# Statistical Aspects of the Fractional Stochastic Calculus


Ciprian A. Tudor[1] Frederi G. Viens[2]

[1]SAMOS-MATISSE, Université de Paris 1 Panthéon-Sorbonne ,
90, rue de Tolbiac, 75634, Paris, France.
tudor@univ-paris1.fr

[2] Dept. Statistics and Dept. Mathematics, Purdue University,
150 N. University St., West Lafayette, IN 47907-2067, USA.
viens@purdue.edu +1 (765) 494 6035


March 11, 2005


**Abstract**

We apply the techniques of stochastic integration with respect to the fractional Brownian motion and the Gaussian theory of regularity and supremum estimation to study the maximum likelihood estimator (MLE) for the drift parameter of stochastic processes satisfying stochastic equations driven by fractional Brownian motion with any level of Hölder-regularity (any *Hurst* parameter). We prove existence and strong consistency of the MLE for linear and nonlinear equations. We also prove that a basic discretized version of the MLE, is still a strongly consistent estimator.



**Key Words and Phrases:** maximum likelihood estimator, fractional Brownian motion, strong consistency, stochastic differential equation, Gaussian regularity theory, Malliavin calculus, Hurst parameter.


## 1 Introduction

Stochastic calculus with respect to fractional Brownian motion (fBm) has recently known an intensive development, motivated by the wide array of applications of this family of stochastic processes. For example, fBm is used as a model in network traffic analysis; recent work and empirical studies have shown that traffic in modern packet-based high-speed networks frequently exhibits fractal behavior over a wide range of time scales; this has major implications for the statistical study of such traffic. An other example of applications is in quantitative finance and econometrics: the fractional Black-Scholes model has been recently introduced (see e.g. [10], [8]) and this motivates the statistical study of stochastic differential equations governed by fBm. We also note the use of the fBm in other branches of science and engineering such as hydrology or biophysics.

The topic of parameter estimation for stochastic differential equations driven by standard Brownian motion is not new. Of particular contemporary interest are works in which an approximate estimator, using only information gleaned from the underlying process in discrete time, is able to do as well as an estimator that uses continuously gathered information. We mention the fundamental seminal works of [17], [11], or the book [13]. More recently the topic was taken up again in [23], and most recently, with the comparison of a number of different techniques in [6], or some sharp probabilistic bounds on the convergence of estimators in [3].

Parameter estimation questions for stochastic differential equations driven by fBm are, in contrast, in their infancy. Some of the main contributions include [15], [14], [25] or [16]. Our purpose is to contribute further to the study of the statistical aspects of the fractional stochastic calculus, by introducing the systematic use of efficient tools from stochastic analysis. We consider the following stochastic equation

$$X_t = \theta \int_0^t b(X_s)ds + B_t^H, \quad X_0 = 0$$

where $B^H$ is a fBm with *Hurst* parameter $H \in (0,1)$ and the nonlinear function $b$ satisfies some regularity conditions. We estimate the parameter $\theta$ on the basis of the observation of the whole trajectory of the process $X$. The parameter $H$, which is assumed to be known, characterizes the local behavior of the process, with Hölder-regularity increasing with $H$; if $H = 1/2$, fBm is standard Brownian motion (BM), and thus has independent increments; if $H > 1/2$, the increments of fBm are positively correlated, and the process is more regular than BM; if $H < 1/2$, the increments are negatively correlated, and the process is less regular than BM. $H$ also characterizes the speed of decay of the correlation between distant increments.

The results we prove in this paper are as follows:

- for every $H \in (0,1)$, we give concrete assumptions on the nonlinear coefficient $b$ to ensure the existence of the maximum likelihood estimator (MLE) for the parameter $\theta$;

- for every $H \in (0,1)$ and under certain hypothesis on $b$ (which include non-linear examples) we prove the strong consistency of the MLE; note that for $H > \frac{1}{2}$ and $b$ linear, this has also been proved in [14];

- for every $H \in (0,1)$, a strategy is established in view of proving that a discrete, Riemann-sum-based version of the MLE also is a strongly consistent estimator of $\theta$; this strategy is successfully implemented in the linear case.

To establish these results, we use the techniques of the *Malliavin* calculus and the so-called *Dudley-Fernique* theory of regularity and supremum estimation for Gaussian and/or sub-Gaussian processes. To our knowledge, our paper is the first instance where the Malliavin derivative of a stochastic process is used to provide specific information about the process's pathwise regularity; it is given here mainly in the sub-Gaussian context (see Lemmas 1 and 2), although some hint at the fact that the method should be applicable beyond the Gaussian realm can be found here as well (Lemma 6). This new technique should have applications and implications in statistics and probability reaching far beyond the question of MLE for fBm.

For example, apart from providing the first proof of strong consistency of the MLE for an fBm-driven differential equation with non-linear drift or with $H < 1/2$, some of the broader implications of our Gaussian and sub-Gaussian methods should include possible applications to classical situations. That is, in general, in (Itô-) diffusion models, the strong consistency of an estimator follows if one can prove that an expression of the type $I_t := \int_0^t f^2(X_s)ds$ tends to $\infty$ as $t \to \infty$ almost surely. To our knowledge, a limited number of methods has been employed to deal with this kind of problem: for example, if $X$ is Gaussian the Laplace transform can be computed explicitly to show that $\lim_{t\to\infty} I_t = \infty$ a.s.; also, if $X$ is an ergodic diffusion, a local time argument can be used to show the above convergence. Particular situations have also been considered in [11], [12]. We believe that our (sub-)Gaussian tools constitute a new possibility, judging by the fact that the case of $H < 1/2$ is well within reach of our tools, in contrast with the other above-mentioned methods, as employed in particular in [14] (see however a general Bayesian-type problem discussed in [15]).

Our goal here is not to present a practical implementation of the MLE studied herein; such an additional development should be highly non-trivial, and will be the subject of a separate article. We wish nevertheless to give evidence that it should be possible to provide a version of the MLE which can be understood using only discrete observations of the solution $X$ of equation (1). To illustrate this idea, we show in the last section of this article that the natural discrete-time Riemann-sum approximation of the MLE is still a strongly consistent estimator of $\theta$. This is achieved by using more Gaussian and sub-Gaussian tools. Our main concrete result, which holds in the linear case, shows that the discretization time-step can be fixed while still allowing for a strongly consistent estimator in large time. This is a good indication that a true discrete-observation-based strongly consistent estimator should be available.

The organization of our paper is as follows. Section 2 contains preliminaries on the fBm and basic elements of the fractional calculus. In Section 3 we show the existence of the maximum likelihood estimator for the parameter $\theta$ in (6) and in Section 4 we study its asymptotic behavior. Section 5 contains some additional results in the case when the drift function is linear, which is pursued further in Section 6, where a discretized version of the MLE is studied.

## 2 Preliminaries on the fractional Brownian motion and fractional calculus

We consider $(B_t^H)_{t\in[0,T]}$, $B_0^H=0$ a fractional Brownian motion with Hurst parameter $H\in(0,1)$. This is a centered Gaussian process with covariance function $R$ given by

$$R(t,s)=\frac{1}{2}\left(t^{2H}+s^{2H}-|t-s|^{2H}\right)\qquad s,t\in[0,T].$$

Let us denote by $K$ the kernel of the fBm such that (see e.g. [20])

$$B_t^H=\int_0^t K(t,s)dW_s$$

where $W$ is a Wiener process (standard Brownian motion). Denote by $\mathcal{E}_H$ the set of step functions on $[0,T]$ and let $\mathcal{H}$ be the canonical Hilbert space of the fBm; that is, $\mathcal{H}$ is the closure of $\mathcal{E}$ with respect to the scalar product

$$\langle 1_{[0,t]},1_{[0,s]}\rangle_{\mathcal{H}}=R(t,s).$$

The mapping $1_{[0,t]}\to B_t^H$ can be extended to a isometry between $\mathcal{H}$ and the Gaussian space generated by $B^H$ and we denote by $B^H(\varphi)$ the image of $\varphi\in\mathcal{H}$ by this isometry.

We also introduce the operator $K^*$ from $\mathcal{E}_H$ to $L^2([0,T])$ defined by

$$(K^*\varphi)(s)=K(T,s)\varphi(s)+\int_s^T(\varphi(r)-\varphi(s))\frac{\partial K}{\partial r}(r,s)dr. \tag{1}$$

With this notation we have $(K^*1_{[0,t]})(s)=K(t,s)$ and hence the process

$$W_t=\int_0^t(K^{*,-1}1_{[0,t]})(s)dB_s^H$$

is a Wiener process (see [2]); in fact, it is *the* Wiener process referred to in formula (2), and for any non-random $\varphi\in\mathcal{H}$, we have

$$B^H(\varphi)=\int_0^T(K^*\varphi)(s)dW(s),$$

where the latter is a standard Wiener integral with respect to $W$.

We also need some elements of the fractional calculus. Let $f$ be a function over the interval $[0,T]$ and $\alpha>0$. Then

$$I_{0+}^{\alpha}f(t)=\frac{1}{\Gamma(\alpha)}\int_0^T\frac{f(s)}{(t-s)^{1-\alpha}}ds$$

are the Riemann-Liouville fractional integrals of order $\alpha$. For $\alpha \in (0,1)$,

$$D^{\alpha}_{0+} f(t) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_0^T \frac{f(s)}{(t-s)^{\alpha}} ds$$

are the Riemann-Liouville fractional derivatives of order $\alpha$. These derivative admits the following Weil representation

$$D^{\alpha}_{0+} f(t) = \frac{1}{\Gamma(1-\alpha)} \left( \frac{f(t)}{t^{\alpha}} + \alpha \int_0^t \frac{f(t)-f(y)}{(t-y)^{\alpha+1}} dy \right)$$

where the convergence of the integrals at $t = y$ holds in the $L^p$-sense.

We can formally define, for negative orders ($-\alpha < 0$), the fractional integral operators as

$$I^{-\alpha}_{\pm} = D^{\alpha}_{\pm}.$$

If $K_H$ is the linear operator (isomorphism) from $L^2([0,T])$ onto $I_+^{H+\frac{1}{2}}(L^2([0,T]))$ whose kernel is $K(t,s)$, then it can be expressed by fractional integrals as (see e.g. [7])

$$K_H f(t) = I^{2H}_{0+} \left( t^{\frac{1}{2}-H} I^{\frac{1}{2}-H}_{0+} \left( s^{H-\frac{1}{2}} f(s) \right)(t) \right)(t), \qquad H \leq \frac{1}{2} \tag{2}$$

and

$$K_H f(t) = I^{1}_{0+} \left( t^{H-\frac{1}{2}} I^{H-\frac{1}{2}} \left( s^{\frac{1}{2}-H} f(s) \right)(t) \right)(t), \qquad H \geq \frac{1}{2}. \tag{3}$$

We recall the expressions, for $h$ differentiable, of the inverse operator of $K_H$ in terms of fractional integrals

$$\left(K_H^{-1} h\right)(s) = s^{H-\frac{1}{2}} I^{\frac{1}{2}-H}_{0+} (s^{\frac{1}{2}-H} h'(s))(s), \qquad H \leq \frac{1}{2} \tag{4}$$

and

$$\left(K_H^{-1} h\right)(s) = s^{H-\frac{1}{2}} D^{H-\frac{1}{2}}_{0+} (s^{\frac{1}{2}-H} h'(s))(s), \qquad H \geq \frac{1}{2}. \tag{5}$$

# 3 The maximum likelihood estimator for fBm-driven stochastic differential equations

We will analyze the estimation of the parameter $\theta \in \Theta \subset \mathbb{R}$ based on the observation of the solution $X$ of the stochastic differential equation

$$X_t = \theta \int_0^t b(X_s) ds + B^H_t, \qquad X_0 = 0 \tag{6}$$

where $B^H$ is a fBm with $H \in (0,1)$ and $b : \mathbb{R} \to \mathbb{R}$ is a measurable function. Let us recall some known results concerning equation (6):

- In [22] the authors proved the existence and uniqueness of a strong solution to equation (6) under the following assumptions on the coefficient $b$:
  - if $H < \frac{1}{2}$, $b$ satisfies the linear growth condition

$$|b(x)| \leq C(1 + |x|). \tag{7}$$

  - if $H > \frac{1}{2}$, $b$ is Hölder-continuous of order $\alpha \in (1 - \frac{1}{2H}, 1)$.

- In [4] an existence and uniqueness result for (6) is given when $H > \frac{1}{2}$ under the hypothesis $b(x) = b_1(x) + b_2(x)$, $b_1$ satisfying the above conditions and $b_2$ being a bounded nondecreasing left (or right) continuous function.

**Remark 1** *The case of the Hölder-continuous drift is elementary: it is not difficult to show that the usual Picard iteration method can be used to prove the existence and uniqueness of a strong solution.*

Throughout the paper, from now on, we will avoid the use of the $H$-dependent constants appearing in the definitions of the operator kernels related to this calculus, since our main interest consists asymptotic properties on the different estimators. In consequence, we will use the notation $C(H), c(H), \ldots$ for generic constants depending on $H$, which may change from line to line.

Our construction is based on the following observation (see [22]). Consider the process

$$\tilde{B}_t^H = B_t^H + \int_0^t u_s ds$$

where the process $u$ is adapted and with integrable paths. Then we can write

$$\tilde{B}_t^H = \int_0^t K(t,s) dZ_s \tag{8}$$

where

$$Z_t = W_t + \int_0^t K_H^{-1}\left(\int_0^{\cdot} u_r dr\right)(s) ds. \tag{9}$$

We have the following Girsanov theorem.

**Theorem 1** *i) Assume that $u$ is an adapted process with integrable paths such that*

$$t \to \int_0^t u_s ds \in I^{H+\frac{1}{2}}\left(L^2([0,T])\right) \ \ a.s.$$

*ii) Suppose that $\mathbf{E}(V_T) = 1$ where*

$$V_T = \exp\left(-\int_0^T K_H^{-1}\left(\int_0^{\cdot} u_r dr\right)(s) dW_s - \frac{1}{2}\int_0^T \left(K_H^{-1}\left(\int_0^{\cdot} u_r dr\right)(s)\right)^2 ds\right). \tag{10}$$

*Then under the probability measure $\tilde{P}$ defined by $d\tilde{P}/dP = V_T$ it holds that the process $Z$ defined in (9) is a Brownian motion and the process $\tilde{B}^H$ (8) is a fractional Brownian motion on $[0,T]$.*

**Hypothesis.** We need to make, at this stage and throughout the remainder of the paper, the following assumption on the drift: $b$ is differentiable with bounded derivative $b'$ and the affine growth condition (7) holds.

As a consequence of the Girsanov theorem, we obtain the following expression of the MLE.

**Proposition 1** *Denote, for every $t \in [0,T]$, by*

$$Q_t = Q_t(X) = K_H^{-1}\left(\int_0^{\cdot} b(X_r) dr\right)(t). \tag{11}$$

*Then $Q \in L^2([0,T])$ almost surely and the MLE is given by*

$$\theta_t = -\frac{\int_0^t Q_s dW_s}{\int_0^t Q_s^2 ds}. \tag{12}$$

In order to prove this proposition, we need to establish a quadratic exponential moment inequality. We begin with an estimation of some relevant Malliavin derivatives. For our purposes, it is sufficient to recall the following two facts from the Malliavin calculus (see [21]) with respect to $W$, where $W$ is the underlying standard Brownian motion we have been working with, i.e. the one found for example in formula (2).

- The Malliavin derivative $D_rF$ at time $r$ of a random variable of the form $F = \int_0^T f(s)\,dW_s$ where $f$ is a non-random function in $L^2([0,T])$, is equal to $f(r)$.
- Let $F$ be a random variable with Malliavin derivative $D_rF$ at time $r$. The random variable $G = \Phi(F)$, where $\Phi$ is a differentiable function such that both $G$ and $\Phi'(F)$ are members of $L^2(\Omega)$, has a Malliavin derivative satisfying $D_rG = \Phi'(F)\,D_rF$.
- Note that the letter $D$ for the Malliavin derivative of a random variable should not be confused with the notation used for the fractional derivative of a non-random function.

**Lemma 1** *Let $X$ be the unique solution to the Langevin equation (6). Let $r \in [0,T]$ and $0 \le t \le t' \le T$. Let*

$$A_r(t,t') = D_r(X_{t'} - X_t).$$

*Then the following estimates hold: there exists a constant $K = K(T,H,b)$ depending only on $T, H,$ and $\|b'\|_\infty$, such that for all $t \le t' \le T$*

$$\|A_\cdot(0,t)\|^2_{L^2(dr)} \le K;$$
$$\|A_\cdot(t,t')\|^2_{L^2(dr)} \le K\,|t'-t|^{2H}.$$

**Proof:** Since we are required to estimate Malliavin derivatives with respect to $W$ rather than $B^H$, it is most convenient to first convert all expressions to formulas involving $W$. This is of course trivial using (2). We thus have

$$X_{t'} - X_t = \int_t^{t'} b(X_s)\,ds + \int_0^T \left(\mathbf{1}_{[0,t']}(s)\,K(t',s) - \mathbf{1}_{[0,t]}(s)\,K(t,s)\right) dW_s$$

and using the above rules for Malliavin differentiation, which commutes with Riemann integration w.r.t. $ds$,

$$\begin{aligned} A_r(t,t') = \mathbf{1}_{[0,t]}(r) \int_t^{t'} b'(X_s)\,A_r(0,s)\,ds + \mathbf{1}_{[t,t']}(r) \int_r^{t'} b'(X_s)\,A_r(0,s)\,ds \\ + \mathbf{1}_{[0,t']}(r)\,K(t',r) - \mathbf{1}_{[0,t]}(r)\,K(t,r). \end{aligned} \tag{13}$$

In particular we get $A_r(0,t) = 0$ for $r \notin [0,t]$, while for $r \in [0,t]$ we have

$$A_r(0,t) = \int_r^t b'(X_s)\,A_r(0,s)ds + K(t,r).$$

Therefore

$$\begin{aligned} \|A_\cdot(0,t)\|^2_{L^2(dr)} &= \int_0^t \left| \int_r^t b'(X_s)\,A_r(0,s)\,ds + K(t,r) \right|^2 dr \\ &\le 2\,\|b'\|^2_\infty \int_0^t dr \left| \int_r^t A_r(0,s)\,ds \right|^2 + 2\int_0^t K(t,r)^2\,dr \\ &= 2\,\|b'\|^2_\infty \int_0^t dr \left| \int_r^t A_r(0,s)\,ds \right|^2 + 2t^{2H} \\ &\le 2\,\|b'\|^2_\infty \int_0^t (t-r) \int_r^t |A_r(0,s)|^2\,dsdr + 2t^{2H} \\ &\le 2\,\|b'\|^2_\infty\, t \int_0^t ds \int_0^s |A_r(0,s)|^2\,dr + 2t^{2H}, \end{aligned}$$

where we made use of Jensen's inequality and Fubini's lemma. Since the last expression above is equal to

$$2\left\|b'\right\|_{\infty}^{2} t\int_{0}^{t}\left\|A_{\cdot}\left(0,s\right)\right\|_{L^{2}(dr)}^{2} ds+2t^{2H}$$

from the standard Gronwall lemma, we obtain immediately

$$\left\|A_{\cdot}\left(0,t\right)\right\|_{L^{2}(dr)}^{2}\leq 2t^{2H}\exp\left(2\left\|b'\right\|_{\infty}^{2} t^{2}\right),$$

which proves the first statement of the Lemma.

For the second statement, we may now use what we have just proved together with equality (13). Since $\mathbf{E}\left[\left(B_{t}^{H}-B^{H}\left(t'\right)\right)^{2}\right]=\left|t-t'\right|^{2H}$, we obtain directly

$$\begin{aligned}\left\|A_{\cdot}\left(t,t'\right)\right\|_{L^{2}(dr)}^{2} &\leq 2\left\|b'\right\|_{\infty}^{2}\int_{0}^{t}\left|\int_{t}^{t'}A_{r}\left(0,s\right)ds\right|^{2} dr\\ &\quad+2\left\|b'\right\|_{\infty}^{2}\int_{t}^{t'}\left|\int_{r}^{t'}A_{r}\left(0,s\right)ds\right|^{2} dr+2\left|t-t'\right|^{2H}.\end{aligned}$$

We evaluate the first term on the right-hand side using Jensen's inequality:

$$\begin{aligned}\int_{0}^{t}\left|\int_{t}^{t'}A_{r}\left(0,s\right)ds\right|^{2} dr &\leq\left(t'-t\right)\int_{0}^{t}\int_{t}^{t'}\left|A_{r}\left(0,s\right)\right|^{2} dsdr\\ &\leq\left(t'-t\right)^{2}\sup_{s\leq t}\left\|A_{\cdot}\left(0,s\right)\right\|_{L^{2}(dr)}^{2}\\ &=K\left|t-t'\right|^{2}.\end{aligned}$$

For the second term on the right-hand side, we have

$$\begin{aligned}\int_{t}^{t'}\left|\int_{r}^{t'}A_{r}\left(0,s\right)ds\right|^{2} dr &\leq\int_{t}^{t'}\left(t'-r\right)\int_{r}^{t'}\left|A_{r}\left(0,s\right)\right|^{2} dsdr\\ &=\int_{t}^{t'}ds\int_{t}^{s}\left|A_{r}\left(0,s\right)\right|^{2}\left(t'-r\right)dr\\ &\leq\left(t'-t\right)^{2}\int_{t}^{t'}ds\sup_{s\leq t}\left\|A_{\cdot}\left(0,s\right)\right\|_{L^{2}(dr)}^{2}\\ &=K\left|t-t'\right|^{2}.\end{aligned}$$

Since $2H<2$, the conclusion of the lemma follows. ■

We now use the lemma we have just proved together with a so-called *Poincaré* inequality on Wiener space. We have the following, which states that $X$ is a *sub-Gaussian* process relative to the metric $K\left|t-t'\right|^{H}$.

**Lemma 2** *Let $X$ be the solution of (6). There exists a constant $K$ depending only on $T,H$, and $\left\|b'\right\|_{\infty}$ such that for any $\lambda\in\mathbf{R}$,*

$$\mathbf{E}\left[\exp\lambda\left(X_{t'}-X_{t}\right)\right]\leq\exp\left(\frac{\lambda^{2}}{2}K\left|t-t'\right|^{2H}\right).$$

**Proof**. This is a trivial application of the result of the previous lemma to the Poincaré inequality [26, Theorem 3 (i) page 76] which states that for any centered random variable $F$ in $L^{2}\left(\Omega\right)$,

$$\mathbf{E}\left[\exp\left(F\right)\right]\leq\mathbf{E}\left[\exp\left(\frac{\pi^{2}}{8}\left|D_{\cdot}F\right|_{L^{2}(dr)}^{2}\right)\right].$$

■

This lemma can be invoked in the context of sub-Gaussian processes: since $X$ is now proved to be a sub-Gaussian process with "canonical metric" bounded above by $K\left|t-t'\right|^{H}$ , this means that the variations of $X$ are bounded above, in distribution, by those of a corresponding centered Gaussian process $Y$ satisfying $\mathbf{E}\left[\left(Y\left(t\right)-Y\left(t'\right)\right)^{2}\right]=\left|t-t'\right|^{H}$. Of course, this process is none other than fractional Brownian motion with parameter $H$. One can consult [18, Chapter 12] for details on the relation between the regularity of Gaussian and sub-Gaussian processes, and [1] for a more self-contained account of Gaussian supremum estimates. In particular we have the following results.

**Lemma 3** *There exists a constant $K$ depending only on $H$ and $\left\|b'\right\|_{\infty}$ such that for all $t>0$*

$$\mu_{t}:=\mathbf{E}\left[\sup_{s\in[0,t]}\left|X_{s}\right|\right]<Kt^{H}. \tag{14}$$

*Also let $\sigma_{t}^{2}:=\sup_{s\in[0,t]}var\left(X\left(s\right)\right)$. Then for any $x>0$,*

$$\mathbf{P}\left[\left|\sup_{s\in[0,t]}\left|X_{s}\right|-\mu_{t}\right|>x\right]\leq4\exp\left(-\frac{x^{2}}{2\sigma_{t}^{2}}\right). \tag{15}$$

*As a consequence, we have that for every positive constant $\alpha<1/\left(2\sigma_{T}^{2}\right)$,*

$$\mathbf{E}\left[\exp\left(\alpha\sup_{s\in[0,T]}\left|X_{s}\right|^{2}\right)\right]<\infty. \tag{16}$$

**Proof.** Inequality (15) is exactly the so-called *Borell-Sudakov* inequality, which holds for any separable centered Gaussian or sub-Gaussian process, so there is nothing to prove: see [1] and [18, Chapter 12]. It is useful only if a sharp estimation of $\mu_{t}$ can be obtained. The last result (16) in this lemma is an application of what is often known as Fernique's theorem in the Gaussian context, but we establish the result in our sub-Gaussian context using (15), as follows. Note that we always have $\mu_{T}>0$. Also note that $\sigma_{T}^{2}$ is finite, as can trivially be inferred from Lemma 2 for example. Let $Z=\sup_{s\in[0,T]}\left|X_{s}\right|$. Then

$$\begin{aligned}
\mathbf{E}\left[\exp\left(\alpha Z^{2}\right)\right] &= \int_{0}^{\infty}\mathbf{P}\left[\exp\left(\alpha Z^{2}\right)>x\right]dx\\
&=1+\int_{1}^{\infty}\mathbf{P}\left[\left|Z\right|>\sqrt{\alpha^{-1}\log x}\right]dx\\
&=1+\int_{0}^{\infty}\mathbf{P}\left[\left|Z\right|>y\right]2\alpha ye^{\alpha y^{2}}dy\\
&\leq1+2\alpha\int_{0}^{\mu_{T}}ye^{\alpha y^{2}}dy\\
&+2\alpha\int_{\mu_{T}}^{\infty}\mathbf{P}\left[Z-\mu_{T}>y-\mu_{T}\right]ye^{\alpha y^{2}}dy\\
&+2\alpha\int_{\mu_{T}}^{\infty}\mathbf{P}\left[Z-\mu_{T}<-\left(y+\mu_{T}\right)\right]ye^{\alpha y^{2}}dy\\
&\leq1+\mu_{T}^{2}e^{\alpha\mu_{T}^{2}}+2\alpha\int_{\mu_{T}}^{\infty}8e^{-\left(y-\mu_{T}\right)^{2}/\left(2\sigma_{T}^{2}\right)}ye^{\alpha y^{2}}dy
\end{aligned}$$

where the last line comes from two applications of inequality (15) since $0<y-\mu_{T}<y+\mu_{T}$. It is now evident that the last integral above is convergent as soon as $\mu_{T}$ is finite and $\alpha<1/\left(2\sigma_{T}^{2}\right)$, so $\alpha$ can indeed be chosen positive, finishing the proof of (16), assuming $\mu_T$ is indeed finite. This last issue follows immediately from (14), which we now prove.

Establishing (14) follows a somewhat classical argument, using the so-called Dudley-Fernique theorem (see [1]), valid also in the sub-Gaussian context (see [18, Chapter 12]). We now provide the details. According to Lemma 2, as mentioned before, $X$ is sub-Gaussian with respect to the canonical metric $\delta^2(s,s') := K|s-s'|^{2H}$. Let now $N(\varepsilon)$ be the smallest number of balls of radius no more than $\varepsilon$ in this canonical metric that are needed to cover the interval $[0,t]$. Here we clearly see that

$$N(\varepsilon) \leq 1 + 2^{-1} t K^{1/(2H)} \varepsilon^{-1/H}.$$

It holds that there exists a universal constant $K_u$ such that

$$\mu_t := \mathbf{E}\left[\sup_{s\in[0,t]} |X_s|\right] \leq K_u \int_0^D \sqrt{\log N(\varepsilon)} d\varepsilon$$

where $D$ is the diameter of $[0,t]$ in the metric $\delta$, i.e. $D = \sqrt{K} t^H$. It follows that

$$\begin{aligned}
\mu_t &\leq K_u \int_0^{\sqrt{K}t^H} \sqrt{\log\left(1 + 2^{-1} t K^{1/(2H)} \varepsilon^{-1/H}\right)} d\varepsilon \\
&\leq H^{-1/2} K_u \int_0^{\sqrt{K}t^H} \sqrt{\log\left(1 + 2^{-1} t^H \sqrt{K} \varepsilon^{-1}\right)} d\varepsilon \\
&= K_u H^{-1/2} \sqrt{K} t^H \int_0^1 \sqrt{\log\left(1 + 2^{-1}\eta\right)} d\eta,
\end{aligned}$$

which is the statement (14), finishing the proof of the lemma. ■

**Remark 2** *The reader of [22] will notice some similarities between our development and that article, in particular the use of the Girsanov theorem. One technical task we have which cannot be found in [22] stems from our need to establish and study the sub-Gaussian property of $X$. We take this route because it is a natural and direct way to obtain the estimates we need to study the MLE as the quantity $d\tilde{P}/dP$ under $P$; quantities of the form $b(X_s)$ – as opposed to $b(B_s^H)$ – appear, making it necessary to look at $X$ as a sub-Gaussian process under $P$, as opposed to being able to exploit the Gaussian property of $B^H$ or $\tilde{B}^H$ under $\tilde{P}$.*

A more explicit, and more immediately obtained, estimate on the expected value of $X$ is the following. It refers only to the constant $C$ in hypothesis (7).

**Lemma 4** *We have for all $t \geq 0$,*

$$\mathbf{E}[|X_t|] \leq \left(Ct + \sqrt{\frac{2}{\pi}} t^H\right) e^{Ct}.$$

**Proof.** Let $M(t) := \mathbf{E}[|X_t|]$. We have from the definition (6) of $X$ that

$$\begin{aligned}
M(t) &= \mathbf{E}\left|\int_0^t b(X_s)\, ds + B_t^H\right| \\
&\leq \sqrt{\frac{2}{\pi}} t^H + C \int_0^t (1 + \mathbf{E}|X_s|)\, ds \\
&= Ct + \sqrt{\frac{2}{\pi}} t^H + C \int_0^t M(s)\, ds,
\end{aligned}$$

from which Gronwall's lemma implies

$$M(t) \leq \left( Ct + \sqrt{\frac{2}{\pi}} t^H \right) e^{Ct},$$

proving the lemma. ■

**Proof of Proposition 1.** Let

$$h(t) = \int_0^t b(X_s)\, ds.$$

We prove that the process $h$ satisfies i) and ii) of Theorem 1. Note first that the application of the operator $K_H^{-1}$ preserves the adaptability. We treat separately the cases when $H$ is bigger or less than one half.

**The case** $H < 1/2$. To prove i), we only need to show that $Q \in L^2([0,T])$ **P**-a.s. Indeed i) is equivalent to the following, almost-surely:

$$\begin{aligned} & h \in I_+^{H+1/2}\left(L^2([0,T])\right) \\ & \Longleftrightarrow K_H^{-1} h \in K_H^{-1}\left(I_+^{H+1/2}\left(L^2([0,T])\right)\right). \end{aligned}$$

Then using the isomorphism property of $K_H$ we see that i) is equivalent to $K_H^{-1}h \in L^2([0,T])$, which means $Q \in L^2([0,T])$ a.s. by definition. Now using relation (4) we thus have, for some constant $C(H)$ which may change from line to line, using the hypothesis $|b(x)| \leq C(1+|x|)$, for all $s \leq T$,

$$\begin{aligned} |Q_s| &\leq C(H) s^{H-1/2} \left| \int_0^s (s-u)^{-1/2-H} u^{1/2-H} b(X_u)\, du \right| \\ &\leq C(H) \left( 1 + \sup_{u \leq s} \mathbf{E}|X_u| \right), \end{aligned} \tag{17}$$

which we can rewrite, thanks to Lemma 4, as

$$\sup_{s \leq T} |Q_s| \leq C(H,T) \left( 1 + \sup_{s \leq T} |X(s)| \right),$$

which, thanks to inequality (14), is of course much stronger than $Q \in L^2([0,T])$ a.s., since $\sup_{s \leq T} |X(s)|$ has moments of all orders.

To prove ii) it suffices to show that there exists a constant $\alpha > 0$ such that

$$\sup_{s \leq T} \mathbf{E}\left(\exp(\alpha Q_s^2)\right) < \infty.$$

Indeed, one can invoke an argument used by Friedman in [9], Theorem 1.1, page 152, showing that this condition implies the so-called Novikov condition (see [24]), itself implying ii). Since $Q$ satisfies (17), the above exponential moment is a trivial consequence of inequality (16).

**The case** $H > 1/2$. Using formula (5) we have in this case that

$$\begin{aligned} Q_s &= c(H) \left[ s^{\frac{1}{2}-H} b(X_s) + \left( H - \frac{1}{2} \right) s^{H-\frac{1}{2}} \int_0^s \frac{b(X_s) s^{\frac{1}{2}-H} - b(X_u) u^{\frac{1}{2}-H}}{(s-u)^{H+\frac{1}{2}}} du \right] \\ &= c(H) \left[ s^{\frac{1}{2}-H} b(X_s) + \left( H - \frac{1}{2} \right) s^{H-\frac{1}{2}} b(X_s) \int_0^s \frac{s^{\frac{1}{2}-H} - u^{\frac{1}{2}-H}}{(s-u)^{H+\frac{1}{2}}} du \right. \\ & \left. + \left( H - \frac{1}{2} \right) s^{H-\frac{1}{2}} \int_0^s \frac{b(X_s) - b(X_u)}{(s-u)^{H+\frac{1}{2}}} u^{\frac{1}{2}-H} du \right] \end{aligned}$$

and using the fact that

$$\int_0^s \left(s^{\frac{1}{2}-H} - u^{\frac{1}{2}-H}\right)(s-u)^{-H-\frac{1}{2}} du = c(H) s^{1-2H}$$

we get

$$\begin{aligned}|Q_s| &\leq c(H)\left(s^{\frac{1}{2}-H}|b(X_s)| + s^{H-\frac{1}{2}} \int_0^s \frac{b(X_s)-b(X_u)}{(s-u)^{H+\frac{1}{2}}} u^{\frac{1}{2}-H} du\right) \\ &:= A(s) + B(s).\end{aligned}$$

The first term $A(s)$ above can be treated as in [22], proof of Theorem 3, due to our Lipschitz assumption on $b$. We will obtain that for every $\lambda > 1$,

$$\mathbf{E}\left(\exp(\lambda \int_0^t A_s^2 ds)\right) < \infty. \tag{18}$$

To obtain the same conclusion for the second summand $B(s)$ we can still use the argument in [22] and our Lemma 3. We only need to point out the fact that, since the process $X$ has obviously Hölder-continuous paths of order $\delta > H$, the random variable $G = \sup_{0 \leq u < s \leq T} |X_s - X_u|/|u-s|^{H-\varepsilon}$ is almost surely finite for every $\varepsilon > 0$. Moreover, since the process $X$ is sub-Gaussian, it follows from a Fernique-type theorem (see the comments in [18], pag. 321-323, but a direct proof, as in Lemma 3 is not difficult) that $G$ has a quadratic exponential moment $\mathbf{E}\left(\exp(cG^2)\right) < \infty$. Since for every fixed $\lambda > 1$ and for $\alpha < 1$, there exists a constant $K = K(c, \lambda)$ such that

$$\exp(\lambda G^{2\alpha}) \leq K \exp(cG^2).$$

combining the above comments with the calculations contained in [22], we obtain a similar inequality to (18) for the term $B(s)$. Now ii) is a consequence of the Novikov criterion (see [24]).

The properties i) and ii) are established for both cases, and we may use Theorem 1. The obtention of expression (12) for the MLE follows a standard calculation, since (recall that $P_\theta$ is the probability measure induced by $(X_s)_{0 \leq s \leq t}$),

$$F(\theta) := \log \frac{dP_\theta}{dP_0} = -\theta \int_0^t Q_s dW_s - \frac{\theta^2}{2} \int_0^t Q_s^2 ds. \tag{19}$$

■

We finish this section with some remarks that will relate our construction to previous works ([15], [14], [25]). More details about these links are given in Section 5.

**Alternative form of the MLE.** By (6) we can write, by integrating the quantity $K^{*,-1} 1_{[0,t]}(s)$ for $s$ between 0 and $t$,

$$\int_0^t \left(K^{*,-1} 1_{[0,t]}(\cdot)\right)(s) dX_s = \theta \int_0^t \left(K^{*,-1} 1_{[0,t]}(\cdot)\right)(s) b(X_s) ds + W_t. \tag{20}$$

On the other hand, by (6) again,

$$X_t = \int_0^t K(t,s) dZ_s \tag{21}$$

where $Z$ is given by (9). Therefore, we have the equality

$$\int_0^t \left(K^{*,-1} 1_{[0,t]}(\cdot)\right)(s) dX_s = Z_t. \tag{22}$$

By combining (20) and (22) we obtain

$$\int_0^t K_H^{-1}\left(\int_0^\cdot b(X_r) dr\right)(s) ds = \int_0^t \left(K^{*,-1} 1_{[0,t]}(\cdot)\right)(s) b(X_s) ds$$

and thus the function

$$t \to \int_0^t \left(K^{*,-1} 1_{[0,t]}(\cdot)\right)(s) b(X_s) ds$$

is absolutely continuous with respect to the Lebesgue measure and

$$Q_t = \frac{d}{dt} \int_0^t \left(K^{*,-1} 1_{[0,t]}(\cdot)\right)(s) b(X_s) ds. \tag{23}$$

By (9) we get that the function (19) can be written as

$$F(\theta) = -\theta \int_0^t Q_s dZ_s + \frac{\theta^2}{2} \int_0^t Q_s^2 ds.$$

As a consequence, the maximum likelihood estimator $\theta_t$ has the equivalent form

$$\theta_t = \frac{\int_0^t Q_s dZ_s}{\int_0^t Q_s^2 ds}. \tag{24}$$

# 4 Asymptotic behavior of the maximum likelihood estimator

This section is devoted to studying the strong consistency of the MLE (12). A similar result has been proven in the case $b(x) \equiv x$ and $H > \frac{1}{2}$ in [14]. We propose here a proof of strong consistency for a class of functions $b$ which contains significant non-linear examples. By replacing (9) in (24), we obtain that

$$\theta_t - \theta = \frac{\int_0^t Q_s dW_s}{\int_0^t Q_s^2 ds}$$

with $Q$ given by (11) or (23). To prove that $\theta_t \to \theta$ almost surely as $t \to \infty$ (which means by definition that the estimator $\theta_t$ is strongly consistent), by the strong law of large numbers we need only show that

$$\lim_{t \to \infty} \int_0^t Q_s^2 ds = \infty \text{ a.s. }. \tag{25}$$

To prove that $\lim_{t\to\infty} \int_0^t Q_s^2 ds = \infty$ in a non-linear case, it is necessary to make some assumption of non-degeneracy on the behavior of $b$. In order to illustrate our method using the least amount of technicalities, we will restrict our study to the case where the function $|b|$ satisfies a simple probabilistic estimate with respect to fractional Brownian motion.

**(C)** There exist positive constants $t_0$ and $K_b$, both depending only on $H$ and the function $b$, such that for all $t \geq t_0$ and all $\varepsilon > 0$, we have $\tilde{\mathbf{P}}\left[\left|Q_t\left(\tilde{\omega}\right)\right| / \sqrt{t} < \varepsilon\right] \leq \varepsilon K_b$, where under $\tilde{\mathbf{P}}$, $\tilde{\omega}$ has the law of fractional Brownian motion with parameter $H$.

## 4.1 The case $H < \frac{1}{2}$

In this paragraph we prove the following result.

**Theorem 2** *Assume that $H < 1/2$ and that Condition (C) holds. Then the estimator $\theta_t$ is strongly consistent, that is,*

$$\lim_{t \to \infty} \theta_t = \theta \text{ almost surely.}$$

Before proving this theorem, we discuss Condition (C). To understand this condition, we first note that with $\mu_H^t$ the positive measure on $[0,t]$ defined by $\mu_H^t(dr) = (r/t)^{1/2-H}(t-r)^{-1/2-H}\,dr$, according to the representation (4), we have

$$Q_t = \int_0^t \mu_H^t(ds)\, b(\tilde{\omega}_s)$$

and therefore, by the change of variables $r = s/t$,

$$\begin{aligned} \frac{Q_t}{\sqrt{t}} &= \int_0^1 \mu_H^1(dr)\, \frac{b(\tilde{\omega}_{tr})}{t^H} \\ &\stackrel{\mathcal{D}}{=} \int_0^1 \mu_H^1(dr)\, \frac{b\left(t^H\tilde{\omega}_r\right)}{t^H}, \end{aligned} \tag{26}$$

where the last inequality is in distribution under $\tilde{\mathbf{P}}$. Now if $b$ has somewhat of a linear behavior, we can easily imagine that $b\left(t^H\tilde{\omega}_r\right)/t^H$ will be of the same order as $b(\tilde{\omega}_r)$. Therefore $Q_t/\sqrt{t}$ should behave, in distribution for fixed $t$, similarly to the universal random variable $\int_0^1 \mu_H^1(dr)\, b(\tilde{\omega}_r)$ (whose distribution depends only on $b$ and $H$). Generally speaking, if this random variable has a bounded density, condition (C) will follow. However, thinking again of the case where $b$ increases roughly linearly, this random variable will be similar to the one where $b$ is replaced by the identity function, for which it is known, by the Arcsine law, that the random variable has indeed a bounded density.

Leaving aside these vague considerations, let us now give some specific examples of situations where (C) holds. Note that none of the examples in the "non-linear class" defined below (see Condition (28)) imply any sort of local regularity for $b$.

- **Linear case**. If $b(x) = cx$ for some constant $c$, we do indeed get

$$\frac{Q_t}{\sqrt{t}} \stackrel{\mathcal{D}}{=} c\int_0^1 \mu_H^1(dr)\, \tilde{\omega}_r. \tag{27}$$

  This random variable is centered and Gaussian, with a positive finite variance, and therefore its density near the origin is bounded above. Property (C) follows trivially.

- **A nonlinear class of examples**. Assume $xb(x)$ has a constant sign for all $x \in \mathbf{R}_+$ and a constant sign for all $x \in \mathbf{R}_-$. Assume

$$|b(x)/x| = c + h(x) \tag{28}$$

  for all $x$, where $c$ is a fixed positive constant, and $\lim_{x\to\infty} h(x) = 0$. *

**Lemma 5** *The second class of nonlinear examples above (condition (28)) satisfies condition (C).*

**Proof.** Define

$$V_t := t^{-H}\int_0^1 \mu_H^1(dr)\, b\left(t^H\tilde{\omega}_r\right) \stackrel{\mathcal{D}}{=} \frac{Q_t}{\sqrt{t}},$$

Our assumption implies four different scenarios in terms of the constant sign of $b$ on $\mathbf{R}_+$ or $\mathbf{R}_-$. We will limit this proof to the situation where $b(x)$ has the same sign as $x$. The other three cases are either similar or easier. Thus we have $b(x) = cx + xh(x)$. Also define

$$\begin{aligned} V_* &= \int_0^1 \mu_H^1(dr)\, c\tilde{\omega}_r, \\ E_t &= V_t - V_*, \end{aligned}$$

*Note that this condition is less restrictive than saying $b$ is asymptotically linear, since it covers the family $b(x) = cx + (|x| \wedge 1)^\alpha$ for any $\alpha \in (0,1)$.

so that $V_t = E_t + V_*$. Now

$$\begin{aligned} E_t &= \int_0^1 \mu_H^1 (dr) \left( b \left( t^H \tilde{\omega}_r \right) t^{-H} - c\tilde{\omega}_r \right) \\ &= t^{-H} \int_0^1 \mu_H^1 (dr) \left( b \left( t^H \tilde{\omega}_r \right) - ct^H \tilde{\omega}_r \right) \\ &= \int_0^1 \mu_H^1 (dr) \, \tilde{\omega}_r h \left( t^H \tilde{\omega}_r \right). \end{aligned}$$

For $\tilde{\mathbf{P}}$-almost every $\tilde{\omega}$, the function $\tilde{\omega}$ is continuous, and thus bounded on $[0,1]$. Therefore, $\tilde{\mathbf{P}}$-almost surely, uniformly for every $r \in [0,1]$, $\lim_{t\to\infty} h \left( t^H \tilde{\omega}_r \right) = 0$. Thus the limit is preserved after integration against $\mu_H^1$, which means that $\tilde{\mathbf{P}}$-almost surely, $\lim_{t\to\infty} E_t = 0$. Now fix $\varepsilon > 0$. There exists $t_0 (\tilde{\omega})$ finite $\tilde{\mathbf{P}}$-almost surely such that for any $t > t_0 (\tilde{\omega})$, $|E_t| \leq \varepsilon$. Thus if $|V_t| < \varepsilon$, we must have $|V_*| = |V_* + E_t - E_t| = |V_t - E_t| \leq |V_t| + |E_t| \leq 2\varepsilon$. This proves that $\tilde{\mathbf{P}}$-almost surely,

$$\limsup_{t\to\infty} \{|V_t| < \varepsilon\} \leq \{|V_*| < 2\varepsilon\}$$

and therefore,

$$\begin{aligned} &\limsup_{t\to\infty} \tilde{\mathbf{P}} \left[ |V_t| < \varepsilon \right] \\ &\leq \tilde{\mathbf{P}} \left[ \limsup_{t\to\infty} \{|V_t| < \varepsilon\} \right] \leq \tilde{\mathbf{P}} \left[ |V_*| < 2\varepsilon \right]. \end{aligned}$$

Now we invoke the fact that $V_*$ is precisely the random variable studied in the first, linear, example, so that $\tilde{\mathbf{P}} \left[ |V_*| < 2\varepsilon \right] \leq 2K\varepsilon$ for some constant $K$ depending only on $c$ and $H$, finishing the proof of the lemma. ■

Here are some additional examples.

- **Additive constant drift**. If we replace $b$ in the previous example by $b(x) + C$ where $C$ is an arbitrary constant, then Condition (C) still holds. Indeed, in the proof of the lemma, this only adds a term to $E_t$ which converges to 0 deterministically.
- It would be satisfying to be able to prove Condition (C) for $b$'s satisfying a cone condition, such as: $b(x)/x$ bounded above and below by constants of the same sign for all $x \in \mathbf{R}_+$, and of the same sign for all $x \in \mathbf{R}_-$. Although we believe this fact, we have not yet been able to prove it.
- The most interesting subcase of the examples above is that in which $b(x)$ has the opposite sign of $x$. It is also the case for which statement (25) might be most likely to fail for physical reasons. In this case the solution $X$ of the non-linear Langevin equation (6) should have a similar behavior to that of the stable fBm-OU process, i.e. $\theta < 0$ and $b(x) = x$.

**Proof of Theorem 2.** Since we only want to show that (25) holds, and since $\int_0^t |Q_s|^2 \, ds$ is increasing, it is sufficient to satisfy condition (25) for $t$ tending to infinity along a sequence $(t_n)_{n\in\mathbf{N}}$. We write, according to the representation (4), for each fixed $t \geq 0$,

$$I_t = I_t(X) := \int_0^t |Q_s(X)|^2 \, ds = \int_0^t \left| \int_0^s \mu_H^s (dr) \, b(X_r) \right|^2 ds$$

where $X$ is the solution of the Langevin equation (6) and the positive measure $\mu_H^s$ is defined by $\mu_H^s(dr) = (r/s)^{1/2-H} (s-r)^{-1/2-H} \, dr$. Recall from the Girsanov Theorem 1 applied to $X$, that with

$$\eta_T = \exp\left( -\int_0^T Q_s(X) \, dW_s - \frac{1}{2} \int_0^T |Q_s(X)|^2 \, ds \right)$$

where $W$ is standard Brownian motion under $\mathbf{P}$, we have that under the probability measure $\tilde{\mathbf{P}}$ defined by its Radon-Nikodym derivative

$$\left.\frac{d\tilde{\mathbf{P}}}{d\mathbf{P}}\right|_{\mathcal{F}_T^X} = \eta_T$$

for all $T \geq 0$, $X$ is a fractional Brownian motion with parameter $H$. Moreover, since $\eta_T$ is a true martingale, this Girsanov transformation can be reversed. See [24, Theorem VIII.1.7]: with $L_t = \int_0^T Q_s(X)\, dW_s$, we can write that

$$\left.\frac{d\mathbf{P}}{d\tilde{\mathbf{P}}}\right|_{\mathcal{F}_T^X} = \tilde{\eta}_T,$$

where under $\tilde{\mathbf{P}}$, $\tilde{\eta}$ can be written as the exponential martingale

$$\tilde{\eta}_t = \exp\left(\tilde{L}_t - \frac{1}{2}\left\langle \tilde{L}\right\rangle_t\right)$$

for some martingale $\tilde{L}$ under $\tilde{\mathbf{P}}$ satisfying

$$\left\langle \tilde{L}\right\rangle_t = \langle L\rangle_t = I_t(X) = \int_0^t |Q_s(X)|^2\, ds.$$

Here, since $X$ is still a fractional Brownian motion with parameter $H$ under $\tilde{\mathbf{P}}$, we will use the notation $\tilde{\omega}$ for $X$, to signify that $X$ does not have the law of $X$ under $\mathbf{P}$.

Thence consider a sequence of constants $(\beta_n)_{n=0}^{\infty}$ which will be chosen later. Using the trivial fact that $\mathbf{1}_{(-\infty,a]}(x) \leq \exp(-\lambda x)\exp(\lambda a)$, and Hölder's inequality, we can write

$$\begin{aligned}
\mathbf{P}\left[I_t(X) < \beta_t\right] &= \tilde{\mathbf{E}}\left[\mathbf{1}_{I_t(\tilde{\omega})<\beta_t}\tilde{\eta}_t(\tilde{\omega})\right] \\
&\leq \exp(\lambda\beta_t)\,\tilde{\mathbf{E}}\left[\exp\left(-\lambda I_t(\tilde{\omega}) + \tilde{L}_t - 2^{-1}I_t(\tilde{\omega})\right)\right] \\
&= \exp(\lambda\beta_t)\,\tilde{\mathbf{E}}\left[\exp\left(-(\lambda-\nu)I_t(\tilde{\omega}) + \tilde{L}_t - \left(\nu + 2^{-1}\right)I_t(\tilde{\omega})\right)\right] \\
&\leq \exp(\lambda\beta_t)\,\tilde{\mathbf{E}}\left[\exp\left(-p(\lambda-\nu)I_t(\tilde{\omega})\right)\right]^{1/p} \\
&\times \tilde{\mathbf{E}}\left[\exp\left(q\tilde{L}_t - q\left(\nu+2^{-1}\right)I_t(\tilde{\omega})\right)\right]^{1/q}.
\end{aligned} \tag{29}$$

where $0 < \nu < \lambda$ are arbitrary fixed positive constants. We may now choose the conjugate Hölder exponents $p^{-1} + q^{-1} = 1$. It will be convenient to allow $p > 1$ to be as close to 1 as possible, hence $q$ will be very large. We also want $q^2/2 = q\left(\nu + 2^{-1}\right)$. This forces us to take $\nu = 2^{-1}(q-1)$, which will also be very large. We then take $\lambda$ to be a fixed value $> \nu$. The choice on $q$ means that the last term in (29) above is equal to 1. Hence, letting

$$y := p(\lambda - \nu)$$

we have

$$\mathbf{P}\left[I_t(X) < \beta_t\right] \leq \exp(\lambda\beta_t)\,\tilde{\mathbf{E}}\left[\exp\left(-yI_t(\tilde{\omega})\right)\right]^{1/p}.$$

To evaluate the above expectation, since $I_t(\tilde{\omega})$ is a random variable in the interval $(0,1)$, we first write

$$\begin{aligned}
\tilde{\mathbf{E}}\left[\exp\left(-yI_t(\tilde{\omega})\right)\right] &= \int_0^1 \tilde{\mathbf{P}}\left[\exp(-yI_t) > x\right] dx \\
&= \int_0^{\infty} e^{-z}\tilde{\mathbf{P}}\left[I_t < \frac{z}{y}\right] dz.
\end{aligned}$$

Now let $t = t_n = n^k$ for some fixed $k \geq 1$, and for all $n \in \mathbf{N}$. We also introduce a positive sequence $b_n$ whose definition will be motivated below. We write

$$\begin{aligned} I_{t_n} &= \int_0^{t_n} |Q_s(\tilde{\omega})|^2 \, ds \\ &\geq \int_{t_n - b_n}^{t_n} |Q_s(\tilde{\omega})|^2 \, ds \\ &\geq 2b_n |Q_{t_n}(\tilde{\omega})|^2 - 2\int_{t_n - b_n}^{t_n} |Q_{t_n}(\tilde{\omega}) - Q_s(\tilde{\omega})|^2 \, ds \\ &\geq 2b_n \left( |Q_{t_n}(\tilde{\omega})|^2 - \sup_{s \in [t_n - b_n, t_n]} |Q_{t_n}(\tilde{\omega}) - Q_s(\tilde{\omega})|^2 \right). \end{aligned} \tag{30}$$

We will need the result of the next lemma in order to control the variations of $Q$ on the interval $[t_n - b_n, t_n]$. It can be considered as a consequence of the fact that $V_t := t^{-1/2} Q_t$ is an asymptotically sub-stationary process in the second Gaussian chaos, although the proof we present below only requires the use of moments of $V$ via the Kolmogorov continuity lemma, because of the fact that we are working in the Hölder scale of fractional Brownian regularity.

**Lemma 6** *Let $V_t(\tilde{\omega}) := t^{-1/2} Q_t(\tilde{\omega})$. The process $V$ is $\tilde{\mathbf{P}}$-almost surely continuous. Moreover, if $b_n \geq 1$ and $b_n \ll t_n$, then for any $m > 2$, there exists a constant $C_{m,H}$ such that*

$$\tilde{\mathbf{E}} \left[ \sup_{s,t \in [t_n - b_n, t_n]} |V_t - V_s|^m \right] \leq C_{m,H,b} \left( \frac{b_n}{t_n} \right)^{Hm}.$$

The proof of this lemma will be given further below. We now use it as follows. Let $x = z/(2y)$. Let $Z_n^2 = \sup_{s \in [t_n - b_n, t_n]} |Q_{t_n}(\tilde{\omega}) - Q_s(\tilde{\omega})|^2 / t_n$. We also introduce another positive sequence $a_n$. From (30), and from Condition (C), we have

$$\begin{aligned} &\tilde{\mathbf{P}} \left[ I_{t_n} < \frac{z}{y} \right] \\ &\leq \tilde{\mathbf{P}} \left[ |Q_{t_n}(\tilde{\omega})|^2 / t_n - |Z_n|^2 < \frac{x}{b_n} \right] \\ &= \tilde{\mathbf{P}} \left[ |Q_{t_n}(\tilde{\omega})|^2 / t_n - n^{1/2} |Z_n|^2 < \frac{x}{b_n}; |Z_n|^2 \geq a_n \right] \\ &+ \tilde{\mathbf{P}} \left[ |Q_{t_n}(\tilde{\omega})|^2 / t_n - |Z_n|^2 < \frac{x}{b_n}; |Z_n|^2 < a_n \right] \\ &\leq \tilde{\mathbf{P}} \left[ |Z_n|^2 \geq a_n \right] + \tilde{\mathbf{P}} \left[ |Q_{t_n}(\tilde{\omega})|^2 / t_n < \frac{x}{b_n} + a_n \right] \\ &\leq C_m \left( \frac{b_n}{t_n} \right)^{Hm} (a_n)^{-m} + K_b \sqrt{\frac{x}{b_n} + a_n}. \end{aligned}$$

We now let $N > 1$, and choose $t_n = n^k$, $b_n = n^{2N}$, $a_n = n^{-2N}$, $k > 2N(1+H)/H$ and $m$ large enough so that we can write, for some other constant $C_m$,

$$\tilde{\mathbf{P}} \left[ I_{t_n} < \frac{z}{y} \right] \leq C_m K_b \sqrt{x+1} \frac{1}{n^N}.$$

Hence we finally obtain

$$\begin{aligned}\mathbf{P}\left[I_{t_n}(X)<\beta_{t_n}\right] &\leq \exp\left(\lambda\beta_{t_n}\right)\left[\int_0^\infty e^{-z}\tilde{\mathbf{P}}\left[I_t<2x\right]dz\right]^{1/p}\\ &\leq \left(C_mK_b\right)^{1/p}\frac{1}{n^{pN}}\exp\left(\lambda\beta_t\right)\left(\int_0^\infty e^{-z}\sqrt{1+\frac{z}{2p\left(\lambda-\nu\right)}}\right)^{1/p}\\ &= C\left(m,p,b\right)\frac{1}{n^{pN}}\exp\left(\lambda\beta_t\right)\end{aligned}$$

where the constant $C(m,p,b)$ can be chosen as depending only on $m$, $p$, and the function $b(\cdot)$ via the constant $K_b$ in condition (C). Note that $N$ can be made arbitrarily large by taking $m$ large enough, but there is little to gain by assuming that $pN$ is much greater than 1.

In order for the above bound to be summable in $n$, it is sufficient to choose $\beta_{t_n}=(pN/2\lambda)\log n$. Therefore, by the Borel-Cantelli lemma, there exists a random integer $n_0(\tilde{\omega})$ depending on the function $b(\cdot)$ via the constant $K_b$, and depending on the constants $p$, $\lambda$, and $N$, such that for all $n>n_0(\tilde{\omega})$,

$$I_{n^k}(\tilde{\omega})\geq\frac{pN}{2\lambda}\log n,$$

where the constants $p$ and $\lambda$ are chosen as described in the lines following inequality (29), and $N$ can be chosen arbitrarily large, implying the result (25) and the Theorem. ■

**Proof of Lemma 6.** To apply the Kolmogorov continuity lemma (see [24, Theorem I.2.1]), we must evaluate the moments of the increments of $V$: let $m>2$ and $s<t$ with $s,t\in[t_n-b_n,t_n]$; abbreviate $\mu:=\mu_H^1$; let $c_H$ denote the mass of $\mu$, or other constants depending only on $H$. We have

$$\begin{aligned}&\tilde{\mathbf{E}}\left[\left|V_t-V_s\right|^m\right]\\ &=\tilde{\mathbf{E}}\left[\left|\int_0^1\mu(dr)\left(\frac{b(\tilde{\omega}_{tr})}{t^H}-\frac{b(\tilde{\omega}_{sr})}{s^H}\right)\right|^m\right]\\ &\leq\left(c_H\right)^m\int_0^1\mu(dr)\,\tilde{\mathbf{E}}\left[\left|\frac{b(\tilde{\omega}_{tr})}{t^H}-\frac{b(\tilde{\omega}_{sr})}{s^H}\right|^m\right]\\ &\leq\left(c_H\right)^m\int_0^1\mu(dr)\left\{t^{-Hm}\tilde{\mathbf{E}}\left[\left|b(\tilde{\omega}_{tr})-b(\tilde{\omega}_{sr})\right|^m\right]+\left(\frac{|t-s|}{s^{1+H}}\right)^m\tilde{\mathbf{E}}\left[\left|b(\tilde{\omega}_{sr})\right|^m\right]\right\}.\end{aligned}$$

Now we use the fact that $b$ is Lipshitz, so for some constant $b_0$, $|b(x)-b(y)|\leq b_0|x-y|$ and $|b(x)|\leq b_0(1+|x|)$; and we use the Gaussian law of $\tilde{\omega}$. For some constant $C_{m,H,b}$ which may change from line to line,

$$\begin{aligned}\tilde{\mathbf{E}}\left[\left|V_t-V_s\right|^m\right] &\leq C_{m,H,b}\int_0^1\mu(dr)\left\{t^{-Hm}b_0^m|t-s|^{mH}+\left(\frac{|t-s|}{s^{1+H}}\right)^m s^{mH}b_0^m\right\}\\ &=C_{m,H,b}\int_0^1\mu(dr)\left\{\left|\frac{t-s}{t}\right|^{mH}+\left(\left|\frac{t-s}{s}\right|\right)^m\right\}.\end{aligned}$$

Now since we assume that $b_n\ll t_n$, it follows that $(t-s)/s\leq 2(t-s)/t$ for $n$ large enough. Hence we have proved that

$$\begin{aligned}\tilde{\mathbf{E}}\left[\left|V_t-V_s\right|^m\right] &\leq C_{m,H,b}\int_0^1\mu(dr)\left|\frac{t-s}{t_n}\right|^{mH}\\ &=C_{m,H,b}\frac{|t-s|^{mH}}{t_n^{mH}}.\end{aligned}$$

If we now define $U$ on the interval $[0,1]$ by $U_u = V_{t_n-b_n+ub_n}$, we see that $U$ satisfies

$$\tilde{\mathbf{E}}\left[|U_u - U_v|^m\right] \le C_{m,H,b}\left(b_n/t_n\right)^{mH}|u-v|^{Hm}.$$

Temporarily normalizing $U$ by the constant $C_{m,H,b}\left(b_n/t_n\right)^{mH}$, and applying [24, Theorem I.2.1], we finally get that the unnormalized $U$ has a continuous version, and for any $\alpha < H - 1/m$ and a universal constant $K$

$$\tilde{\mathbf{E}}\left[\sup_{u,v\in[0,1]}\left(\frac{|U_u - U_v|}{|u-v|^{\alpha}}\right)^m\right] \le K C_{m,H,b}\left(b_n/t_n\right)^{mH}.$$

Since $1 \le |u-v|^{-\alpha}$, the statement of the Lemma follows. ■

## 4.2 The case $H > \frac{1}{2}$

Due to the fact that the function $Q$ is less regular in this case, we need to consider a more restrictive situation. We look specifically only at the nonlinear class of examples given in (28).

**Theorem 3** *Assume that $H > \frac{1}{2}$ and $b$ is of the form (28) with $h$ Lipshitz. Then the maximum likelihood estimator $\theta_t$ is strongly consistent.*

**Proof:** Recall from the proof of Proposition 1 that we can write

$$Q_t = c(H)t^{\frac{1}{2}-H}b(X_t) + c'(H)\int_0^t \mu_H^t(dr)\left(b(X_t) - b(X_r)\right). \tag{31}$$

We note that in this case the expression $\mu_H^t(dr)$ does not determine a measure, but we still use this notation to simplify the presentation; the Lipschitz assumption on $b$ and the Hölder property of $X$ do ensure the existence of the integral.

One can actually follow the proof in the case $H < \frac{1}{2}$ line by line. All we have to do here is to prove an equivalent of Lemma 6 on the variations of $Q$, this being the only point where the form of $Q$, which differs depending on whether $H$ is bigger or less than $1/2$, is used. We will illustrate how the second summand of $Q$ in (31) (which is the most difficult to handle) can be treated.

Denoting by $Q_t' = \int_0^t \mu_H^t(dr)\left(b(X_t) - b(X_r)\right)$, it holds

$$\begin{aligned} Q_t' t^{-\frac{1}{2}} &= t^{-\frac{1}{2}}\int_0^t \mu_H^t(dr)\left(b(X_t) - b(X_r)\right) \\ &\overset{\mathcal{D}}{=} \int_0^1 \mu_H^1(dr)\frac{b(t^H\tilde{\omega}_1) - b(t^H\tilde{\omega}_r)}{t^H}. \end{aligned}$$

where again $\overset{\mathcal{D}}{=}$ denotes equality in distribution. Now, if $V_t' := t^{-\frac{1}{2}}Q_t'$, we have

$$\begin{aligned} V_t' - V_s' &= \int_0^1 \mu_H^1(dr)t^{-H}\left(b(t^H\tilde{\omega}_1) - b(t^H\tilde{\omega}_r) - b(s^H\tilde{\omega}_1) + b(s^H\tilde{\omega}_r)\right) \\ &\quad + \left(t^{-H} - s^{-H}\right)\int_0^1 \mu_H^1(dr)\left(b(s^H\tilde{\omega}_1) - b(s^H\tilde{\omega}_r)\right) \\ &:= J_1 + J_2 \end{aligned}$$

and it holds that

$$\begin{aligned} \tilde{\mathbf{E}}\,|J_2|^m &\le |t^{-H} - s^{-H}|^m b_0^m s^{Hm}\tilde{\mathbf{E}}\left(\int_0^1 \mu_H^1(dr)\,|\tilde{\omega}_r - \tilde{\omega}_1|\right)^m \\ &\le C_{m,H,b}\left(\frac{|t-s|}{s^{1+H}}\right)^m s^{Hm} \end{aligned}$$

and it has been already proved that this is bounded by $C_{m,H,b}\left(|t-s|/t_n\right)^{mH}$. For the term denoted by $J_1$ one can obtain the same bound by observing that if $b$ has the form (28) with $h$ Lipshitz, then, for every $a, b$,

$$\left|b(t^H a) - b(t^H b) - b(s^H a) + b(s^H b)\right| \leq c|t^H - s^H||a-b|,$$

and thus

$$\begin{aligned}\tilde{\mathbf{E}}\left|J_1\right|^m &\leq C_{m,b,H}\left(\frac{|t^H - s^H|}{t^H}\right)^m \tilde{\mathbf{E}}\left(\int_0^1 \mu_H^1(dr)\left|\tilde{\omega}_r - \tilde{\omega}_1\right|\right)^m \\ &\leq C_{m.H,b}\left(\frac{|t-s|}{t_n}\right)^{mH}.\end{aligned}$$

■

# 5 The linear case

In this section we present some comments in the case when the drift $b$ is linear. We will assume that $b(x) \equiv x$ to simplify the presentation. In this case, the solution $X$ to equation (6) is the fractional Ornstein-Uhlenbeck process and it is possible to prove more precise results concerning the asymptotic behavior of the maximum likelihood estimator.

**Remark 3** *In [5], it is shown that there exists an unique almost surely continuous process $X$ that satisfies the Langevin equation (6) for any $H \in (0,1)$. Moreover the process $X$ can be represented as*

$$X_t = \int_0^t e^{\theta(t-u)} dB_u^H, \qquad t \in [0,T] \tag{32}$$

*where the above integral is a Wiener integral with respect with $B^H$ (which exists also as a pathwise Riemann-Stieltjes integral). It follows from the stationarity of the increments of $B^H$ that $X$ is stationary and the decay of its auto-covariance function is like a power function. The process $X$ is ergodic, and for $H > \frac{1}{2}$, it exhibits a long-range dependence.*

Let us briefly recall the method employed in [14] to estimate the drift parameter of the fractional OU process. Let us consider the function , for $0 < s < t \leq 1$,

$$k(t,s) = c_H^{-1} s^{\frac{1}{2}-H}(t-s)^{\frac{1}{2}-H} \text{ with } c_H = 2H\Gamma(\frac{3}{2}-H)\Gamma(H+\frac{1}{2}) \tag{33}$$

and let us denote its Wiener integral with respect to $B^H$ by

$$M_t^H = \int_0^t k(t,s) dB_s^H. \tag{34}$$

It has been proved in [20] that $M^H$ is a Gaussian martingale with bracket

$$\langle M^H \rangle_t := \omega_t^H = \lambda_H^{-1} t^{2-2H} \text{ with } \lambda_H = \frac{2H\Gamma(3-2H)\Gamma(H+\frac{1}{2})}{\Gamma(\frac{3}{2}-H)}. \tag{35}$$

The authors called $M^H$ the fundamental martingale associated to fBm. Therefore, observing the process $X$ given by (6) is the same thing as observing the process

$$Z_t^{KB} = \int_0^t k(t,s) dX_s$$

which is actually a semimartingale with the decomposition

$$Z_t^{KB} = \theta \int_0^t Q_s^{KB} d\omega_s^H + M_t^H \tag{36}$$

where

$$Q_t^{KB} = \frac{d}{d\omega^H} \int_0^t k(t,s) X_s ds, \qquad t \in [0,T]. \tag{37}$$

By using Girsanov's theorem (see [20] and [14]) we obtain that the MLE is given by

$$\theta_t := \theta_t^{KB} = \frac{\int_0^t Q_s^{KB} dZ_s^{KB}}{\int_0^t (Q_s^{KB})^2 d\omega_s^H}. \tag{38}$$

**Remark 4** *We can observe that our operator (12) or (24) coincides (possibly up to a multiplicative constant) with the one used in [14] and given by (38). Assume that $H < \frac{1}{2}$; the case $H > \frac{1}{2}$ is just a little more technical.*

**Proof**. Using relations (11) and (33) we can write

$$\begin{aligned} Q_t &= C(H) t^{H-\frac{1}{2}} \int_0^t s^{\frac{1}{2}-H} (t-s)^{-\frac{1}{2}-H} b(X_s) ds \\ &= C(H) t^{H-\frac{1}{2}} \int_0^t \frac{d}{dt} k(t,s) b(X_s) ds \\ &= C(H) t^{H-\frac{1}{2}} \frac{d}{dt} \int_0^t k(t,s) b(X_s) ds. \end{aligned}$$

It is not difficult to see that $\frac{d}{dt}\int_0^t k(t,s)b(X_s)ds = C(H)t^{1-2H}Q_t^{KB}$ and therefore

$$Q_t = C(H) t^{1/2-H} Q_t^{KB}. \tag{39}$$

On the other hand, it can be similarly seen that

$$Z_t^{KB} = C(H) \int_0^t s^{\frac{1}{2}-H} dZ_s. \tag{40}$$

and the estimation given by (38) and (24) coincide up to a constant. ■

To compute the expression of the bias and of the mean square error and to prove the strong consistency of the estimator, one has the option, in this explicit linear situation, to compute the Laplace transform of the quantity $\int_0^t (Q_s^{KB})^2 d\omega_s^H$. This is done for $H > 1/2$ in [14], Section 3.2, and the following properties are obtained:

- the estimator $\theta_t$ is strongly consistent, that is,

$$\theta_t \to \theta \text{ almost surely when } t \to \infty,$$

- the *bias* and the *mean square error* are given by
  - If $\theta < 0$, when $t \to \infty$, then

$$\mathbf{E}(\theta_t - \theta) \backsim \frac{2}{t}\ ,\ \mathbf{E}(\theta_t - \theta)^2 \backsim \frac{2}{t}|\theta|, \tag{41}$$

– If $\theta > 0$, when $t \to \infty$, then

$$\mathbf{E}(\theta_t - \theta) \backsim -2\sqrt{\pi \sin \pi H}\theta^{\frac{3}{2}} e^{-\theta t}\sqrt{t} \tag{42}$$

$$\mathbf{E}(\theta_t - \theta)^2 \backsim 2\sqrt{\pi \sin \pi H}\theta^{\frac{5}{2}} e^{-\theta t}\sqrt{t}. \tag{43}$$

**Remark 5** *It is very interesting to realize that the rate of convergence of the bias and of the mean square error does not depends on $H$. In fact the only difference between the classical case (see [19]) and the fractional case is the presence of the constant $\sqrt{\pi H}$ in (41), (42), and (43). It is natural to expect the same results if $H < \frac{1}{2}$. We prove it below.*

**Proposition 2** *If $H < \frac{1}{2}$, then (41), (42), and (43) hold.*

**Proof:** To avoid tedious calculations with fractional integrals and derivatives, we will take advantage of the calculations performed in [14] when $H > \frac{1}{2}$; nevertheless we believe that a direct proof is also possible. Actually the only moment when the authors of [14] use the fact that $H$ is bigger than $\frac{1}{2}$ is the computation of the process $Q$. By relations (21) and (23) we can write

$$\begin{aligned} Q_t &= \frac{d}{dt}\int_0^t \left(K^{*,-1}1_{[0,t]}(\cdot)\right)(s) X_s ds \\ &= \frac{d}{dt}\int_0^t \left(K^{*,-1}1_{[0,t]}(\cdot)\right)(s)\left(\int_0^s K(s,v)dZ_v\right)ds \\ &= \frac{d}{dt}\int_0^t \int_v^t \left(K^{*,-1}1_{[0,t]}(\cdot)\right)(s)K(s,v)dsdZ_v \end{aligned}$$

Note that from the formulas presented in Section 2, we have

$$\left(K^{*,-1}1_{[0,t]}(\cdot)\right)(s) = c(H)s^{\frac{1}{2}-H}\int_s^t u^{\frac{1}{2}-H}(u-s)^{-H-\frac{1}{2}}, \qquad H < \frac{1}{2},$$

$$\left(K^{*,-1}1_{[0,t]}(\cdot)\right)(s) = c(H)s^{\frac{1}{2}-H}\frac{d}{ds}\int_s^t u^{\frac{1}{2}-H}(u-s)^{-H+\frac{1}{2}}, \qquad H > \frac{1}{2}$$

To unify the notation, we write

$$\left(K^{*,-1}1_{[0,t]}(\cdot)\right)(s) = c(H)s^{\frac{1}{2}-H}\frac{d}{ds}\int_s^t u^{\frac{1}{2}-H}(u-s)^{-H+\frac{1}{2}}, \qquad H \in (0,1)$$

and we just observe that the constant $c(H)$ above is analytic with respect to $H$. Let us consider, for $v \le t$ a function $A(v,t)$ such that

$$\int_v^t A(v,s)ds = \int_v^t \left(K^{*,-1}1_{[0,t]}(\cdot)\right)(s)K(s,v)ds.$$

Then, obviously,

$$Q_t = \int_0^t A(t,v)dZ_v.$$

On the other hand, it has been proved in [14] (see relations (3.4) and (3.5) therein) that for $H > \frac{1}{2}$,

$$Q_t^{KB} = \int_0^t A^{KB}(t,v)dZ_v^{KB}$$

with

$$A^{KB}(t,s) = c(H)(t^{2H-1} + s^{2H-1}).$$

Using the relations between $Q$ and $Q^{KB}$ and between $Z$ and $Z^{KB}$ (see Remark 4), it follows that, for every $H > \frac{1}{2}$, and $s < t$,

$$A(s,t) = c(H)\left[\left(\frac{s}{t}\right)^{\frac{1}{2}-H} + \left(\frac{t}{s}\right)^{\frac{1}{2}-H}\right]. \tag{44}$$

We show that the above relation (44) is true for $H < \frac{1}{2}$ as well. We use an argument inspired by [7], proof of Theorem 3.1. We observe that the functions

$$H \in (0,1) \to A(s,t) \text{ and } H \in (0,1) \to c(H)\left[\left(\frac{s}{t}\right)^{\frac{1}{2}-H} + \left(\frac{t}{s}\right)^{\frac{1}{2}-H}\right]$$

are analytic with respect to $H$ and coincide on $(1/2, 1)$. Moreover, both are well-defined for every $H \in (0,1)$ (in fact it follows from [14] that $A$ is well-defined for $H > \frac{1}{2}$ and it is more regular for $H \leq \frac{1}{2}$). To conclude (44) for every $H \in (0,1)$, we invoke the fact that if $f, g : (a,b) \to \mathbb{R}$ are two analytic functions and the set $\{x \in (a,b); f(x) = g(x)\}$ has an accumulation point in $(a,b)$, then $f = g$.

As a consequence, (44) holds for every $H$ and this shows that

$$\int_0^t Q_s dZ_s = \int_0^t Q_s^{KB} dZ_s^{KB} = c(H)\left(Z_t^{KB}\int_0^t r^{2H-1} dZ_r^{KB} - t\right)$$

and all the calculations contained in [14], Sections 3.2, 4 and 5 hold for every $H \in (0,1)$. ■

# 6 Discretization

In this last section we present a discretization result which represents an important step towards a full understanding of how to implement the calculation of an MLE for an fBm-driven stochastic differential equation. As noted in the introduction, fully addressing this implementation issue is beyond the scope of the present article; it will be the subject of a separate paper.

For practical aspects, even if continuous information is available, to compute the MLE $\theta_t$ given by (12), the integrals appearing in it expression have to be approximated by Riemann sums. In [17] the folowing estimator is introduced in the case $H = 1/2$

$$\theta_{N,T} = -\frac{\sum_{i=0}^{N-1} Q_{t_i}(W_{t_{i+1}} - W_{t_i})}{\sum_{i=0}^{N-1} |Q_{t_i}|^2 (t_{i+1} - t_i)}$$

based on the discrete observations $t_0, ...., t_N$ . It has been proved in [17] that if $\delta_N \max_{i=1,..,N} |t_i - t_{i-1}|$ tends to zero, then $\theta_{N,T} \to \theta_T$ in probability as $N \to \infty$ and $\delta_N^{-1/2}(\theta_{N,T} - \theta_T)$ is bounded in probability. We also choose to work with the formula (12) by replacing the stochastic integral in the numerator and the Riemann integral in the denominator by their corresponding approximate Riemann sums, using discrete integer time. Specifically we define for any integer $n \geq 1$,

$$\bar{\theta}_n := -\frac{\sum_{m=0}^{n} Q_m (W_{m+1} - W_m)}{\sum_{m=0}^{n} |Q_m|^2}. \tag{45}$$

Our goal in this section is to prove that $\bar{\theta}_n$ is actually a consistent estimator for $\theta$. One could also consider the question of the discretization of $\theta_t$ using a fine time mesh for fixed $t$, and showing that this discretization converges almost surely to $\theta_t$; by time-scaling such a goal is actually equivalent to our own, and both are therefore stronger than the convergence in [17].

Let $\langle M\rangle_n$ denote the quadratic variation at time $n$ of a square-integrable martingale $M$. We introduce the following two martingales:

$$A_t := \int_0^t Q_s dW_s$$
$$B_t := \int_0^t Q_{[s]} dW_s$$

where $[s]$ denotes the integer part of $s$. We clearly have $B_n = \sum_{m=0}^{n-1} Q_m \left(W_{m+1} - W_m\right)$ and

$$\theta_n = \frac{A_n}{\langle A\rangle_n} \text{ and } \bar{\theta}_n = \frac{B_n}{\langle B\rangle_n}.$$

The following proposition defines a strategy for proving that $\bar{\theta}_n$ is a consistent estimator for $\theta$.

**Proposition 3** *Let $H \in (0,1)$. If there exists a constant $\alpha > 0$ such that*

- *$n^\alpha \langle A - B\rangle_n / \langle B\rangle_n$ is bounded almost surely for $n$ large enough,*
- *for all $k \geq 1$, for some constant $K > 0$, almost surely, for large $n$, $\langle B\rangle_n^k \geq K\mathbf{E}\left[\langle B\rangle_n^k\right]$,*
- *and for all $k > 1$, $\mathbf{E}\left[\left|\langle A - B\rangle_n\right|^k\right] \leq n^{-k\alpha}\mathbf{E}\left[\left|\langle B\rangle_n\right|^k\right]$,*

*then almost surely $\lim_{n\to\infty} \bar{\theta}_n = \theta$.*

**Proof**. By our Theorems 2 and 3, it is of course sufficient to prove that

$$\lim_{n\to\infty} \left(\bar{\theta}_n - \theta_n\right) = 0.$$

In preparation for this, we first note that by classical properties for quadratic variations, and using our hypothesis, for large enough $n$, we have

$$\begin{aligned} \left|\langle B\rangle_n - \langle A\rangle_n\right| &= \left|\langle (B-A), (B+A)\rangle_n\right| \\ &\leq \left|\langle B + A\rangle_n\right|^{1/2} \left|\langle B - A\rangle_n\right|^{1/2} \\ &\leq \sqrt{2} n^{-\alpha} \left|\langle B\rangle_n\right|^{1/2} \left|\langle A\rangle_n + \langle B\rangle_n\right|^{1/2}. \end{aligned} \tag{46}$$

Now we prove that (46) implies almost surely,

$$\lim_{n\to\infty} \frac{\langle A\rangle_n}{\langle B\rangle_n} = 1. \tag{47}$$

Indeed let $x_n = \langle A\rangle_n / \langle B\rangle_n$. Then we can write

$$\begin{aligned} |x_n - 1| &= \frac{\left|\langle B\rangle_n - \langle A\rangle_n\right|}{\langle B\rangle_n} \\ &\leq \sqrt{2} n^{-\alpha} \left|\langle B\rangle_n\right|^{-1/2} \left|\langle A\rangle_n + \langle B\rangle_n\right|^{1/2} \\ &= \sqrt{2} n^{-\alpha} \left|1 + x_n\right|^{1/2}. \end{aligned}$$

Let $\varepsilon > 0$ be given; it is elementary to check that the inequality $(x-1)^2 \leq 2\varepsilon (x+1)$ is equivalent to

$$|x - (1+\varepsilon)| \leq \sqrt{4\varepsilon + \varepsilon^2}.$$

For us this implies immediately $|x_n - 1| \le 6n^{-\alpha}$, proving the claim (47).

Now we have

$$\theta_n - \bar{\theta}_n = \frac{A_n}{\langle A \rangle_n} - \frac{B_n}{\langle B \rangle_n}$$
$$= \frac{A_n - B_n}{\langle B \rangle_n} + A_n \frac{\langle B \rangle_n - \langle A \rangle_n}{\langle A \rangle_n \langle B \rangle_n}. \tag{48}$$

Using (46) we have that the second term in (48) is bounded above in absolute value by

$$\sqrt{2} n^{-\alpha} \frac{A_n}{\langle A \rangle_n} \frac{|\langle A \rangle_n + \langle B \rangle_n|^{1/2}}{|\langle B \rangle_n|^{1/2}}$$
$$= \sqrt{2} n^{-\alpha} \frac{A_n}{\langle A \rangle_n} \left( \frac{\langle A \rangle_n}{\langle B \rangle_n} + 1 \right)^{1/2}.$$

By Theorems 2 and 3, $A_n / \langle A \rangle_n$ converges to the finite constant $\theta$. By the limit (47), the last term in the above expression converges to 2, so that the entire expression converges to 0.

Let $k$ and $\gamma$ be fixed positive values. For the first term in (48), using our hypotheses, by Chebyshev's theorem and the Burkholder-Davis-Gundy inequality, we have

$$\mathbf{P}\left[ |A_n - B_n|^k > n^{-k\gamma} \mathbf{E}\left[ \langle B \rangle_n^k \right] \right] \le n^{\gamma k} \mathbf{E}^{-1}\left[ \langle B \rangle_n^k \right] \mathbf{E}\left[ |A_n - B_n|^k \right]$$
$$\le n^{\gamma k} n^{-k\alpha}.$$

Thus picking a positive value $\gamma < \alpha$ and choosing $k$ large enough, by the Borell-Cantelli lemma, almost surely, for $n$ large enough

$$|A_n - B_n| \le n^{-\gamma} \mathbf{E}\left[ \langle B \rangle_n^k \right]^{1/k} \le \frac{1}{K} n^{-\gamma} \langle B \rangle_n ,$$

which finishes the proof of the proposition. ■

In order to apply this proposition to a concrete situation, we specialize to the linear case. Generalize the next result to non-linear cases would presumably require a hypothesis such as (28); we will not comment on this further.

**Theorem 4** *Assume $b(x) = x$. Then for all $H \in (0,1)$, almost surely $\lim_{n\to\infty} \bar{\theta}_n = \theta$ where the discretization $\bar{\theta}_n$ of the maximum likelihood estimator $\theta_n$ is defined in (45).*

The proof of this theorem proceeds through several lemmas. We begin with some definitions. For the sake of notational convenience, we reindex the sequence $Q$ so that $Q_{m+1}$ replaces $Q_m$. Let

$$M(n) := \mathbf{E}[|Q_n|],$$
$$R(n) := \frac{|Q_n| - M(n)}{\sqrt{n}},$$
$$S(n) := \sum_{n=1}^{m} Q_m,$$
$$Z(m) := \sup_{s \in [m, m+1]} |Q_s - Q_m|.$$

**Lemma 7** *$R$ is a centered sub-Gaussian stochastic process on $\mathbf{N}$ relative to the canonical metric corresponding to the centered Gaussian stochastic process $\{Q_m/\sqrt{m} : m \in \mathbf{N}\}$. In other words, let*

$$\delta^2(n, n') = \mathbf{E}\left[ \left( \frac{Q_n}{\sqrt{n}} - \frac{Q_{n'}}{\sqrt{n'}} \right)^2 \right];$$

*then for every* $\lambda \in \mathbf{R}$,

$$\mathbf{E}\left[\exp\left(\lambda\left(R\left(n\right)-R\left(n'\right)\right)\right)\right] \leq \exp\left(\frac{\lambda^2}{2}\delta^2\left(n,n'\right)\right).$$

*Moreover,* $M\left(n\right)=\sqrt{n}c_H$ *for some constant* $c_H$*.*

**Proof**. The last statement of the lemma is trivial using the scaling property of fBm and the equality (27). For the remainder, note that

$$\begin{aligned}\left(R\left(n\right)-R\left(n'\right)\right)^2 &= \left(\frac{\left|Q_n\right|}{\sqrt{n}}-\frac{\left|Q_{n'}\right|}{\sqrt{n'}}\right)^2\\ &= \frac{\left|Q_n\right|^2}{n}+\frac{\left|Q_{n'}\right|^2}{n'}-2\frac{\left|Q_n\right|\left|Q_{n'}\right|}{\sqrt{nn'}}\\ &\leq \left(\frac{Q_n}{\sqrt{n}}-\frac{Q_{n'}}{\sqrt{n'}}\right)^2.\end{aligned}$$

The announced result follows immediately. ■

**Lemma 8** *The variance of the centered Gaussian random variable* $S\left(n\right)$ *is bounded, for some constant* $c'_H$*, as follows:*

$$s_n := \mathbf{E}\left[\left|S\left(n\right)\right|^2\right] \leq n^2 c'_H.$$

**Proof**. The proof will proceed by induction on $n$. We have, using the constant $c_H$ in the previous lemma,

$$\begin{aligned}s_{n+1} &= \mathbf{E}\left[\left|S\left(n\right)+Q_{n+1}\right|^2\right]\\ &= s_n+\left(c_H\right)^2\left(n+1\right)+2\mathbf{E}\left[S\left(n\right)Q_{n+1}\right].\end{aligned}$$

We calculate the expectation in this last expression. Using the defining formula (26), and the covariance for fBm, we have

$$\mathbf{E}\left[S\left(n\right)Q_{n+1}\right] = \sum_{m=1}^{n} m^{1/2-H}\left(n+1\right)^{1/2-H}\int_0^1\int_0^1 drdr'\left|\left(mr\right)^{2H}+\left(\left(n+1\right)r\right)^{2H}-\left|mr+\left(n+1\right)r\right|^{2H}\right|.$$

Some tedious calculations yield, with $x=m/\left(n+1\right)$

$$\mathbf{E}\left[S\left(n\right)Q_{n+1}\right] = \sum_{m=1}^{n} x^{3/2-H}\left(n+1\right)^{2H}\left[\frac{1}{2H+1}\left(x^{2H+1}+\left(1-x\right)^{2H+1}-1\right)+x\right].$$

We see that this quantity is, up to a factor $\left(n+1\right)^{2H+1}$, the Riemann-sum approximation for the integral

$$I_H := \int_0^1 x^{3/2-H}\left[\frac{1}{2H+1}\left(x^{2H+1}+\left(1-x\right)^{2H+1}-1\right)+x\right].$$

A calculation shows that this value $I_H$ is strictly negative for any value of $H\in(0,1)$. Since for some $n_{0,H}$ depending only on $H$, for all $n\geq n_{0,H}$,

$$\left|\mathbf{E}\left[S\left(n\right)Q_{n+1}\right]-I_H\right| \leq I_H/2,$$

we immediately obtain that for all $n\geq n_{0,H}$,

$$\mathbf{E}\left[S\left(n\right)Q_{n+1}\right] < 0.$$

This allows us to write, for all $n \geq n_{0,H}$,

$$s_{n+1} \leq s_n + (c_H)^2 (n+1) .$$

Now for $n = n_{0,H}$, $s_n$ is a value depending only on $H$. Therefore we can write for some constant $c''_H$ depending only on $H$ that the induction hypothesis (conclusion of the lemma) holds at $n = n_{0,H}$ if $c'_H \geq c''_H$. We then obtain by induction that for all $n \geq n_{0,H}$,

$$\begin{aligned} s_{n+1} &\leq n^2 c'_H + (c_H)^2 (n+1) \\ &\leq n^2 c'_H + (c_H)^2 (n+1) + c'_H. \end{aligned}$$

This is consistent with the induction hypothesis as soon as $(c_H)^2 \leq 2c'_H$. Therefore, if we take $c'_H = \max\left(c''_H; (c_H)^2/2\right)$, the proof is complete. ■

**Lemma 9** *For any $H \in (0,1)$, for any $\beta \in (0,H)$, almost surely, there exists a finite integer $n_0$ such that for $n \geq n_0$,*

$$|Z_m| \leq m^{1/2-(H-\beta)}.$$

Proof. This follows easily from the result of Lemma 6 and the Borel-Cantelli lemma. The details are omitted. ■

**Proof of Theorem 4**. First note that we have

$$\begin{aligned} \langle A-B \rangle_n &= \int_0^n \left|Q_s - Q_{[s]}\right|^2 ds \\ &= \sum_{m=0}^{n-1} \int_m^{m+1} \left|Q_s - Q_{[s]}\right|^2 ds \\ &\leq \sum_{m=0}^{n-1} |Z_m|^2 = \sum_{m=0}^{n_0-1} |Z_m|^2 + \sum_{m=n_0}^{n-1} |Z_m|^2 . \end{aligned}$$

We recall that from the proofs of Theorems 2 and 3 that almost surely, $\lim_{n\to\infty} \langle B \rangle_n = +\infty$. Therefore to prove that $\langle A-B \rangle_n / \langle B \rangle_n$ converges to 0 as in the first condition of Proposition 3, we only need to consider the tail of the above series (terms for $m \geq n_0$). By the result of Lemma 9, almost surely,

$$\sum_{m=n_0}^{n-1} |Z_m|^2 \leq \sum_{m=0}^{n-1} m^{1-2(H-\beta)} \leq C_{H,\beta} n^{2-2(H-\beta)}.$$

Therefore it is sufficient to prove that that for some $\alpha > 0$, almost surely

$$\lim_{n\to\infty} \frac{n^{2-2(H-\beta)+\alpha}}{\langle B \rangle_n} = 0.$$

We can write, using Jensen's inequality (and using the reindexed $Q$), that

$$\sqrt{\langle B \rangle_n} = \sqrt{\sum_{m=1}^{n} |Q_m|^2} \geq \frac{1}{\sqrt{n}} \sum_{m=1}^{n} |Q_m| . \tag{49}$$

Therefore we only need to show, for some $\alpha > 0$, almost surely,

$$\lim_{n\to\infty} \frac{n^{3/2-(H-\beta)+\alpha}}{\sum_{m=1}^{n} |Q_m|} = 0. \tag{50}$$

Using Lemma 7, the denominator in the above expression can be rewritten as

$$\begin{aligned}\sum_{m=1}^{n}|Q_m| &= \sum_{m=1}^{n}\left(c_H\sqrt{m}+\sqrt{m}R(m)\right)\\ &\le c_H n^{3/2}+\sum_{m=1}^{n}\sqrt{m}R(m)\\ &:= q_1+q_2,\end{aligned} \tag{51}$$

which means that this expression is the sum of $q_1 := c_H n^{3/2}$ and of a random variable $q_2 = q_2(R,n)$ which is a linear functional of the centered sub-Gaussian process $R$ whose canonical metric is bounded above by that of the process $Q_\cdot/\sqrt{\cdot}$. By the sub-Gaussian theory (see [18, Chapter 12]), and given the fact that $q_2(R,1)$ has a density which is bounded above by that of a multiple of $Q_1$, we can conclude for example that for some constant $c>0$, for all $\lambda>0$,

$$\begin{aligned}\mathbf{E}\left[\exp\lambda q_2(R,n)\right] \le \mathbf{E}\left[\exp\lambda q_2\left(cQ_\cdot/\sqrt{\cdot},n\right)\right] &= \mathbf{E}\left[\exp\lambda c\sum_{m=1}^{n}Q_m\right]\\ &= \mathbf{E}\left[\exp\lambda cS(n)\right] = \exp\left(\frac{1}{2}\lambda^2c^2s_n\right)\\ &\le \exp\left(\frac{1}{2}\lambda^2c^2n^2c'_H\right),\end{aligned}$$

where the last estimate is by Lemma 8.

Now from inequality (51) and the previous exponential moment, for some constant $c_H$ we can write, for a fixed $\gamma>0$

$$\begin{aligned}\mathbf{P}\left[\sum_{m=1}^{n}|Q_m| \le c_H n^{3/2}-n^{\gamma}\right] &\le \mathbf{P}\left[|q_2(R,n)|\ge n^{\gamma}\right]\\ &\le \exp\left(\frac{1}{2}\lambda^2n^2c'_H\right)\exp\left(-\lambda n^{\gamma}\right).\end{aligned}$$

Choosing $\lambda = 2n^{\gamma-2}$ yields that the right-hand side in the last expression is equal to $\exp\left(-\frac{1}{2}n^{2\gamma-2}\right)$. If we take $\gamma>1$, then the above expression is summable in $n$, implying, by the Borel-Cantelli lemma, that for $n$ large enough,

$$\begin{aligned}\sum_{m=1}^{n}|Q_m| &\ge c_H n^{3/2}-n^{\gamma}\\ &\ge \frac{1}{2}c_H n^{3/2}\end{aligned} \tag{52}$$

The limit in (50) follows, and therefore also the convergence of $\langle A-B\rangle_n/\langle B\rangle_n$ to 0 almost surely at the speed $n^{-\alpha}$ for any $\alpha<H$ (by choosing $\beta$ arbitrarily small), as required in the first condition of Proposition 3.

The second and third conditions of Proposition 3 can be established using similar arguments to what we have presented in this section, including the sub-Gaussian property of Lemma 7. For example, the second condition for $k=1$ is obtained as follows. By combining (52) and (49) we get almost surely for large $n$,

$$\langle B\rangle_n \ge \frac{1}{n}\left(\sum_{m=1}^{n}|Q_m|\right)^2 \ge \left(\frac{c_H}{2}\right)^2 n^2.$$

On the other hand, we can calculate explicitly

$$\mathbf{E}\left[\langle B\rangle_n\right] = \sum_{m=1}^{n} mc_H \leq c_H n^2.$$

We leave it to the reader to check the second and third conditions of the proposition do hold. The theorem then immediately follows from the proposition. ■

# References


[1] R. Adler (1990). *Introduction to continuity, extrema, and related topics for general Gaussian processes.* Ins. Math. Stat, Hayward, CA.

[2] E. Alos, O. Mazet and D. Nualart (2001). Stochastic calculus with respect to Gaussian processes. *Annals of Probability*, **29**, 766-801.

[3] J.P. Bishwal; A. Bose (2001). Rates of convergence of approximate maximum likelihood estimators in the Ornstein-Uhlenbeck process. *Comput. Math. Appl.* **42**, no. 1-2, 23–38.

[4] B. Boufoussi and Y. Ouknine (2003). On a stochastic equation driven by a fBm with discontinuous drift. *Electronic Comm. in Probab.*, **8**, 122-134.

[5] P. Cheridito, H. Kawaguchi, M. Maejima (2003). Fractional Ornstein-Uhlenbeck processes. *Electronic Journal of Probability*, **8**, paper no. 3, 1-14.

[6] E.M. Cleur (2001). Maximum likelihood estimates of a class of one-dimensional stochastic differential equation models from discrete data. *J. Time Ser. Anal.* **22**, no. 5, 505–515.

[7] L. Decreusefond and A.S. Ustunel (1999). Stochastic analysis of the fractional Brownian motion. *Potential Analysis*, **10**, 177-214.

[8] B. Djehiche and M. Eddahbi (2001). Hedging options in market models modulated by fractional Brownian motion. *Stochastic Analysis and Applications*, **19**(5), 753-770.

[9] A. Friedman (1975). *Stochastic differential equations and applications.* Academic Press.

[10] Y. Hu and B. Oksendhal (2003). Fractional white noise calculus and applications to finance. *IDAQP*, **6**(1), 1-32.

[11] Yu. A. Kutoyants (1977). On a property of estimator of parameter of trend coefficient. *Izv. Akad. Nauk Arm. SSR., Mathematika*, **12**, 245-251.

[12] Yu. A. Kutoyants (1977). Estimation of the trend parameter of a diffusion process. *Theory Probab. Appl.*, **22**, 399-405.

[13] Yu. A. Kutoyants (1980). *Parameter Estimation for Stochastic Processes*, Heldermann, Berlin, 1984. (Russian edition 1980).

[14] M.L Kleptsyna and A. Le Breton (2002). Statistical analysis of the fractional Ornstein-Uhlenbeck type processes. *Statistical inference for stochastic processes*, **5**, 229-248.

[15] M. Kleptsyna, A. Le Breton and M.C. Roubaud (2000). Parameter estimation and optimal filtering for fractional type stochastic systems. *Statistical inference for stochastic processes*, **3**, 173-182.

[16] A.Kukush, Y. Mishura and E. Valkeila (2005). Statistical inference with fractional Brownian motion. *Statistical inference for stochastic processes*, **8**, 71-93.

[17] A. Le Breton (1976). On continuous and discrete sampling for parameter estimation in diffusion type processes. *Mathematical Programming studies*, **5**, 124-144.

[18] M. Ledoux and M. Talagrand (1990). *Probability on Banach Spaces.* Springer-Verlag.

[19] R.S. Liptser and A.N. Shiryayev (1978). *Statistics of random processes II. Applications* Springer-Verlag.

[20] I. Norros, E. Valkeila and J. Virtamo (1999). An elementary approach to a Girsanov formula and other analytical results on fractional Brownian motion. *Bernoulli*, **5**(4), 571-587.

[21] D. Nualart (1995). *The Malliavin calculus and related topics.* Springer-Verlag.

[22] D. Nualart and Y. Ouknine (2002). Regularization of differential equations by fractional noise. *Stoc. Proc. Appl.*,**102**, 103-116.

[23] A.R. Pedersen, (1995). A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations. *Scand. J. Statist.* **22**, no. 1, 55–71.

[24] D. Revuz and M. Yor (1999). *Continuous Martingales and Brownian Motion. Third edition.* Springer-Verlag.

[25] B.L.S. Prakasa Rao (2003). Parameter estimation for linear stochastic differential equations driven by fractional Brownian motion. *Random Oper. Stoc. Eqs.*, **11**(3), 229-242.

[26] A.S. Üstünel (1985): *An Introduction to Analysis on Wiener Space.* Lecture Notes in Mathematics, **1610**. Springer-Verlag.