_0) \right) du$$

and moreover using the notation

$$\delta_{\vartheta_0 \vartheta} b(x) = b(x; \vartheta) - b(x; \vartheta_0), \quad \text{and} \quad \delta F_{t_k t_{k+1}} = F_{t_{k+1}} - F_{t_k},$$

it is readily checked that

$$\begin{aligned} Q_n(\vartheta) = & \frac{1}{n\alpha_n^2} \sum_{k=0}^{n-1} |\delta_{\vartheta_0 \vartheta} b(Y_{t_k})|^2 \alpha_n^2 - \frac{2}{n\alpha_n^2} \sum_{k=0}^{n-1} \langle \delta_{\vartheta_0 \vartheta} b(Y_{t_k}), \delta F_{t_k t_{k+1}} \rangle \alpha_n \\ & + \frac{1}{n\alpha_n^2} \sum_{k=0}^{n-1} \left(|\delta F_{t_k t_{k+1}}|^2 - \|\sigma\|^2 \alpha_n^{2H} \right) + \frac{1}{n\alpha_n^2} \sum_{k=0}^{n-1} |r_k|^2 \\ & - \frac{2}{n\alpha_n^2} \sum_{k=0}^{n-1} \langle \delta_{\vartheta_0 \vartheta} b(Y_{t_k}), r_k \rangle \alpha_n + \frac{2}{n\alpha_n^2} \sum_{k=0}^{n-1} \langle \delta F_{t_k t_{k+1}}, r_k \rangle. \end{aligned} \tag{10}$$

Note that our assumptions on the drift coefficient imply that

$$\sup_{\vartheta \in \Theta} |b(x; \vartheta) - b(y; \vartheta)| \leq c\big(1 + |x|^N + |y|^N\big) \cdot |x - y|$$

for all $x, y \in \mathbb{R}^d$ and

$$|b(x; \vartheta_1) - b(x; \vartheta_2)| \leq c\big(1 + |x|^N\big) \cdot |\vartheta_1 - \vartheta_2|$$

for all $x \in \mathbb{R}^d$ and $\vartheta_1, \vartheta_2 \in \Theta$. So, straightforward estimations using Proposition 2.2 give

$$\mathbf{E}|r_k|^p \leq c \cdot \alpha_n^{p(1+H)}.$$

Hence for all $p \geq 1$ it holds

$$\mathbf{E} \left| \sum_{k=0}^{n-1} |r_k|^2 \right|^p \leq c \cdot n^p \alpha_n^{2p(1+H)},$$

and Lemma 2.6 implies

$$\lim_{n \to \infty} \frac{1}{n\alpha_n^2} \sum_{k=0}^{n-1} |r_k|^2 = 0 \qquad \mathbf{P}\text{-}a.s. \tag{11}$$

Using Proposition 2.2 and Lemma 2.6 again, it follows similarly

$$\lim_{n \to \infty} \sup_{\vartheta \in \Theta} \frac{2}{n\alpha_n^2} \left| \sum_{k=0}^{n-1} \langle \delta_{\vartheta_0 \vartheta} b(Y_{t_k}), r_k \rangle \alpha_n \right| = 0 \qquad \mathbf{P}\text{-}a.s. \tag{12}$$

and, since $H > 1/2$, we also have

$$\lim_{n\to\infty} \frac{2}{n\alpha_n^2} \left| \sum_{k=0}^{n-1} \langle \delta F_{t_k t_{k+1}}, r_k \rangle \right| = 0 \qquad \mathbf{P}\text{-}a.s. \tag{13}$$

Plugging relations (11)–(13) into (10), we have obtained that

$$Q_n(\vartheta) = Q_n^{(1)}(\vartheta) - 2Q_n^{(2)}(\vartheta) + Q_n^{(3)} + R_n(\vartheta), \tag{14}$$

where $\lim_{n\to\infty} \sup_{\vartheta\in\Theta} |R_n(\vartheta)| = 0$ in the $\mathbf{P}$-almost sure sense and

$$Q_n^{(1)}(\vartheta) = \frac{1}{n} \sum_{k=0}^{n-1} |\delta_{\vartheta_0 \vartheta} b(Y_{t_k})|^2, \quad Q_n^{(2)}(\vartheta) = \frac{1}{n\alpha_n} \sum_{k=0}^{n-1} \langle \delta_{\vartheta_0\vartheta} b(Y_{t_k}), \delta F_{t_k t_{k+1}} \rangle$$

and

$$Q_n^{(3)} = \frac{1}{n\alpha_n^2} \sum_{k=0}^{n-1} \left(|\delta F_{t_k t_{k+1}}|^2 - \|\sigma\|^2 \alpha_n^{2H} \right).$$

The treatment of the terms $Q_n^{(1)}(\vartheta), Q_n^{(2)}(\vartheta)$ and $Q_n^{(3)}$ will be carried out in the following series of Lemmata. We first show a discrete version of Proposition 2.3:

Lemma 3.1. *Let $f \in \mathcal{C}^{1,1}(\mathbb{R}^d \times \Theta; \mathbb{R}^d)$ be a function such that*

$$|f(x;\vartheta)| \le c\left(1 + |x|^N\right), \quad |\partial_x f(x;\vartheta)| \le c\left(1 + |x|^N\right), \quad |\partial_\vartheta f(x;\vartheta)| \le c\left(1 + |x|^N\right)$$

for some $c > 0, N \in \mathbb{N}$, independent of $\vartheta \in \Theta$. Then we have

$$\sup_{\vartheta\in\Theta} \left| \frac{1}{n} \sum_{k=0}^{n-1} |f(Y_{t_k};\vartheta)|^2 - \mathbf{E}|f(\overline{Y};\vartheta)|^2 \right| \to 0 \qquad \mathbf{P}\text{-}a.s.$$

In particular, we have

$$\sup_{\vartheta\in\Theta} \left| Q_n^{(1)}(\vartheta) - \mathbf{E}|\delta_{\vartheta_0\vartheta} b(\overline{Y})|^2 \right| \to 0 \qquad \mathbf{P}\text{-}a.s.$$

Proof. Let $T_n = n\alpha_n$ and set

$$V_n(\vartheta) = \frac{1}{T_n} \int_0^{T_n} |f(Y_s;\vartheta)|^2 \, ds.$$

The ergodicity of Y yields that there exists a set $A_1 \in \mathcal{F}$ with full measure such that

$$\lim_{n\to\infty} V_n(\vartheta)(\omega) = \mathbf{E}|f(\overline{Y};\vartheta)|^2$$

for all $\vartheta \in \Theta \cap \mathbb{Q}^q$ and all $\omega \in A_1$. The assumptions on f give

$$|V_n(\vartheta_1) - V_n(\vartheta_2)| \le c \cdot \left(1 + \frac{1}{T_n} \int_0^{T_n} |Y_s|^{2N} \, ds \right) \cdot |\vartheta_1 - \vartheta_2|, \tag{15}$$

so V_n is Lipschitz continuous in ϑ and thus

$$\sup_{\vartheta\in\Theta} \left| V_n(\vartheta) - \mathbf{E}|f(\overline{Y};\vartheta)|^2 \right| = \sup_{\vartheta\in\Theta\cap\mathbb{Q}^q} \left| V_n(\vartheta) - \mathbf{E}|f(\overline{Y};\vartheta)|^2 \right|.$$

However, from (15) and the ergodicity of Y, it also follows that there exists a set $A_2 \in \mathcal{F}$ with $\mathbf{P}(A_2) = 1$ in which the family of random functions $V_n : \Theta \to \mathbb{R}$, $n \in \mathbb{N}$, is equicontinuous, and hence the Arzela-Ascoli Theorem yields the desired uniform convergence, i.e.

$$\lim_{n\to\infty} \sup_{\vartheta\in\Theta\cap\mathbb{Q}^q} \left| V_n(\vartheta) - \mathbf{E}|f(\overline{Y};\vartheta)|^2 \right| = 0 \qquad \mathbf{P}\text{-}a.s. \tag{16}$$

Setting

$$G_n(t;\vartheta) = |f(Y_t;\vartheta)|^2 - |f(Y_{t_k};\vartheta)|^2, \qquad t \in [t_k, t_{k+1}), \qquad k = 0, 1, \ldots,$$

it remains to show that

$$\frac{1}{T_n} \int_0^{T_n} \sup_{\vartheta\in\Theta} |G_n(t;\vartheta)| \, dt \to 0 \qquad \mathbf{P}\text{-}a.s.$$

To this aim, the assumptions on f imply that

$$\sup_{\vartheta\in\Theta} |G_n(t;\vartheta)| \le c \cdot (1 + Y_t^{2N} + Y_{t_k}^{2N}) \cdot |Y_t - Y_{t_k}|.$$

Using Proposition 2.2 and Hölder's inequality we obtain

$$\sup_{t\ge 0} \mathbf{E} \sup_{\vartheta\in\Theta} |G_n(t;\vartheta)|^p \le c \cdot \alpha_n^{pH} \tag{17}$$

for all $p \ge 1$. Now, Jensen's inequality gives

$$\mathbf{E} \left| \frac{1}{T_n} \int_0^{T_n} \sup_{\vartheta\in\Theta} |G_n(t;\vartheta)| \, dt \right|^p \le \frac{1}{T_n} \int_0^{T_n} \mathbf{E} \sup_{\vartheta\in\Theta} |G_n(t;\vartheta)|^p \, dt,$$

so (17) yields

$$\mathbf{E} \left| \frac{1}{T_n} \int_0^{T_n} \sup_{\vartheta\in\Theta} |G_n(t;\vartheta)| \, dt \right|^p \le c \cdot \alpha_n^{pH}$$

for all $p \ge 1$. Lemma 2.6 implies

$$\frac{1}{T_n} \int_0^{T_n} \sup_{\vartheta\in\Theta} |G_n(t;\vartheta)| \, dt \to 0 \qquad \mathbf{P}\text{-}a.s.$$

for $n \to \infty$. $\square$

We have a similar ergodic result for weighted sums of the increments of the process F.

Lemma 3.2. *Let $f \in \mathcal{C}^{1,1}(\mathbb{R}^d \times \Theta; \mathbb{R}^d)$ be a function such that*

$$|f(x;\vartheta)| \le c\left(1 + |x|^N\right), \quad |\partial_x f(x;\vartheta)| \le c\left(1 + |x|^N\right), \quad |\partial_\vartheta f(x;\vartheta)| \le c\left(1 + |x|^N\right)$$

for some $c > 0, N \in \mathbb{N}$, independent of $\vartheta \in \Theta$. Assume moreover that there exists a function $U \in \mathcal{C}^{2,1}(\mathbb{R}^d \times \Theta; \mathbb{R})$ such that

$$\partial_x U(x;\vartheta) = f(x;\vartheta), \qquad x \in \mathbb{R}^d, \ \vartheta \in \Theta,$$

i.e. f is of gradient type. Then, for $H > 1/2$, we have

$$\sup_{\vartheta\in\Theta} \left| \frac{1}{n\alpha_n} \sum_{k=0}^{n-1} \langle f(Y_{t_k};\vartheta), \delta F_{t_k t_{k+1}} \rangle + \mathbf{E}\langle b(\overline{Y};\vartheta_0), f(\overline{Y};\vartheta)\rangle \right| \to 0 \qquad \mathbf{P}\text{-}a.s.$$

In particular,

$$\sup_{\vartheta\in\Theta} \left| Q_n^{(2)}(\vartheta) + \mathbf{E}\langle b(\overline{Y};\vartheta_0), \delta_{\vartheta_0\vartheta} b(\overline{Y})\rangle \right| \to 0 \qquad \mathbf{P}\text{-}a.s.$$

Proof. Let $T_n = n\alpha_n$. First note that the chain of variable and density formula for Riemann-Stieltjes integrals, see (5) and(6) in Subsection 2.2, gives that

$$\frac{1}{T_n}\left(U(Y_{T_n};\vartheta) - U(y_0;\vartheta)\right) = \frac{1}{T_n}\int_0^{T_n} \langle f(Y_u;\vartheta), b(Y_u;\vartheta_0)\rangle\, du + \frac{1}{T_n}\int_0^{T_n} \langle f(Y_u;\vartheta), dF_u\rangle.$$

Now the properties of f, Proposition 2.2 and Lemma 2.6 imply that

$$\sup_{\vartheta\in\Theta} \frac{1}{T_n} |U(Y_{T_n};\vartheta) - U(y_0;\vartheta)| \to 0 \qquad \mathbf{P}\text{-}a.s.$$

Moreover, we have

$$\sup_{\vartheta\in\Theta} \left| \frac{1}{T_n}\int_0^{T_n} \langle f(Y_u;\vartheta), b(Y_u;\vartheta_0)\rangle\, du - \mathbf{E}\langle f(\overline{Y};\vartheta), b(\overline{Y};\vartheta_0)\rangle \right| \to 0 \qquad \mathbf{P}\text{-}a.s.,$$

which can be derived completely analogously to (16). It follows

$$\sup_{\vartheta\in\Theta} \left| \frac{1}{T_n}\int_0^{T_n} \langle f(Y_u;\vartheta), dF_u\rangle + \mathbf{E}\langle f(\overline{Y};\vartheta), b(\overline{Y};\vartheta_0)\rangle \right| \to 0 \qquad \mathbf{P}\text{-}a.s.$$

So, it remains to show that

$$\sup_{\vartheta\in\Theta} \frac{1}{T_n} \left| \int_0^{T_n} \langle G_n(t;\vartheta), dF_t\rangle \right| \to 0 \qquad \mathbf{P}\text{-}a.s. \tag{18}$$

where

$$G_n(t;\vartheta) = f(Y_t;\vartheta) - f(Y_{t_k};\vartheta), \qquad t\in[t_k,t_{k+1}), \qquad k = 0,1,\ldots.$$

Applying Proposition 2.4 and using the polynomial Lipschitz continuity of f yields, for all $\lambda < H$,

$$\left| \int_0^{T_n} \langle G_n(t;\vartheta), dF_t\rangle \right| \le c\cdot\alpha_n^{2\lambda}\cdot \sum_{j=1}^{m}\sum_{k=0}^{n-1} \sup_{t\in[t_k,t_{k+1}]} (1+|Y_t|^N)|Y|_{\lambda;[t_k;t_{k+1}]}|B^{(j)}|_{\lambda;[t_k;t_{k+1}]}.$$

From the Garcia-Rademich-Rumsey inequality, see Lemma 2.5, and Proposition 2.2 we have that

$$\left(\mathbf{E}|Y|^p_{\lambda;[t_k;t_{k+1}]}\right)^{1/p} \le c\cdot\alpha_n^{H-\lambda}$$

and also

$$\left(\mathbf{E}|B^{(j)}|^p_{\lambda;[t_k;t_{k+1}]}\right)^{1/p} \le c\cdot\alpha_n^{H-\lambda}.$$

Since moreover

$$\sup_{t\in[t_k,t_{k+1}]} (1+|Y_t|^N) \le c\cdot\left(Y_{t_k}^N + \alpha_n^{\lambda N}\cdot|Y|^N_{\lambda;[t_k,t_{k+1}]}\right)$$

and $\sup_{t\ge 0}\mathbf{E}|Y_t|^p < \infty$ for all $p \ge 1$, it follows that

$$\left(\mathbf{E}\sup_{\vartheta\in\Theta}\left|\frac{1}{T_n}\int_0^{T_n}\langle G_n(t;\vartheta), dF_t\rangle\right|^p\right)^{1/p} \le c\cdot\alpha_n^{2H-1} \tag{19}$$

Now Lemma 2.6 implies (18), since $H > 1/2$.

□

Remark 3.3. As mentioned in the introduction, the condition $H > 1/2$ is used in our proofs. Specifically, it is invoked in the convergence of the weighted stochastic integral (19) and also in order to derive (13).

The following Lemma deals with the remaining term, i.e. the quadratic variations of the process F:

Lemma 3.4. *We have*

$$\lim_{n\to\infty} Q_n^{(3)} = \lim_{n\to\infty} \frac{1}{n\alpha_n^2} \sum_{k=0}^{n-1} \left(|\delta F_{t_k t_{k+1}}|^2 - \|\sigma\|^2 \alpha_n^{2H} \right) = 0 \qquad \mathbf{P}\text{-}a.s.$$

with $\|\sigma\|^2 = \sum_{j=1}^m |\sigma_j|^2$.

Proof. We have

$$|\delta F_{t_k t_{k+1}}|^2 - \|\sigma\|^2 \alpha_n^{2H} = \underbrace{\sum_{j=1}^m |\sigma_j|^2 \left(|\delta B^{(j)}_{t_k t_{k+1}}|^2 - \alpha_n^{2H} \right)}_{=I_k^{(1)}} + \underbrace{\sum_{i,j=1,\, i\neq j}^m \langle \sigma_i, \sigma_j \rangle \delta B^{(i)}_{t_k t_{k+1}} \delta B^{(j)}_{t_k t_{k+1}}}_{=I_k^{(2)}}.$$

Owing to the scaling property of fBm it follows that

$$\mathbf{E} \left| \sum_{k=0}^{n-1} I_k^{(1)} \right|^p = \alpha_n^{2Hp} \mathbf{E} \left| \sum_{k=0}^{n-1} \sum_{j=1}^m |\sigma_j|^2 \left[|\delta B^{(j)}_{kk+1}|^2 - 1 \right] \right|^p.$$

Since all moments of random variables in a finite Gaussian chaos are equivalent, it follows from (7)-(9) that

$$\mathbf{E} \left| \sum_{k=0}^{n-1} \left[|\delta B^{(j)}_{kk+1}|^2 - 1 \right] \right|^p \leq c \cdot |\log(n)|^p \cdot n^{p(2H-1)}$$

and consequently

$$\mathbf{E} \left| \sum_{k=0}^{n-1} I_k^{(1)} \right|^p \leq c \cdot \alpha_n^{2Hp} \cdot |\log(n)|^p \cdot n^{p(2H-1)}.$$

Since $\alpha_n = \kappa \cdot n^{-\alpha}$ with $\alpha \in (0,1)$ Lemma 2.6 now implies that

$$\lim_{n\to\infty} \frac{1}{n\alpha_n^2} \left| \sum_{k=0}^{n-1} I_k^{(1)} \right| = 0 \qquad \mathbf{P}\text{-}a.s.$$

So it remains to consider the off-diagonal terms, i.e. $I_k^{(2)}$. Here we can exploit the following trick: Let β and $\tilde\beta$ be two independent fractional Brownian motions with the same Hurst index. From (7)-(9) we clearly have that

$$V_n = \sum_{k=0}^{n-1} \left(|\delta_{t_k t_{k+1}} \beta|^2 - |\delta_{t_k t_{k+1}} \tilde\beta|^2 \right)$$

satisfies

$$\mathbf{E}|V_n|^p \leq c \cdot \alpha_n^{2Hp} \cdot |\log(n)|^p \cdot n^{p(2H-1)}.$$

However, setting $B^{(i)} = (\beta + \widetilde\beta)/\sqrt{2}$ and $B^{(j)} = (\beta - \widetilde\beta)/\sqrt{2}$, then $B^{(i)}$ and $B^{(j)}$ are two independent fractional Brownian motions and

$$V_n \stackrel{\mathcal{L}}{=} 2 \sum_{k=0}^{n-1} \delta_{t_k t_{k+1}} B^{(i)} \delta_{t_k t_{k+1}} B^{(j)}.$$

Now we can easily conclude that

$$\lim_{n\to\infty} \frac{1}{n\alpha_n^2} \left| \sum_{k=0}^{n-1} I_k^{(2)} \right| = 0 \qquad \mathbf{P}\text{-}a.s.$$

□

Proof of Theorem 1.5. Let us go back to expression (10) and (14). Applying Lemma 3.1 and 3.2 we obtain that

$$\lim_{n\to\infty} \sup_{\vartheta\in\Theta} \Big| \big(Q_n^{(1)}(\vartheta) - 2Q_n^{(2)}(\vartheta)\big) - \big(\mathbf{E}|b(\overline{Y};\vartheta)|^2 - \mathbf{E}|b(\overline{Y};\vartheta_0)|^2\big) \Big| = 0$$

almost surely. Furthermore, recall that Lemma 3.4 asserts that $\lim_{n\to\infty} Q_n^{(3)} = 0$ almost surely. The proof is now finished.

□

Proof of Theorem 1.7. Let us first recall the following result ([6, 16]):

Proposition 3.5. *Assume that the family of random variables $L_n(\vartheta)$, $n \in \mathbb{N}$, $\vartheta \in \Theta$, satisfies:*

(1) *With probability one, $L_n(\vartheta) \to L(\vartheta)$ uniformly in $\vartheta \in \Theta$ as $n \to \infty$.*
(2) *The limit L is non-random and $L(\vartheta_0) \leq L(\vartheta)$ for all $\vartheta \in \Theta$.*
(3) *It holds $L(\vartheta) = L(\vartheta_0)$ if and only if $\vartheta = \vartheta_0$.*

Then, we have

$$\widehat{\vartheta}_n \to \vartheta_0 \qquad \mathbf{P}\text{-}a.s.$$

for $n \to \infty$, where

$$L_n(\widehat{\vartheta}_n) = \min_{\vartheta\in\Theta} L_n(\vartheta).$$

The strong consistency of the zero squares estimator follows now from Theorem 1.5 and an application of Proposition 3.5 to $|Q_n(\vartheta)|$.

□

Remark 3.6. In the case $H = 1/2$ we have under similar assumptions that

$$\sup_{\vartheta\in\Theta} \big| \big(Q_n(\vartheta) - Q_n(\vartheta_0)\big) - \big(\mathbf{E}\,|b(\overline{Y}_0;\vartheta_0) - b(\overline{Y}_0;\vartheta)|^2\big) \big| \to 0$$

in the $\mathbf{P}$-almost sure sense, see e.g. [5]. Since

$$Q_n(\vartheta_0) = \frac{1}{n\alpha_n^2} \sum_{k=0}^{n-1} \big(|\delta F_{t_k t_{k+1}}|^2 - \|\sigma\|^2\alpha_n\big) + \frac{1}{n\alpha_n^2} \sum_{k=0}^{n-1} |r_k|^2 + \frac{2}{n\alpha_n^2} \sum_{k=0}^{n-1} \langle \delta F_{t_k t_{k+1}}, r_k\rangle,$$

an application of Lemma 3.1 and 3.4 (which are also valid for $H = 1/2$) yield that

$$\lim_{n\to\infty} Q_n(\vartheta_0) = \lim_{n\to\infty} \frac{2}{n\alpha_n^2} \sum_{k=0}^{n-1} \langle \delta F_{t_k t_{k+1}}, r_k\rangle \qquad \mathbf{P}\text{-}a.s.$$

However, using the Itô-isometry and the Burkholder-Davis-Gundy inequality we have

$$\mathbf{E} \left| \frac{1}{n\alpha_n^2} \sum_{k=0}^{n-1} \langle r_k, \delta_{t_k t_{k+1}} F\rangle \right|^p \leq c \cdot \frac{1}{n^p \alpha_n^{2p}} \cdot T_n^{1+p/2} \alpha_n^{3p/2} \leq c \cdot n^{-p/2+1-\alpha}$$

and Lemma 2.6 thus gives

$$\lim_{n\to\infty} Q_n(\vartheta_0) = 0 \qquad \mathbf{P}\text{-}a.s.$$

Hence we end up with

$$\sup_{\vartheta\in\Theta} \left| Q_n(\vartheta) - \left(\mathbf{E}\, |b(\overline{Y}_0;\vartheta_0) - b(\overline{Y}_0;\vartheta)|^2 \right) \right| \to 0,$$

so the limit of the statistics Q_n is different for $H = 1/2$, where one obtains the standard least square estimator.

References

[1] A. Bégyn (2005): Quadratic variations along irregular subdivisions for Gaussian processes. *Electronic J. Probab.* **10**, 691–717.

[2] R. Belfadli, K. Es-Sebaiy, Y. Ouknine (2011): Parameter Estimation for Fractional Ornstein-Uhlenbeck Processes: Non-ergodic Case. *Arxiv* Preprint.

[3] A. Chronopoulou, S. Tindel (2011): On inference for fractional differential equations. *Arxiv* Preprint.

[4] J.F. Coeurjolly (2001): Estimating the Parameters of a Fractional Brownian Motion by Discrete Variations of its Sample Paths. *Stat. Infer. Stoch. Process.* **4**, no. 2, 199–227.

[5] D. Florens-Zmirou (1989): Approximate discrete-time schemes for statistics of diffusion processes. *Statistics* **20**, no. 4, 547–557.

[6] R. Frydman (1980): A proof of the consistency of maximum likelihood estimators of non-linear regression models with autocorellated errors. *Econometrica* **48**, 853–860.

[7] 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.

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

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

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

[11] M. Hairer, A. Majda (2010): A simple framework to justify linear response theory. *Nonlinearity* **23**, no. 4, 909–922.

[12] M. Hairer, A. Ohashi (2007): Ergodicity theory of SDEs with extrinsic memory. *Ann. Probab.* **35**, no. 5, 1950–1977.

[13] M. Hairer, S. Pillai (2011): Ergodicity of hypoelliptic SDEs driven by fractional Brownian motion. *Ann. Inst. Henri Poincaré, Probab. Stat.* **47**, no. 2, 601–628.

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

[15] J. Istas, G. Lang (1994): Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Ann. Inst. Poincaré* **33**, 407–436.

[16] R.A. Kasonga (1988): The consistency of a non-linear least squares estimator from diffusion processes. *Stoch. Proc. Appl.* **30**, 263–275.

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

[18] 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.

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

[20] A. Papavasiliou, C. Ladroue (2011): Parameter estimation for rough differential equations. *Ann. Statist.*, to appear.

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

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

[23] C. Tudor, F. Viens (2009): Variations and estimators for self-similarity parameters via Malliavin calculus. *Ann. Probab.* **37**, no. 6, 2093–2134.
[24] Young, L.C. (1936). An inequality of Hölder type connected with Stieltjes integration. *Acta Math.* **67** 251–282.
[25] Zähle, M. (2005). Stochastic differential equations with fractal noise. *Math. Nachr.* **278**, no. 9, 1097–1106.

Andreas Neuenkirch, Fachbereich Mathematik, TU Kaiserslautern, Postfach 3049, D-67663 Kaiserslautern, Germany neuenkirch@mathematik.uni-kl.de

Samy Tindel, Institut Élie Cartan Nancy, Université de Nancy 1, B.P. 239, 54506 Vandœuvre-lès-Nancy Cedex, France tindel@iecn.u-nancy.fr