April 2, 2008

ASYMPTOTIC PROPERTIES OF THE MAXIMUM LIKELIHOOD ESTIMATOR FOR STOCHASTIC PARABOLIC EQUATIONS WITH ADDITIVE FRACTIONAL BROWNIAN MOTION

IGOR CIALENCO, SERGEY V. LOTOTSKY, AND JAN POSPISIL

Abstract. A parameter estimation problem is considered for a diagonaliazable stochastic evolution equation using a finite number of the Fourier coefficients of the solution. The equation is driven by additive noise that is white in space and fractional in time with the Hurst parameter $H \geq 1/2$. The objective is to study asymptotic properties of the maximum likelihood estimator as the number of the Fourier coefficients increases. A necessary and sufficient condition for consistency and asymptotic normality is presented in terms of the eigenvalues of the operators in the equation.

1. Introduction

In the classical statistical estimation problem, the starting point is a family $\mathbf{P}^\theta$ of probability measures depending on the parameter θ in some subset Θ of a finite-dimensional Euclidean space. Each $\mathbf{P}^\theta$ is the distribution of a random element. It is assumed that a realization of one random element corresponding to one value $\theta = \theta_0$ of the parameter is observed, and the objective is to estimate the values of this parameter from the observations.

The intuition is to select the value θ corresponding to the random element that is *most likely* to produce the observations. A rigorous mathematical implementation of this idea leads to the notion of the regular statistical model [4]: the statistical model (or estimation problem) $\mathbf{P}^\theta,\ \theta \in \Theta$, is called regular, if the following two conditions are satisfied:

- there exists a probability measure $\mathbf{Q}$ such that all measures $\mathbf{P}^\theta$ are absolutely continuous with respect to $\mathbf{Q}$;
- the density $d\mathbf{P}^\theta/d\mathbf{Q}$, called the likelihood ratio, has a special property, called local asymptotic normality.

If at least one of the above conditions is violated, the problem is called singular.

2000 *Mathematics Subject Classification.* Primary 60H15; Secondary 62F12.

Key words and phrases. Asymptotic normality, ergodicity, parameter estimation, stochastic evolution equations.

SVL acknowledges support from the NSF CAREER award DMS-0237724, as well as hospitality and support of the Scuola Normale Superiore (Pisa, Italy). SVL and JP are grateful to the Institut Mittag-Leffler (Djursholm, Sweden) for the hospitality and support. The work of JP was also partially supported by MSMT Research Plan MSM 4977751301.

In regular models, the estimator $\widehat{\theta}$ of the unknown parameter is constructed by maximizing the likelihood ratio and is called the maximum likelihood estimator (MLE). Since, as a rule, $\widehat{\theta} \neq \theta_0$, the consistency of the estimator is studied, that is, the convergence of $\widehat{\theta}$ to θ_0 as more and more information becomes available. In all known regular statistical problems, the amount of information can be increased in one of two ways: (a) increasing the sample size, for example, the observation time interval (large sample asymptotic); (b) reducing the amplitude of noise (small noise asymptotic).

In finite-dimensional models, the only way to increase the sample size is to increase the observation time. In infinite-dimensional models, in particular, those provided by stochastic partial differential equations (SPDEs), another possibility is to increase the dimension of the spatial projection of the observations. Thus, a consistent estimator can be possible on a finite time interval with fixed noise intensity. This possibility was first suggested by Huebner at al. [2] for parabolic equations driven by additive space-time white noise, and was further investigated by Huebner and Rozovskii [3], where a necessary and sufficient condition for the existence of a consistent estimator was stated in terms of the orders of the operators in the equation.

The objective of the current paper is to extend the model from [3] to parabolic equations in which the time component of the noise is fractional with the Hurst parameter $H \geq 1/2$. More specifically, we consider an abstract evolution equation

$$u(t) + \int_0^t (\mathcal{A}_0 + \theta\mathcal{A}_1)u(s)ds = W^H(t), \tag{1.1}$$

where $\mathcal{A}_0$, $\mathcal{A}_1$ are known linear operators and $\theta \in \Theta \subseteq \mathbb{R}$ is the unknown parameter; the zero initial condition is taken to simplify the presentation. The noise $W^H(t)$ is a formal series

$$W^H(t) = \sum_{j=1}^{\infty} w_j^H(t)h_j, \tag{1.2}$$

where $\{w_j^H,\ j \geq 1\}$ are independent fractional Brownian motions with the same Hurst parameter $H \geq 1/2$ and $\{h_j,\ j \geq 1\}$ is an orthonormal basis in a Hilbert space $\mathbf{H}$; $H = 1/2$ corresponds to the usual space-time white noise. Existence and uniqueness of the solution for such equations are well-known for all $H \in (0, 1)$ (see, for example, Tindel et al. [14, Theorem 1]).

The main additional assumption about (1.1), both in [3] and in the current paper, is that the equation is *diagonalizable*: $\{h_j,\ j \geq 1\}$ from (1.2) is a common system of eigenfunction of the operators $\mathcal{A}_0$ and $\mathcal{A}_1$:

$$\mathcal{A}_0 h_j = \rho_j h_j,\ \mathcal{A}_1 = \nu_j h_j. \tag{1.3}$$

Under certain conditions on the numbers $\rho_j,\ \nu_j$, the solution of (1.1) is a convergent Fourier series $u(t) = \sum_{j\geq 1} u_j(t)h_j$, and each $u_j(t)$ is a fractional Ornstein-Uhlenbeck (OU) process. An N-dimensional projection of the solution is then an N-dimensional fractional OU process with independent components. A Girsanov-type formula (for example, from Kleptsyna et al. [7, Theorem 3]) leads to a maximum likelihood estimator $\hat{\theta}_N$ of θ based on the first N Fourier coefficients $u_1, \ldots, u_N$ of the solution of (1.1). An explicit expression for this estimator exists but requires a number of additional notations; see formula (3.8) on page 8 below.

The following is the main results of the paper.

Theorem 1.1. *Define $\mu_j = \theta\nu_j + \rho_j$ and assume that the series $\sum_j (1+|\mu_j|)^{-\gamma}$ converges for some $\gamma > 0$. Then the maximum likelihood estimator $\hat{\theta}_N$ of θ is strongly consistent and asymptotically normal, as $N \to \infty$, if and only if the series $\sum_j \nu_j^2 \mu_j^{-1}$ diverges; the rate of convergence of the estimator is given by the square root of the partial sums of this series: as $N \to \infty$, the sequence $\left(\sum_{j\le N} \nu_j^2\mu_j^{-1}\right)^{1/2} (\hat{\theta}_N - \theta)$ converges in distribution to a standard Gaussian random variable.*

If the operators $\mathcal{A}_0$ and $\mathcal{A}_1$ are elliptic of orders m_0 and m_1 on $L_2(M)$, where M is a d-dimensional manifold, and $2m = \max(m_0, m_1)$, then the condition of the theorem becomes $m_1 \ge m - (d/2)$; in the case $H = 1/2$ this is known from [3]. Thus, beside extending the results of [3] to fractional-in-time noise, we also generalize the necessary and sufficient condition for consistency of the estimator.

While parameter estimation for the finite-dimensional fractional OU and similar processes has been recently investigated by Tudor and Viens [15] for all $H \in (0,1)$, our analysis in infinite dimensions requires more delicate results: an explicit expression for the Laplace transform of a certain functional of the fractional OU process, as obtained by Kleptsyna and Le Brenton [6], and for now this expression exists only for $H \ge 1/2$.

2. Stochastic Parabolic Equations with Additive FBM

In this section we introduce a diagonalizable stochastic parabolic equation depending on a parameter and study the main properties of the solution.

Let $\mathbf{H}$ be a separable Hilbert space with the inner product $(\cdot,\cdot)_0$ and the corresponding norm $\|\cdot\|_0$. Let Λ be a densely-defined linear operator on $\mathbf{H}$ with the following property: there exists a positive number c such that $\|\Lambda u\|_0 \ge c\|u\|_0$ for every u from the domain of Λ. Then the operator powers Λ^γ, $\gamma \in \mathbb{R}$, are well defined and generate the spaces $\mathbf{H}^\gamma$: for $\gamma > 0$, $\mathbf{H}^\gamma$ is the domain of Λ^γ; $\mathbf{H}^0 = \mathbf{H}$; for $\gamma < 0$, $\mathbf{H}^\gamma$ is the completion of $\mathbf{H}$ with respect to the norm $\|\cdot\|_\gamma := \|\Lambda^\gamma \cdot\|_0$ (see for instance Krein at al. [8]). By construction, the collection of spaces $\{\mathbf{H}^\gamma,\ \gamma \in \mathbb{R}\}$ has the following properties:

- $\Lambda^\gamma(\mathbf{H}^r) = \mathbf{H}^{r-\gamma}$ for every $\gamma, r \in \mathbb{R}$;
- For $\gamma_1 < \gamma_2$ the space $\mathbf{H}^{\gamma_2}$ is densely and continuously embedded into $\mathbf{H}^{\gamma_1}$: $\mathbf{H}^{\gamma_2} \subset \mathbf{H}^{\gamma_1}$ and there exists a positive number c_{12} such that $\|u\|_{\gamma_1} \le c_{12}\|u\|_{\gamma_2}$ for all $u \in \mathbf{H}^{\gamma_2}$;
- For every $\gamma \in \mathbb{R}$ and $m > 0$, the space $\mathbf{H}^{\gamma-m}$ is the dual of $\mathbf{H}^{\gamma+m}$ relative to the inner product in $\mathbf{H}^\gamma$, with duality $\langle\cdot,\cdot\rangle_{\gamma,m}$ given by

$$\langle u_1, u_2\rangle_{\gamma,m} = (\Lambda^{\gamma-m}u_1, \Lambda^{\gamma+m}u_2)_0,\ \text{where } u_1 \in \mathbf{H}^{\gamma-m},\ u_2 \in \mathbf{H}^{\gamma+m}.$$

In the above construction, the operator Λ can be bounded, and then the norms in all the spaces $\mathbf{H}^\gamma$ will be equivalent. A more interesting situation is therefore when Λ is unbounded and plays the role of the first-order operator.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{w_j^H,\ j \geq 1\}$ be a collection of independent fractional Brownian motions on this space with the same Hurst parameter $H \in (0,1)$:

$$\mathbb{E}w_j^H(t) = 0,\quad \mathbb{E}\big(w_j^H(t)w_j^H(s)\big) = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right).$$

Consider the following equation:

$$\begin{cases} du(t) + (\mathcal{A}_0 + \theta\mathcal{A}_1)u(t)dt = \sum\limits_{j\geq 1} g_j(t)dw_j^H(t),\ 0 < t \leq T, \\ u(0) = u_0 \end{cases} \tag{2.1}$$

where $\mathcal{A}_0$, $\mathcal{A}_1$ are linear operators, g_j are non-random, and θ is a scalar parameter belonging to an open set $\Theta \subset \mathbb{R}$.

Definition 2.1.

(1) *Equation* (2.1) *is called* `diagonalizable` *if the operators* $\mathcal{A}_0$, $\mathcal{A}_1$, *have a common system of eigenfunctions* $\{h_j,\ j \geq 1\}$ *such that* $\{h_j,\ j \geq 1\}$ *is an orthonormal basis in* $\mathbf{H}$ *and each* h_j *belongs to* $\bigcap_{\gamma\in\mathbb{R}} \mathbf{H}^\gamma$.

(2) *Equation* (2.1) *is called* (m,γ)-`parabolic` *for some numbers* $m \geq 0$ *and* $\gamma \in \mathbb{R}$ *if*
   - *the operator* $\mathcal{A}_0 + \theta\mathcal{A}_1$ *is uniformly bounded from* $\mathbf{H}^{\gamma+m}$ *to* $\mathbf{H}^{\gamma-m}$ *for* $\theta \in \Theta$: *there exists a positive real number* C_1 *such that*
   $$\|(\mathcal{A}_0 + \theta\mathcal{A}_1)v\|_{\gamma-m} \leq C_1\|v\|_{\gamma+m} \tag{2.2}$$
   for all $\theta \in \Theta$, $v \in \mathbf{H}^{\gamma+m}$;
   - *there exists a positive number* δ *and a real number* C *such that, for every* $v \in \mathbf{H}^{\gamma+m}$, $\theta \in \Theta$,
   $$-2\langle(\mathcal{A}_0 + \theta\mathcal{A}_1)v, v\rangle_{\gamma,m} + \delta\|v\|^2_{\gamma+m} \leq C\|v\|^2_\gamma. \tag{2.3}$$

Remark 2.2. If equation (2.1) is (m,γ)-parabolic, then condition (2.3) implies that

$$\langle(2\mathcal{A}_0 + 2\theta\mathcal{A}_1 + CI)v, v\rangle_{\gamma,m} \geq \delta\|v\|^2_{\gamma+m},$$

where I is the identity operator. The Cauchy-Schwartz inequality and the continuous embedding of $\mathbf{H}^{\gamma+m}$ into $\mathbf{H}^\gamma$ then imply

$$\|(2\mathcal{A}_0 + 2\theta\mathcal{A}_1 + CI)v\|_\gamma \geq \delta_1\|v\|_\gamma$$

for some $\delta_1 > 0$ uniformly in $\theta \in \Theta$. As a result, we can take $\Lambda = (2\mathcal{A}_0 + 2\theta^*\mathcal{A}_1 + CI)^{1/(2m)}$ for some fixed $\theta^* \in \Theta$. If the operator $\mathcal{A}_0 + \theta\mathcal{A}_1$ is unbounded, it is natural to say that $\mathcal{A}_0 + \theta\mathcal{A}_1$ has order $2m$ and Λ has order 1.

From now on, if equation (2.1) *is* (m,γ)*-parabolic and diagonalizable, we will assume that the operator* Λ *has the same eigenfunctions as the operators* $\mathcal{A}_0$, $\mathcal{A}_1$*; by Remark 2.2, this leads to no loss of generality.*

For a diagonalizable equation, condition (2.3) can be expressed in terms of the eigenvalues of the operators in the equation.

Theorem 2.3. *Assume that equation* (2.1) *is diagonalizable and*

$$\mathcal{A}_0 h_j = \rho_j h_j, \ \mathcal{A}_1 h_j = \nu_j h_j.$$

With no loss of generality (see Remark 2.2)*, we also assume that*

$$\Lambda h_j = \lambda_j h_j.$$

Then equation (2.1) *is* (m,γ)*-parabolic if and only if there exist positive real numbers* δ, C_1 *and a real number* C_2 *such that, for all* $j \geq 1$ *and* $\theta \in \Theta$*,*

$$\lambda_j^{-2m}|\rho_j + \theta\nu_j| \leq C_1; \tag{2.4}$$

$$-2(\rho_j + \theta\nu_j) + \delta\lambda_j^{2m} \leq C_2. \tag{2.5}$$

Proof. We show that, for a diagonalizable equation, (2.4) is equivalent to (2.2) and (2.5) is equivalent to (2.3). Indeed, note that for every $\gamma, \ r \in \mathbb{R}$,

$$\|h_j\|_{\gamma+r} = \|\Lambda^r h_j\|_\gamma = \lambda_j^r \|h_j\|_\gamma.$$

Then (2.4) is (2.2) and (2.5) is (2.3), with $v = h_j$. Since both (2.4) and (2.5) are uniform in j and the collection $\{h_j, \ j \geq 1\}$ is dense in every $\mathbf{H}^\gamma$, the proof of the theorem is complete. □

Remark 2.4. (a) As conditions (2.4), (2.5) do not involve γ, we conclude that a diagonalizable equation is (m,γ)-parabolic *for some* γ if and only if it is (m,γ)-parabolic *for every* γ. As a result, in the future we will simply say that the equation is m-parabolic.

(b) If the operators $\mathcal{A}_0 + \theta\mathcal{A}_1$ and Λ are unbounded, then (2.5) implies that $\mu_j(\theta) = \rho_j + \theta\nu_j$ is positive for all sufficiently large j.

From now on we will assume that equation (2.1) is diagonalizable and fix the basis $\{h_j, \ j \geq 1\}$ in $\mathbf{H}$. Since each h_j belongs to every $\mathbf{H}^\gamma$ and, by construction, $\bigcap_\gamma \mathbf{H}^\gamma$ is dense in $\bigcup_\gamma \mathbf{H}^\gamma$, every element f of $\bigcup_\gamma \mathbf{H}^\gamma$ has a unique expansion $\sum_{j\geq 1} f_j h_j$, where $f_j = \langle f, h_j\rangle_{0,m}$ for a suitable m.

Definition 2.5. *The* `space-time fractional Brownian motion` W^H *is an element of* $\bigcup_{\gamma\in\mathbb{R}} \mathbf{H}^\gamma$ *with the expansion*

$$W^H(t) = \sum_{j\geq 1} w_j^H(t) h_j. \tag{2.6}$$

Definition 2.6. *Let* W^H *be a space-time fractional Brownian motion. The solution of the diagonalizable equation*

$$\begin{cases} du(t) + (\mathcal{A}_0 + \theta\mathcal{A}_1)u(t)dt = dW^H(t), \ 0 < t \leq T, \\ u(0) = u_0 \end{cases} \tag{2.7}$$

$u_0 \in \mathbf{H}$*, is a random process with values in* $\bigcup_\gamma \mathbf{H}^\gamma$ *and an expansion*

$$u(t) = \sum_{j\geq 1} u_j(t) h_j, \tag{2.8}$$

where

$$u_j(t) = (u_0, h_j)_0\, e^{-(\theta\nu_j+\rho_j)t} + \int_0^t e^{-(\theta\nu_j+\rho_j)(t-s)} dw_j^H(s). \tag{2.9}$$

Notice that, due to the special structure of the equation, Definition 2.6 implies both existence and uniqueness of the solution.

To simplify further notations we write

$$\mu_j(\theta) = \theta\nu_j + \rho_j. \tag{2.10}$$

By (2.5), if equation (2.1) is m-parabolic and diagonalizable, then, for every $\theta \in \Theta$, there exists a positive integer J such that

$$\mu_j(\theta) > 0 \quad \text{for all} \quad j \geq J.$$

Theorem 2.7. *Assume that*

(1) $H \geq 1/2$;
(2) *equation* (2.1) *is m-parabolic and diagonalizable;*
(3) *There exists a positive real number γ such that*

$$\sum_{j\geq 1} (1 + |\mu_j(\theta)|)^{-\gamma} < \infty. \tag{2.11}$$

Then, for every $t > 0$,

(1) $W^H(t) \in L_2(\Omega; \mathbf{H}^{-m\gamma})$;
(2) $u(t) \in L_2(\Omega; \mathbf{H}^{-m\gamma+2mH})$.

Proof. Condition (2.11) implies that $\lim_{j\to\infty} |\mu_j| = \infty$, and consequently the operators $\mathcal{A}_0 + \theta\,\mathcal{A}_1$ and Λ are unbounded. The parabolicity assumption and Theorem 2.3 then imply that, for all sufficiently large j,

$$1 + |\mu_j(\theta)| \leq C_2 \lambda_j^{2m},$$

uniformly in $\theta \in \Theta$.

$$\mathbb{E}\|W^H(t)\|_{-m\gamma}^2 = t^{2H} \sum_{j\geq 1} \lambda_j^{-2m\gamma} \leq C_2 t^{2H} \sum_{j\geq 1} (1 + |\mu_j(\theta)|)^{-\gamma} < \infty.$$

Next, the properties of the fractional Brownian motion imply

$$\mathbb{E}u_j^2(t) = H(2H-1)e^{-2\mu_j(\theta)t} \int_0^t \int_0^t e^{\mu_j(\theta)(s_1+s_2)} |s_1 - s_2|^{2H-2} ds_1 ds_2;$$

see, for example, Pipiras and Taqqu [11, formulas (4.1), (4.2)]. By direct computation,

$$\lim_{j\to\infty} |\mu_j(\theta)|^{2H} \mathbb{E}u_j^2(t) = H(2H-1) \int_0^\infty x^{2H-2} e^{-x} dx = H(2H-1)\Gamma(2H-1). \tag{2.12}$$

Consequently,

$$\sum_{j=1}^\infty (1 + |\mu_j(\theta)|)^{-\gamma+2H} \mathbb{E}|u_j(t)|^2 < \infty, \tag{2.13}$$

and the second conclusion of the theorem follows. □

Example 2.8. (a) For $0 < t \leq T$ and $x \in (0,1)$, consider the equation

$$du(t,x) - \theta\, u_{xx}(t,x)dt = dW^H(t,x) \tag{2.14}$$

with periodic boundary conditions, where $u_{xx} = \partial^2 u/\partial x^2$. Then $\mathbf{H}^\gamma$ is the Sobolev space on the unit circle (see, for example, Shubin [13, Section I.7]) and $\Lambda = \sqrt{I - \mathbf{\Delta}}$, where $\mathbf{\Delta}$ is the Laplace operator on $(0,1)$ with periodic boundary conditions. Direct computations show that equation (2.14) is diagonalizable; it is 1-parabolic if and only if $\theta > 0$. Also, $\mu_j = -\theta\pi^2 j^2$, so that, by Theorem 2.7 the solution $u(t)$ of (2.14) is an element of $L_2(\Omega; \mathbf{H}^{-\gamma+2H})$ for every $t > 0$, $\gamma > 1/2$, and $\theta > 0$.

(b) Let G be a smooth bounded domain in $\mathbb{R}^d$. Let $\mathbf{\Delta}$ be the Laplace operator on G with zero boundary conditions. It is known (for example, from Shubin [13]), that

(1) the eigenfunctions $\{h_j,\ j \geq 1\}$ of $\mathbf{\Delta}$ are smooth in G and form an orthonormal basis in $L_2(G)$;
(2) the corresponding eigenvalues $\sigma_j,\ j \geq 1$, can be arranged so that $0 < -\sigma_1 \leq -\sigma_2 \leq \ldots$, and there exists a number $c > 0$ such that $|\sigma_j| \sim c j^{2/d}$, that is,

$$\lim_{j\to\infty} |\sigma_j| j^{-2/d} = c. \tag{2.15}$$

We take $\mathbf{H} = L_2(G)$, $\Lambda = \sqrt{I - \mathbf{\Delta}}$, where I is the identity operator. Then $\|\Lambda u\|_0 \geq \sqrt{1 - \sigma_1}\|u\|_0$ and the operator Λ generates the Hilbert spaces $\mathbf{H}^\gamma$, and, for every $\gamma \in \mathbb{R}$, the space $\mathbf{H}^\gamma$ is the closure of the set of smooth compactly supported function on G with respect to the norm

$$\left(\sum_{j\geq 1}(1+j^2)^\gamma |\varphi_j|^2\right)^{1/2}, \quad \text{where } \varphi_j = \int_G \varphi(x)h_j(x)dx,$$

which is an equivalent norm in $\mathbf{H}^\gamma$. Then, for every $\theta \in \mathbb{R}$, the stochastic equation

$$du - \big(\mathbf{\Delta} u + \theta u\big)dt = dW^H(t,x) \tag{2.16}$$

is diagonalizable and 1-parabolic. Indeed, we have $\mathcal{A}_1 = I$, $\mathcal{A}_0 = -\mathbf{\Delta}$, and

$$-2\langle \mathcal{A}_0 v, v\rangle_{\gamma,1} = -2\|v\|^2_{\gamma+1} + 2\|u\|^2_\gamma,$$

so that (2.3) holds with $\delta = 2$ and $C = 2 - \theta$. Finally, by (2.15) we see that (2.11) holds for every $\gamma > d/2$. As a result, by Theorem 2.7, the solution $u(t)$ of (2.16) is an element of $L_2(\Omega; \mathbf{H}^{-\gamma+2H})$ for every $t > 0$, $\gamma > d/2$, and $\theta \in \mathbb{R}$.

3. The Maximum Likelihood Estimator and its Properties

Consider the diagonalizable equation

$$du(t) + (\mathcal{A}_0 + \theta\mathcal{A}_1)u(t)dt = dW^H(t) \tag{3.1}$$

with solution $u(t) = \sum_{j\geq 1} u_j(t)h_j$ given by (2.9); for simplicity, we assume that $u(0) = 0$. Suppose that the processes $u_1(t), \ldots, u_N(t)$ can be observed for all $t \in [0,T]$. The problem is to estimate the parameter θ using these observations.

Recall the notation $\mu_j(\theta) = \rho_j + \nu_j\theta$, where ρ_j and ν_j are the eigenvalues of $\mathcal{A}_0$ and $\mathcal{A}_1$, respectively. Then each u_j is a fractional Ornstein-Uhlenbeck process satisfying

$$du_j(t) = -\mu_j(\theta)u_j(t)dt + dw_j^H(t),\ u_j(0) = 0, \tag{3.2}$$

and, because of the independence of w_j^H for different j, the processes $u_1, \dots, u_N$ are (statistically) independent.

Let Γ denote the Gamma-function (see (2.12)). Following Kleptsyna and Le Brenton [6], we introduce the notations

$$\kappa_H = 2H\Gamma\left(\frac{3}{2} - H\right)\Gamma\left(H + \frac{1}{2}\right),\ k_H(t,s) = \kappa_H^{-1}s^{\frac{1}{2}-H}(t-s)^{\frac{1}{2}-H}; \tag{3.3}$$

$$\lambda_H = \frac{2H\Gamma(3-2H)\Gamma\left(H+\frac{1}{2}\right)}{\Gamma\left(\frac{3}{2}-H\right)},\ \mathbf{w}_H(t) = \lambda_H^{-1}t^{2-2H}; \tag{3.4}$$

$$M_j^H(t) = \int_0^t k_H(t,s)dw_j^H(s),\ Q_j(t) = \frac{d}{d\mathbf{w}_H(t)}\int_0^t k_H(t,s)u_j(s)ds; \tag{3.5}$$

$$Z_j(t) = \int_0^t k_H(t,s)du_j(s). \tag{3.6}$$

By a Girsanov-type formula (see, for example, Kleptsyna et al. [7, Theorem 3]), the measure in the space of continuous, $\mathbb{R}^N$-valued functions, generated by the process $(u_1, \dots, u_N)$ is absolutely continuous with respect to the measure generated by the process $(w_1^H, \dots, w_N^H)$, and the density is

$$\exp\left(-\sum_{j=1}^N \mu_j(\theta)\int_0^T Q_j(s)dZ_j(s) - \sum_{j=1}^N \frac{|\mu_j(\theta)|^2}{2}\int_0^T Q_j^2(s)d\mathbf{w}_H(s)\right). \tag{3.7}$$

Maximizing this density with respect to θ gives the Maximum Likelihood Estimator (MLE):

$$\widehat{\theta}_N = -\frac{\sum_{j=1}^N \int_0^T \nu_j Q_j(s)\big(dZ_j(s) + \rho_j Q_j(s)d\mathbf{w}_H(s)\big)}{\sum_{j=1}^N \int_0^T \nu_j^2 Q_j^2(s)d\mathbf{w}_H(s)}. \tag{3.8}$$

An important feature of (3.8) is that the process Z_j is a semi-martingale ([6, Lemma 2.1]), and so there is no stochastic integration with respect to fractional Brownian motion: $\int_0^T \nu_j Q_j(s)dZ_j(s)$ is an Itô integral. Notice that, when $H = 1/2$, we have $k_H = 1$, $\mathbf{w}_H(s) = s$, $Q_j(s) = Z_j(s) = u_j(s)$, and (3.8) becomes

$$\widehat{\theta}_N = -\frac{\sum_{j=1}^N \int_0^T \nu_j u_j(s)\big(du_j(s) + \rho_j u_j(s)ds\big)}{\sum_{j=1}^N \int_0^T \nu_j^2 u_j^2(s)du_j(s)}, \tag{3.9}$$

which is the MLE from [3].

Let us also emphasize that an implementation of (3.8) is impossible without the knowledge of H.

The following is the main result of the paper.

Theorem 3.1. *Under the assumptions of Theorem 2.7, the following conditions are equivalent:*

$$(1) \quad \sum_{j=J}^{\infty} \frac{\nu_j^2}{\mu_j(\theta)} = +\infty; \tag{3.10}$$

$$(2) \quad \lim_{N\to\infty} \widehat{\theta}_N = \theta \text{ with probability one,} \tag{3.11}$$

where $J = \min\{j : \mu_i(\theta) > 0 \text{ for all } i \geq j\}$.

Proof. Following Kleptsyna and Le Brenton [6, Equation (4.1)], we conclude that

$$\widehat{\theta}_N - \theta = -\frac{\sum_{j=1}^{N} \int_0^T \nu_j Q_j(s) dM_j^H(s)}{\sum_{j=1}^{N} \int_0^T \nu_j^2 Q_j^2(s) d\mathbf{w}_H(s)}. \tag{3.12}$$

Both the top and the bottom on the right-hand side of (3.12) are sums of independent random variables; moreover, it is known from [6, page 242] that

$$\mathbb{E}\left(\int_0^T Q_j(s) dM_j^H(s)\right)^2 = \mathbb{E}\int_0^T Q_j^2(s) d\mathbf{w}_H(s) ds. \tag{3.13}$$

From the expression for the Laplace transform of $\int_0^T Q_j^2(s) d\mathbf{w}_H(s) ds$ (see [6, Equation (4.2)]) direct computations show that

$$\lim_{j\to\infty} \mu_j(\theta) \mathbb{E}\int_0^T Q_j^2(s) d\mathbf{w}_H(s) ds = \frac{T}{2} > 0 \tag{3.14}$$

and, with $\mathrm{Var}(\xi)$ denoting the variance of the random variable ξ,

$$\lim_{j\to\infty} \mu_j^3(\theta) \mathrm{Var}\left(\int_0^T Q_j^2(s) d\mathbf{w}_H(s) ds\right) = \frac{T}{2} > 0; \tag{3.15}$$

a detailed derivation of (3.14) and (3.15) is given in the appendix, Lemmas A.1 and A.2 respectively.

We now see that if (3.10) does not hold, then, by (3.14), the series

$$\sum_{j\geq 1} \int_0^T \nu_j^2 Q_j^2(s) d\mathbf{w}_H(s)$$

converges with probability one, which, by (3.12), means that (3.11) cannot hold.

On the other hand, if (3.10) holds, then

$$\sum_{n\geq J} \frac{\nu_n^2 \mu_n^{-1}}{\left(\sum_{j=1}^{n} \nu_j^2 \mu_j^{-1}\right)^2} < \infty. \tag{3.16}$$

Indeed, setting $a_n = \nu_n^2\mu_n^{-1}$ and $A_n = \sum_{j=1}^n a_j$, we notice that

$$\sum_{n\geq J} \frac{a_n}{A_n^2} \leq \sum_{n\geq J+1} \left(\frac{1}{A_n} - \frac{1}{A_{n-1}}\right) = \frac{1}{A_J}.$$

Then the strong law of large numbers, together with the observation

$$\mathbb{E}\int_0^T Q_j(s)dM_j^H(s) = 0,\ j\geq 1,$$

implies

$$\lim_{N\to\infty} \frac{\sum_{j=1}^N \int_0^T \nu_j Q_j(s)dM_j(s)}{\sum_{j=1}^N \mathbb{E}\int_0^T \nu_j^2 Q_j^2(s)d\mathbf{w}_H(s)} = 0 \quad \text{with probability one.}$$

Next, it follows from (3.16) and (2.11) that

$$\sum_{n\geq J} \frac{\nu_n^4\mu_n^{-3}}{\left(\sum\limits_{j=J}^n \nu_j^2\mu_j^{-1}\right)^2} < \infty. \tag{3.17}$$

Then another application of the strong law of large numbers implies that

$$\lim_{N\to\infty} \frac{\sum_{j=1}^N \int_0^T \nu_j^2 Q_j^2(s)d\mathbf{w}_H(s)}{\sum_{j=1}^N \mathbb{E}\sum_{j=1}^N \int_0^T \nu_j^2 Q_j^2(s)d\mathbf{w}_H(s)} = 1 \tag{3.18}$$

with probability one, and (3.11) follows. □

Corollary 3.2. *Under assumptions of Theorem 2.7, if* (3.10) *holds, then*

$$\lim_{N\to\infty} \sqrt{\sum_{j=J}^N \frac{\nu_j^2}{\mu_j(\theta)}}\left(\widehat{\theta}_N - \theta\right) = \zeta \tag{3.19}$$

in distribution, where ζ is a Gaussian random variable with mean zero.

Proof. This follows from (3.12), (3.18), and the central limit theorem for the sum of independent random variables. □

Let us now consider a more general equation

$$du = (\mathcal{A}_0 + \theta\mathcal{A}_1)udt + \mathcal{B}dW^H(t),$$

where $\mathcal{B}$ is a linear operator. If $\mathcal{B}^{-1}$ exists, the equation reduced to (3.1) by considering $v = \mathcal{B}^{-1}u$. If $\mathcal{B}^{-1}$ does not exist, we have two possibilities:

(1) $(u_0, h_i)_0 = 0$ for every i such that $\mathcal{B}h_i = 0$. In this case, $u_i(t) = 0$ for all $t > 0$, so that we can factor out the kernel of $\mathcal{B}$ and reduce the problem to invertible $\mathcal{B}$.

(2) $(u_0, h_i)_0 \neq 0$ for some i such that $\mathcal{B}h_i = 0$. In this case, $u_i(t) = u_i(0)e^{-\rho_i t - \nu_i\theta t}$ and θ is determined exactly from the observations of $u_i(t)$:

$$\theta = \frac{1}{\nu_i(t-s)} \ln\frac{u_i(s)}{u_i(t)} - \frac{\rho_i}{\nu_i},\ t\neq s.$$

Let $\mathcal{A}_0$, $\mathcal{A}_1$ be differential or pseudo-differential operators, either on a smooth bounded domain in $\mathbb{R}^d$ or on a smooth compact d-dimensional manifold, and let m_0, m_1, be the orders of $\mathcal{A}_0$, $\mathcal{A}_1$ respectively, so that $2m = \max(m_0, m_1)$. Then, under rather general conditions we have

$$\lim_{j\to\infty} |\nu_j| j^{m_1/d} = c_1, \quad \lim_{j\to\infty} \mu_j(\theta) j^{2m/d} = c(\theta) \tag{3.20}$$

for some positive numbers $c_1, c(\theta)$; see, for example, Il'in [5] or Safarov and Vassiliev [12]. In particular, this is the case for the operators in equations (2.14) and (2.16).

If (3.20) holds, then condition (3.10) becomes

$$m_1 \geq m - (d/2), \tag{3.21}$$

which, in the case $H = 1/2$, was established by Huebner and Rozovskii [3]. In particular, (3.21) holds for equation (2.14) (where $2m = m_1 = 2$), and for equation (2.16) if $d \geq 2$ (where $2m = 2$, $m_1 = 0$).

Note that, at least as long as $H \geq 1/2$, conditions (3.10) and (3.21) do not involve H.

The maximum likelihood estimator (3.8) has three features that are clearly attractive: consistency, asymptotic normality, and absence of stochastic integration with respect to fractional Brownian motion. On the other hand, actual implementation of (3.8) is problematic: when $H > 1/2$, computing the processes Q_j and Z_j is certainly nontrivial. Estimator (3.9) is defined for all $H \geq 1/2$ and contains only the processes u_j, but, when $H > 1/2$, is not an MLE and is even harder to implement because of the stochastic integral with respect to u_j.

With or without condition (3.10), a consistent estimator of θ is possible in the large time asymptotic: for every $j \geq 1$,

$$\lim_{T\to\infty} \frac{\int_0^T \nu_j Q_j(s)\big(dZ_j(s) + \rho_j Q_j(s) d\mathbf{w}_H(s)\big)}{\int_0^T \nu_j^2 Q_j^2(s) d\mathbf{w}_H(s)} = -\theta \tag{3.22}$$

with probability one ([6, Proposition 2.2]). For $H > 1/2$, implementation of this estimator is essentially equivalent to the implementation of (3.8).

An alternative to (3.22) was suggested by Maslowski and Pospíšil [10] using the ergodic properties of the OU process. Let us first illustrate the idea on a simple example.

If $a > 0$ and $w = w(t)$ is a standard one-dimensional Brownian motion, then the OU process $dX = -aX(t)dt + dw(t)$ is ergodic and its unique invariant distribution is normal with zero mean and variance $(2a)^{-1}$. In particular,

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T X^2(t)dt = \frac{1}{2a} \tag{3.23}$$

with probability one, and so

$$\tilde{a}(T) = \frac{T}{2\int_0^T X^2(t)dt} \tag{3.24}$$

is a consistent estimator of a in the long-time asymptotic. Note that the maximum likelihood estimator in this case is

$$\hat{a}(T) = -\frac{\int_0^T X(t)dX(t)}{\int_0^T X^2(s)ds} \tag{3.25}$$

and is strongly consistent for every $a \in \mathbb{R}$ [9, Theorem 17.4].

Similarly, if $a > 0$, then the fractional OU process

$$dX(t) = -aX(t)dt + dw^H(t),\ X(0) = 0 \tag{3.26}$$

is Gaussian, and, by (2.12) on page 6, converges in distribution, as $t \to \infty$, to the Gaussian random variable with zero mean and variance $c(H)a^{-2H}$, where

$$c(H) = H(2H-1)\Gamma(2H-1); \tag{3.27}$$

notice that, in the limit $H \searrow 1/2$, we recover the result for the usual OU process. Further investigation shows that, similar to (3.23),

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T X^2(s)ds = \frac{c(H)}{a^{2H}}$$

(see [10]). As a result, for every j such that $\theta\nu_j + \rho_j > 0$, we have

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T u_j^2(t)dt = \frac{c(H)}{(\theta\nu_j + \rho_j)^{2H}} \tag{3.28}$$

with probability one. Under an additional assumption that $\nu_j \neq 0$, we get an estimator of θ

$$\tilde{\theta}^{(j)}(T) = \frac{1}{\nu_j}\left(\frac{c(H)T}{\int_0^T u_j^2(t)dt}\right)^{\frac{1}{2H}} - \frac{\rho_j}{\nu_j}. \tag{3.29}$$

This estimator is strongly consistent in the long time asymptotic: $\lim_{T\to\infty} |\tilde{\theta}^{(j)}(T) - \theta| = 0$ with probability one ([10, Theorem 5.2]). While not a maximum likelihood estimator, (3.29) is easier to implement computationally than (3.8). If, in Theorem 2.7 on page 6, we have $\mathcal{A}_0 = 0$, $\nu_j > 0$, and $\gamma < 2H$, then a version of (3.30) exists using all the Fourier coefficients $u_j,\ j \geq 1$:

$$\tilde{\theta}(T) = \left(\frac{c(H)T\sum_{j=1}^{\infty} \nu_j^{-2H}}{\sum_{j=1}^{\infty}\int_0^T u_j^2(t)dt}\right)^{\frac{1}{2H}}; \tag{3.30}$$

see [10, Theorem 5.2].

An interesting open question related to both (3.8) and (3.29), (3.30) is how to combine estimation of θ with estimation of H.

References

[1] L. C. Andrews. *Special functions for engineers and applied mathematicians.* Macmillan Co., New York, 1985.

[2] M. Huebner, R. Khasminskii, and B. L. Rozovskii. Two examples of parameter estimation. In S. Cambanis, J. K. Ghosh, R. L. Karandikar, and P. K. Sen, editors, *Stochastic Processes: A volume in honor of G. Kallianpur*, pages 149–160. Springer, New York, 1992.

[3] M. Huebner and B. L. Rozovskiĭ. On asymptotic properties of maximum likelihood estimators for parabolic stochastic PDE's. *Probab. Theory Related Fields*, 103(2):143–163, 1995.
[4] I. A. Ibragimov and R. Z. Khas′minskiĭ. *Statistical estimation*, volume 16 of *Applications of Mathematics*. Springer-Verlag, New York, 1981.
[5] V. A. Il′in. *Spectral theory of differential operators*. Consultants Bureau, New York, 1995.
[6] M. L. Kleptsyna and A. Le Breton. Statistical analysis of the fractional Ornstein-Uhlenbeck type process. *Stat. Inference Stoch. Process.*, 5(3):229–248, 2002.
[7] M. L. Kleptsyna, A. Le Breton, and M.-C. Roubaud. Parameter estimation and optimal filtering for fractional type stochastic systems. *Stat. Inference Stoch. Process.*, 3(1-2):173–182, 2000. 19th "Rencontres Franco-Belges de Statisticiens" (Marseille, 1998).
[8] S. G. Kreĭn, Yu. Ī. Petunīn, and E. M. Semënov. *Interpolation of linear operators*, volume 54 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, R.I., 1982.
[9] R. S. Liptser and A. N. Shiryayev. *Statistics of random processes, II: Applications*. Springer-Verlag, New York, 2nd edition, 2001.
[10] B. Maslowski and J. Pospíšil. Ergodicity and parameter estimates for infinite-dimensional fractional Ornstein-Uhlenbeck process. *Appl. Math. Optim.*, 0(0):1–29, 2008. Published online 11/2007, to appear in print.
[11] V. Pipiras and M. S. Taqqu. Integration questions related to fractional Brownian motion. *Probab. Theory Related Fields*, 118(2):251–291, 2000.
[12] Yu. Safarov and D. Vassiliev. *The asymptotic distribution of eigenvalues of partial differential operators*, volume 155 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1997. Translated from the Russian manuscript by the authors.
[13] M. A. Shubin. *Pseudodifferential operators and spectral theory*. Springer-Verlag, Berlin, second edition, 2001.
[14] S. Tindel, C. A. Tudor, and F. Viens. Stochastic evolution equations with fractional Brownian motion. *Probab. Theory Related Fields*, 127(2):186–204, 2003.
[15] C. A. Tudor and F. Viens. Statistical aspects of the fractional stochastic calculus. *Ann. Statist.*, 35(3):1183–1212, 2007.

Appendix

Below, we prove equalities (3.14) and (3.15).

Lemma A.1. *For every* $\theta \in \Theta$ *and* $H \in [1/2, 1)$,

$$\lim_{j\to\infty} \mu_j(\theta)\mathbb{E}\int_0^T Q_j^2(s)d\mathbf{w}_H(s) = \frac{T}{2}\,.$$

Proof. Denote by $\Psi_T^H(a,\mu_j)$ the Laplace transform of $\int_0^T Q_j^2(s)d\mathbf{w}_H(s)$, namely

$$\Psi_T^H(a,\mu_j(\theta)) = \mathbb{E}\exp\left\{-a\int_0^T Q_j^2(s)d\mathbf{w}_H(s)\right\}, \quad a>0. \tag{A.1}$$

We will use the expression for Ψ_T^H from [6, page 242], and write it as follows

$$\Psi_T^H(a,\mu_j) = \alpha e^{\frac{(\mu_j-\alpha)T}{2}}\left[\Delta_T^H(\mu_j,\alpha)\right]^{-\frac{1}{2}}$$

where $\mu_j = \mu_j(\theta),\ \alpha := \sqrt{\mu_j^2 + 2a}$,

$$\Delta_T^H(\mu_j, \alpha) = \frac{\pi\alpha T e^{-\alpha T}(\alpha^2 - \mu_j^2)}{4\sin(\pi H)} I_{-H}\left(\frac{\alpha T}{2}\right) I_{H-1}\left(\frac{\alpha T}{2}\right) + e^{-\alpha T}\left[\alpha \sinh\left(\frac{\alpha T}{2}\right) + \mu_j \cosh\left(\frac{\alpha T}{2}\right)\right]^2,$$

and I_p is the modified Bessel function of the first kind and order p.

Note that

$$\mathbb{E}\int_0^T Q_j^2(s)d\mathbf{w}_H(s) = -\frac{\partial \Psi_T^H(a,\mu_j)}{\partial a}\Big|_{a=0}.$$

Direct evaluations (for example, using `Mathematica` computer algebra system) give

$$\frac{\partial \Psi_T^H(a,\mu_j)}{\partial a}\Big|_{a=0} = \frac{2 + 2e^{\mu_j T}(1-\mu_j T) - \mu_j \pi T I_{H-1}\left(\frac{\mu_j T}{2}\right) I_{-H}\left(\frac{\mu_j T}{2}\right)\csc(H\pi)}{4\mu_j^2 e^{\mu_j T}},$$

where $\csc(x) = 1/\sin(x)$. By combining formulas (6.106), (6.155), and (6.162) in [1], we conclude that, for all $p \in (-1,1),\ p \neq 0$, we have $I_p(x) \sim e^x/\sqrt{2\pi x}$, $x \to \infty$, that is,

$$\lim_{x\to+\infty} \sqrt{2\pi x}\ e^{-x} I_p(x) = 1. \tag{A.2}$$

Therefore

$$\frac{\partial \Psi_T^H(a,\mu_j)}{\partial a}\Big|_{a=0} \sim \frac{2 + 2e^{\mu_j T}(1-\mu_j T) - e^{\mu_j T}\csc(H\pi)}{4\mu_j^2 e^{\mu_j T}} \sim -\frac{T}{2\mu_j},\ j \to \infty,$$

$$\lim_{j\to\infty} \mu_j \frac{\partial \Psi_T^H(a,\mu_j)}{\partial a}\Big|_{a=0} = -\frac{T}{2},$$

and the lemma is proved. □

Lemma A.2. *For every $\theta \in \Theta$ and $H \in [1/2, 1)$*

$$\lim_{j\to\infty} \mu_j^3(\theta)\mathrm{Var}\left(\int_0^T Q_j^2(s)d\mathbf{w}_H(s)\right) = \frac{T}{2}.$$

Proof. Note that

$$\mathbf{V} := \mathrm{Var}\left(\int_0^T Q_j^2(s)d\mathbf{w}_H(s)\right) = \left[\frac{\partial^2 \Psi_T^H(a,\mu_j)}{\partial a^2} - \left(\frac{\partial \Psi_T^H(a,\mu_j)}{\partial a}\right)^2\right]_{a=0}, \tag{A.3}$$

with Ψ_T^H from (A.1). Direct evaluation of the right hand side of (A.3) (for example, using `Mathematica` computer algebra system) gives

$$\mathbf{V} = \frac{1}{8\mu_j^4 e^{2T\mu_j}}\Big(2 - 8e^{\mu_j T}(1+\mu_j T) + 2e^{2\mu_j T}(-5+2\mu_j T) \\ + \pi\mu_j T\csc(\pi H)\big[-2e^{\mu_j T}\mu_j T I_{1-H}\left(\frac{\mu_j T}{2}\right) I_{H-1}\left(\frac{\mu_j T}{2}\right) \\ + I_{-H}\left(\frac{\mu_j T}{2}\right)\{4(-1+e^{\mu_j T}(1+\mu_j T))I_{H-1}\left(\frac{\mu_j T}{2}\right) \\ - 2e^{\mu_j T}\mu_j T I_H\left(\frac{\mu_j T}{2}\right) + \pi\mu_j T I_{H-1}^2\left(\frac{\mu_j T}{2}\right) I_{-H}\left(\frac{\mu_j T}{2}\right)\csc(H\pi)\}\big]\Big),$$

where $\csc(x) = 1/\sin(x)$ and I_p is the modified Bessel function of the first kind and order p.

Using (A.2), we conclude that

$$\lim_{j\to\infty}\mu_j^3(\theta)\mathbf{V} = \lim_{j\to\infty}\mu_j^3\Big(\frac{-10+4\csc(H\pi)+\csc^2(H\pi)}{8\mu_j^4} + \frac{1}{4\mu_j^4 e^{2\mu_j T}} \\ - \frac{\csc(H\pi)+2+2\mu_j T}{2\mu_j^4 e^{\mu_j T}} + \frac{T}{2\mu_j^3}\Big) \\ = \frac{T}{2}$$

and complete the proof of the lemma. □

Department of Applied Mathematics, Illinois Institute of Technology, 10 West 32nd Str, Bld E1, Room 208, Chicago, IL 60616, USA

E-mail address: igor@math.iit.edu, http://math.iit.edu/∼igor

Department of Mathematics, University of Southern California, 3620 S. Vermont Avenue, KAP 108, Los Angeles, CA 90089, USA

E-mail address: lototsky@math.usc.edu, http://math.usc.edu/∼lototsky

Department of Mathematics, Faculty of Applied Sciences, University of West Bohemia, Univerzitní 22, 306 14, Plzeň, Czech Republic

E-mail address: honik@kma.zcu.cz, http://www.kma.zcu.cz/Pospisil