PARAMETER ESTIMATION FOR SPDES WITH MULTIPLICATIVE FRACTIONAL NOISE

IGOR CIALENCO

Abstract. We study parameter estimation problem for diagonalizable stochastic partial differential equations driven by a multiplicative fractional noise with any Hurst parameter $H \in (0,1)$. Two classes of estimators are investigated: traditional maximum likelihood type estimators, and a new class called closed-form exact estimators. Finally the general results are applied to stochastic heat equation driven by a fractional Brownian motion.

AMS 2010: Primary 60H15; Secondary 62F12, 60G22

Keywords: Asymptotic normality, parameter estimation, stochastic PDE, multiplicative noise, singular models, stochastic evolution equations

1. Introduction

Parameter estimation problem for stochastic partial differential equation has been of great interest in the past decade, and besides being a challenging theoretical problem, it finds its roots and motivations from various applied problems: fluid dynamics [12, 31], biology [9, 10], finance [1, 2, 8], meteorology [5] etc. At general level the problem is to find or estimate the model parameter $\vartheta$ (could be a vector) based on observations of the underlying process $u_\vartheta$ which is assumed to be a solution of a stochastic evolution equation in finite or infinite dimensional space. We will follow traditional continuous time approach and assume that the solution $u_\vartheta(t)$ is observed continuously in time $t \in [0,T]$. From statistical point of view, we suppose that there exists a family of probability measures $P_\vartheta$ that depends on parameter $\vartheta \in \Theta \subset \mathbb{R}^n$, and each $P_\vartheta$ is the distribution of a random element. Assuming that a realization of one random element corresponds to a particular value $\vartheta_0$, the goal is to estimate this parameter from given observations. One approach is to select parameter $\vartheta$ that most likely produces the observations. This method assumes that the problem is regular or absolutely continuous, which means that there exists a reference probability measure $Q$ such that all measures $P_\vartheta$, $\vartheta \in \Theta$, are absolutely continuous with respect to $Q$. Then Radon-Nikodym derivative $dP_\vartheta/dQ$, also called the likelihood

---

Date: May 26, 2010.
Research supported in part by the the NSF grant DMS-0908099.
ratio, exists, and the Maximum Likelihood Estimator (MLE) $\widehat{\vartheta}$ of the parameter of interest is computed by maximizing the likelihood ratio with respect to $\vartheta$. Usually $\widehat{\vartheta} \neq \vartheta$ and the problem is to study the convergence of MLE to the true parameter as more information arrives (for example as time passes or by decreasing the amplitude of noise). If the measures $P_{\vartheta}$ are singular for different parameters $\vartheta$, then the model is called singular, and usually the parameter can be found exactly, at least theoretically. While all regular models are to some extend the same, each singular model requires individual approach. For example, estimating the drift coefficient for finite-dimensional stochastic differential equations is typically a regular problem, and the parameter can be estimated by means of MLEs, while estimating the diffusion (volatility) coefficient is a singular problem and one can find the diffusion coefficient exactly through quadratic variation of the underlying process. For some finite-dimensional systems, estimating the “drift coefficient” is also a singular problem, and as shown in Khasminskii et al. [18] the estimators have nothing to do with MLEs. Generally speaking statistical inference for finite-dimensional diffusions has been studied widely, and there are established necessary and sufficient conditions for absolute continuity of corresponding measures (see, for example [23], [21] and references therein). Some of these results have been extended to infinite dimensional systems in particular to parabolic Stochastic Partial Differential Equations (SPDE). It turns out that in many cases the estimation of drift coefficient for SPDEs is a singular problem, and as general theory suggests one can find the parameter “exactly”. One of the first fundamental result in this area that explores this singularity is due to Huebner et al. [17]. The idea is to approximate the original singular problem by a sequence of regular problems for which MLEs exist. The approximation was done by considering Galerkin-type projections of the solution on a finite-dimensional space where the estimation problem becomes regular, and it was proved that as dimension of the projection increases the corresponding MLE will converge to the true parameter. In [15, 16, 26, 27], the problem was extended to a general parabolic SPDE driven by additive noise and the convergence of the estimators was given in terms of the order of the corresponding differential operators. For recent developments and other types of inference problems in SPDEs see a survey paper by Lototsky [24] and references therein. Statistical inference for SPDEs driven by multiplicative noise is a more challenging problem. First and only attempt to study equations with multiplicative noise is given in [6], by considering Wiener (not fractional) type noise without spatial correlation structure. Besides MLE type estimators, a completely new class of exact estimators were found due essentially to the very singular nature of the problem.

The aim of this note is to study parameter estimation problem for stochastic evolution equations driven by a multiplicative fractional noise with
following dynamics

\[
\begin{align*}
  u(t) &= u(0) + \int_0^t (A_0 + \theta A_1)u(s)\,ds + \int_0^t M u(s)\,dW^H(s),
\end{align*}
\]

where \( A_0, A_1 \) and \( M \) are some known linear operators, \( W^H \) is a fractional Brownian motion with a Hurst parameter \( H \in (0,1) \), and \( \theta \) is a real parameter belonging to a bounded set \( \Theta \subset \mathbb{R} \). For now, assume that the stochastic integral with respect to fractional Brownian Motion \( W^H \) is well-defined, while the exact meaning will be specified in Section 2.1. The main goal is to estimate the parameter \( \theta \) based on the observations of the underlying process \( u(t), t \in [0,T] \). Similar problem for SPDEs driven by additive space-time fractional noise was investigated in [7, 28, 32]. Estimation of drift coefficient for finite-dimensional fractional Ornstein-Uhlenbeck and similar processes has been investigated by Tudor and Viens [33] for \( H \in (0,1) \), Kleptsyna and Le Breton [19] for \( H \in [1/2,1) \), by developing Girsanov type theorems and finding MLEs. Berzin and Leon [3] estimate simultaneously both drift and diffusion coefficients. Least square estimators for drift coefficients were established by Hu and Nualart [13], and MLE type estimators for discretely observed process by Hu et al. [14]. For a general theory, including Girsanov Theorem and some results on statistical inference, for finite dimensional diffusions driven by fractional noise see also the monograph by Mishura [29].

In this paper we continue to explore the impact of the noise in infinite-dimensional evolution equations and its implications on statistical inference. Besides its theoretical roots, this problem is also motivated by increasing demand in modeling various phenomena by SPDEs driven by fractional noise [5,11]. We assume that the solution of (1) is observed at every \( t \in [0,T] \), and hence each Fourier coefficient \( u_k(t) = (u_k(t), h_k)_H \) is observable for every \( t \in [0,T] \), where \( H \) is a Hilbert space in which the solution leaves and \( h_k, k \geq 1 \), is a CONS in \( H \). All results are stated in terms of Fourier coefficients \( u_k \). In the first part of Section 2 we set up the problem and establish the existence and uniqueness of the solution of the corresponding SPDE. In Subsection 2.2 we introduce the main notations and find the MLE for fractional Geometrical Brownian Motion (which is not covered explicitly in any other sources, at our best knowledge). In Section 3 we study the estimators of drift coefficient \( \theta \) of equation (1) based on MLE of the corresponding Fourier coefficients. We establish sufficient conditions on operators \( A_1, A_1, A_1 \) and \( M \), that guarantee efficiency and asymptotic normality of the estimators and some of their versions. Section 4 is dedicated to investigation of a new type of estimators called \textit{closed-form exact} estimators, similar to those studied in [6]. We show that \( \theta \) can be found \textit{exactly} by knowing just several (usually two) Fourier coefficients. Moreover, by the same technics we found an exact estimator of the Hurst parameter \( H \) too, in both regimes, \( \theta \) known and unknown. Of course there are many other methods of finding the Hurst parameter, but it is out of scope of this publication to apply them to our equation. Some of
the results follow from simple algebraic evaluations, but the very existence of such estimators is amazing and gives a better understanding of the nature of the problem’s singularity. Also, we want to mention that, despite of memory property of the fractional Brownian Motion which is spilled over the solution too, the exact estimators are based only on observations at time zero and some future time $T$. In contrast, the MLEs require observation of the whole trajectory $u(t), t \in [0, T]$. We conclude the paper with two examples which are of interest along: one dimensional stochastic heat equation with parameter $\theta$ next to Laplace operator, and a multi-dimensional version of stochastic heat equation with $\theta$ next to a lower order operator.

While we assume that data is sampled continuously in time, in practice usually this is not the case. For the MLEs derived in Section 3 the problem is reduced to approximate some integrals of a deterministic function with respect to the solution $u$ and eventually to the fractional Brownian motion. However, the Exact Estimators from Section 4 depend only on the values of the solution at initial time $t = 0$ and some future time $t = T$, and thus do not depend on how the solution is observed in time.

2. Preliminary results

2.1. The equation and existence of the solution. Let $H$ be a separable Hilbert space with the inner product $(\cdot, \cdot)_0$ and the corresponding norm $\| \cdot \|_0$. Let $\Lambda$ be a densely-defined linear operator on $H$ with the following property: there exists a positive number $c$ such that $\|\Lambda u\|_0 \geq c \|u\|_0$ for every $u$ from the domain of $\Lambda$. Then the operator powers $\Lambda^\gamma$, $\gamma \in \mathbb{R}$, are well defined and generate the spaces $H^\gamma$: for $\gamma > 0$, $H^\gamma$ is the domain of $\Lambda^\gamma$; $H^0 = H$; for $\gamma < 0$, $H^\gamma$ is the completion of $H$ with respect to the norm $\| \cdot \|_\gamma := \|\Lambda^{-\gamma} \cdot\|_0$ (see for instance Krein at al. [20]). By construction, the collection of spaces $\{H^\gamma, \gamma \in \mathbb{R}\}$ has the following properties:

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

$$\langle u_1, u_2 \rangle_{\gamma,m} = \langle \Lambda^{\gamma-m} u_1, \Lambda^{\gamma+m} u_2 \rangle_0,$$

where $u_1 \in H^{\gamma-m}$, $u_2 \in H^{\gamma+m}$.

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ be a stochastic basis with usual assumptions.

**Definition 1.** A fractional Brownian motion with a Hurst parameter $H \in (0, 1)$ is a Gaussian process $W^H$ with zero mean and covariance

$$\mathbb{E}W^H(t)W^H(s) = \frac{1}{2}(t^{2H} + s^{2H} - |t-s|^{2H}), \quad t, s \geq 0.$$
Consider the following evolution equation

\[
\begin{align*}
\frac{du(t)}{dt} &= (A_0 + \theta A_1)u(t) + Mu(t)dW^H(t), \quad 0 < t < T, \\
u(0) &= u_0,
\end{align*}
\]

where \(A_0, A_1, M\) are linear operators in \(H\), \(u_0 \in H\), \(W^H\) is a fractional Brownian Motion with Hurst parameter \(H \in (0, 1)\), and \(\theta\) is a scalar parameter belonging to an open set \(\Theta \subset \mathbb{R}\).

**Definition 2.** Equation (2) is called diagonalizable if the operators \(A_0, A_1\) and \(M\) have point spectrum and a common system of eigenfunctions \(\{h_j, j \geq 1\}\).

Denote by \(\rho_k, \nu_k, \) and \(\mu_k\) the eigenvalues of the operators \(A_0, A_1,\) and \(M\):

\[
A_0 h_k = \rho_k h_k, \quad A_1 h_k = \nu_k h_k, \quad M h_k = \mu_k h_k, \quad k \geq 1,
\]

and also denote by \(\alpha_k(\theta) := \rho_k + \theta \nu_k, k \geq 1\), the eigenvalues of operator \(A_0 + \theta A_1\). Without loss of generality we assume that the operator \(\Lambda\) has the same eigenfunctions as operators \(A_0, A_1, M\): \(\Lambda h_k = \lambda_k h_k, k \geq 1\).

**Theorem 3.** Assume that equation (2) is diagonalizable and the initial conditions \(u_0\) is deterministic and belongs to \(H^\gamma\). Then the process \(u\) defined by

\[
u(t) = \sum_{k \geq 1} u_k(t)h_k,
\]

where

\[
u_k(t) = u_k(0) \exp\left(\alpha_k(\theta)t - \frac{1}{2}\mu_k^2 t^{2H} + \mu_k W^H(t)\right),
\]

is an \(H^\gamma\)-valued stochastic process if and only if there exists a constant \(C = C(t, H, \theta)\) such that

\[
2\alpha_k(\theta)t + \mu_k^2 t^{2H} \leq C, \quad k \geq 1.
\]

**Proof.** Since \(W^H(t)\) is a Gaussian random variable with zero mean and variance \(t^{2H}\), we have

\[
\mathbb{E}[u_k(t)]^2 = u_k(0)^2 \exp\left(2\alpha_k(\theta)t + \mu_k^2 t^{2H}\right).
\]

Hence,

\[
\mathbb{E}\|u(t)\|_{\gamma}^2 = \sum_{k \geq 1} \lambda_k^{2\gamma} \mathbb{E}[|u_k(t)|^2] = \sum_{k \geq 1} \lambda_k^{2\gamma} u_k^2(0) \exp\left(2\alpha_k(\theta)t + \mu_k^2 t^{2H}\right).
\]

The last series converges if and only if (6) is satisfied uniformly in \(k\), and the theorem follows. \(\square\)
Note that in particular if $H = 1/2$ and the equation (2) is parabolic in the triple $(H^{\gamma+m}, H^{\gamma}, H^{\gamma-m})$, for some positive $m$ and real $\gamma$, i.e. there exist positive real numbers $\delta$, $C_1$ and a real number $C_2$ such that, for all $k \geq 1$ and all $\theta \in \Theta$,

\[
\begin{align*}
\lambda_k^{-2m}\rho_k + \theta \nu_k &\leq C_1; \\
2(\rho_k + \theta \nu_k) + \mu_k^2 + \delta \lambda_k^{2m} &\leq C_2,
\end{align*}
\]

then (6) is satisfied for every $t \geq 0$. If in addition to parabolicity conditions (7) we assume that $\lim_{k \to \infty} \mu_k^2/|\alpha_k| = 0$, then (6) is satisfied uniformly in $t \in [0,T]$ for every $H \in [0,1]$. For more discussions on conditions similar to (6) see for instance [25].

The functions $u_k$ formally represent the Fourier coefficients of the solution of equation (2) with respect to the basis $\{h_k\}_{k \geq 1}$ and the uniqueness of $u$ follows. Since the equation is diagonalizable, naturally we conclude that formally $u_k$ has the following dynamics

\[(8) \quad du_k(t) = (\theta \nu_k + \rho_k)u_k(t)dt + \mu_k u_k(t)dW^H(t), \quad k \geq 1, t \geq 0.\]

Specifying the stochastic integration in (2) is equivalent to specifying in what sense we understand the integration with respect to fractional Brownian Motion for the Fourier coefficients (8). Consequently, since the equation has constant coefficients, specifying the solution of (8) is the same as to stipulate the sense of stochastic integration in (8). If the integration is understood in Wick sense then $u_k$, $k \geq 1$, defined in (5) is the unique solution of equation (8) for all $H \in (0,1)$ (see for instance [4], Theorem 6.3.1). All results stated here are easily transferable to any other form of integration, by caring out the relationship between different form of integration and consequently adjusting the form of the solution of equation (8) (for comparison of various form of integration with respect to fBM see [4], Chapter 6). Our choice was just to have a unified theory and same formulas for all $H \in (0,1)$.

**Definition 4.** The process $u$ constructed in Theorem 3 is called the solution of equation (2).

It should be mentioned that the above result, with some obvious adjustments, also holds true for diagonalizable equations driven by several independent fractional Brownian Motions, even with different Hurst parameters.

### 2.2. Parameter estimation for geometric fractional Brownian motion

In this section we will present some auxiliary results about parameter estimation for one dimensional diffusion processes driven by multiplicative fractional noise. For similar results for equations with additive noise see for instance Kleptsyna and Le Breton [19], Tudor and Viens [33], or Mishura [29], Chapter 6. The results essentially follow from Girsanov type theorem for diffusions driven by fractional Brownian motion.
Let $\Gamma$ denote the Euler Gamma-function. Following Mishura [29] we introduce the following notations

$$C_H = \left( \frac{\Gamma(3 - 2H)}{2H\Gamma(\frac{3}{2} - H)^3\Gamma(\frac{1}{2} + H)} \right)^{\frac{1}{2}},$$

(9)

$$l_H(t, s) = C_H s^{\frac{1}{2} - H}(t - s)^{\frac{1}{2} - H} \mathbb{1}_{0 < s < t},$$

$$M_t^H := \int_0^t l_H(t, s) dW_t^H,$$

where $H \in (0, 1)$, and the integration with respect to fractional Brownian Motion is understood in Wiener sense (for more details see [29], Chapter 1). The process $M_t^H$ is a martingale, also called the fundamental martingale associated with fractional Brownian motion (see for instance [30] or [29], Theorem 1.8.1). $M_t^H$ has quadratic characteristic $\langle M^H \rangle_t = t^{2 - 2H}$, and by Lévy theorem, there exists a Wiener process $\{B_t, t \geq 0\}$ on the same probability space such that

$$M_t^H = (2 - 2H)^{\frac{1}{2}} \int_0^t s^{\frac{1}{2} - H} dB_s.$$

Moreover, $\sigma(W_t^H, 0 \leq s \leq t) = \sigma(B_s, 0 \leq s \leq t)$.

Let us consider the stochastic process of the form

$$X_t = X_0 \exp \left( \theta t - \frac{1}{2}\sigma^2 t^{2H} + \sigma W_t^H \right), \quad t \geq 1,$$

which can be called the Geometric Fractional Brownian Motion, and as mentioned in the previous subsection it is the unique solution of the stochastic equation

$$dX_t = \theta X_t dt + \sigma X_t dW_t^H, \quad X_0 = x_0, \quad t \in [0, T].$$

Let $Y_t := \ln(X_t/X_0) = \theta t - \frac{1}{2}\sigma^2 t^{2H} + \sigma W_t^H$, and consider the process $\tilde{Y}_t := \int_0^t l_H(t, s) dY_s$. Note that observing one path of the process $\{Y_s, 0 \leq s \leq t\}$ implies that the one path of process $\{\tilde{Y}_s, 0 \leq s \leq t\}$ is observable too. By (9) we have

$$\tilde{Y}_t = \sigma M_t^H + \theta b_1 t^{2-2H} - \sigma^2 H b_2 t, \quad t > 0,$$

where $b_1 = C_H B(3/2 - H, 3/2 - H)$, $b_2 = C_H B(1/2 + H, 3/2 - H)$, and $B(x, y) = \Gamma(x)\Gamma(y)/\Gamma(x + y)$ is the Euler beta function.

For a fixed parameter $\theta \in \Theta$, let us denote by $\mathbb{P}_\theta$ the distribution of the process $\tilde{Y}_t$ and by $\mathbb{P}_0$ the distribution of the process $\tilde{Y}_t^0 := \sigma M_t^H = \sigma b_0 \int_0^t s^{1/2-2H} dB_s$. The measure $\mathbb{P}_\theta$ is absolutely continuous with respect to $\mathbb{P}_0$ and the Radon-Nikodym derivative, or the likelihood ratio, has the following form (see for instance [23], Theorem 7.19 or apply classical Girsanov
Theorem for martingales)

\[
\frac{dP_\theta}{dP_0}(\tilde{Y}_t) = \exp\left(-\int_0^t \frac{\theta(2-2H)b_1 s^{1-2H} - \sigma^2 H b_2^2}{\sigma^2 b_0^2 s^{1-2H}} d\tilde{Y}_s + \frac{1}{2} \int_0^t \frac{[\theta(2-2H)b_1 s^{1-2H} - \sigma^2 H b_2^2]^2}{\sigma^2 b_0^2 s^{1-2H}} ds\right).
\]

The MLE is obtained by maximizing the log-likelihood ratio with respect to \(\theta\). Since

\[
\frac{\partial}{\partial\theta} \ln \frac{dP_\theta}{dP_0}(\tilde{Y}_t) = -\frac{(2-2H)b_1}{\sigma^2 b_0} \tilde{Y}_t + \theta \frac{(2-2H)b_2^2 t^{2-2H}}{\sigma^2 b_0^2} - \frac{(2-2H)b_1 H b_2 t}{b_0^2},
\]

the MLE for parameter \(\theta\) has the form

\[
\hat{\theta}_t = \frac{\tilde{Y}_t}{b_1 t^{2-2H}} + \frac{\sigma^2 H b_2}{b_1 t^{1-2H}}.
\]

**Proposition 1.** The estimator \(\hat{\theta}_t, t > 0\), is an unbiased estimator for parameter \(\theta_0\): \(\lim_{t \to \infty} \hat{\theta}_t = \theta_0\) with probability one, i.e. \(\hat{\theta}_t\) is a strong consistent estimator of \(\theta_0\); \(t^{1-H}(\hat{\theta}_t - \theta_0)\) converges in distribution to a Gaussian random variable with zero mean and variance \(\sigma^2 / b_1^2\).

**Proof.** Using the definition of the process \(\tilde{Y}_t\), we represent \(\hat{\theta}_t\) as follows

\[
\hat{\theta}_t = \frac{Y_t}{b_1 t^{2-2H}} + \frac{\sigma M_t^H}{b_1 t^{1-2H}},
\]

where \(\theta_0\) is the true parameter.

The unbiasedness and asymptotic normality follows immediately from (12) and the fact that \(M_t^H\) is a Gaussian random variables with zero mean and variance \(t^{2-2H}\). Since \(M_t^H\) is a square integrable martingale with unbounded quadratic characteristic \(t^{2-2H} \to \infty\), as \(t \to \infty\) a.s., by Law of Large Numbers for Martingales [22], Theorem 2.6.10, \(M_t^H / (M^H)_t \to 0\) a.s., and hence consistency follows.

Note that, in particular, for \(H = 1/2\) we have \(b_1 = b_2 = 1\), and we recover the classical estimator for the drift coefficient of geometric Brownian Motion

\[
\hat{\theta}_t = \frac{Y_t}{t} + \frac{\sigma^2}{2} = \frac{\ln(X_t/X_0)}{t} + \frac{\sigma^2}{2} = \theta_0 + \frac{\sigma W_t}{t}, \quad t > 0,
\]

and its corresponding asymptotic behavior.

### 3. Maximum Likelihood Estimator for SPDEs

Consider the diagonalizable equation

\[
du(t) = (A_0 + A_1)u(t)dt + M_u(t)dW^H(t),
\]

with solution \(u(t) = \sum_{k \geq 1} u_k(t)h_k\) given by (5). As mentioned in Introduction, if \(u\) is observable, then all its Fourier coefficients \(u_k\) can be computed. Thus, we assume that the processes \(u_1(t), \ldots, u_N(t)\) can be observed for all
PARAMETER ESTIMATION FOR SPDES WITH MULTIPlicative FRACTIONAL NOISE

Let $t \in [0, T]$ and the problem is to estimate the parameter $\theta$ based on these observations. Also, we assume that the Hurst parameter $H \in (0, 1)$ is known for now.

By Definition 4 of the solution of equation (13), the Fourier coefficients $u_k$, $k \in \mathbb{N}$, have the following dynamics

$$du_k(t) = \alpha_k(\theta)u_k(t)dt + \mu_k u_k(t)dW^H(t), \quad t \in [0, T],$$

where $\alpha_k(\theta) = \rho_k + \theta \nu_k$, $k \in \mathbb{N}$.

For every non-zero $u_k(0)$, $k \in \mathbb{N}$, denote by $v_k(t) = \ln(u_k(t)/u_k(0))$, and

$$\tilde{v}_k(t) = \int_0^t l(t, s)dv_k(s),$$

where $l(\cdot, \cdot)$ is defined in (9). By results of Section 2.2, it follows that there exists a Maximum Likelihood Estimator for $\alpha_k(\theta)$ and it has the form

$$\hat{\alpha}_k(\theta) = \frac{\tilde{v}_k(t)}{b_1 t^{2-H}} + \frac{H b_2 \mu_k^2}{b_1 t^{1-2H}}, \quad k \geq 1.$$

Since $\alpha_k(\theta)$ is a strictly monotone function in $\theta$, by invariant principle of MLE under invertible transformations, we can find an MLE for the parameter $\theta$

$$\hat{\theta}_{k,T} = \frac{\tilde{v}_k(t)}{v_k b_1 t^{2-H}} + \frac{H b_2 \mu_k^2}{v_k b_1 t^{1-2H}} - \frac{\rho_k}{v_k}, \quad k \geq 1, \quad t \in [0, T].$$

Using the definition of the process $\tilde{v}_k$, the estimator $\hat{\theta}_{k,T}$ can be represented as follows

$$\hat{\theta}_{k,T} = \theta_0 + \frac{\mu_k M^H_T}{b_1 \nu_k T^{2-2H}},$$

and by similar arguments to the proof of Proposition 4, we have the following result.

**Theorem 5.** Assume that equation (13) is diagonalizable and $u_0 \in H^\gamma$ for some $\gamma \in \mathbb{R}$. Then,

1. For every $k \geq 1$ and $T > 0$, $\hat{\theta}_{k,T}$ is an unbiased estimator of $\theta_0$.
2. For every fixed $k \geq 1$, as $T \to \infty$, the estimator $\hat{\theta}_{k,T}$ converges to $\theta_0$ with probability one and $T^{1-H}(\hat{\theta}_{k,T} - \theta_0)$ converges in distribution to a Gaussian random variable with zero mean and variance $\frac{\mu_k^2}{b_1^2 \nu_k^2}$.
3. If, in addition,

$$\lim_{k \to \infty} \left| \frac{\nu_k}{\mu_k} \right| = 0,$$

then for every fixed $T > 0$, $\lim_{k \to \infty} \hat{\theta}_{k,T} = \theta_0$ with probability one, and $|\nu_k/\mu_k(\hat{\theta}_{k,T} - \theta_0)$ converges in distribution to a Gaussian random variable with zero mean and variance $T^{2H-2}/b_1^2$.

**Remark 6.** The parabolicity conditions (7) and MLE consistency condition (18) in general are not connected. In terms of operator’s order, parabolicity states that the order of operator $\mathcal{M}$ from the diffusion term is smaller than
half of the order of the operators $\mathcal{A}_0$ and $\mathcal{A}_1$ from deterministic part. Condition (18), that guarantees the consistency of MLE as number of Fourier coefficients increases, assumes that the order of operator $\mathcal{M}$ from the diffusion part does not exceed the order of the operator $\mathcal{A}_1$ from deterministic part that contains the parameter of interest $\theta$.

By Theorem 5 it follows that the consistency and asymptotic normality of the estimators $\hat{\theta}_{k,T}$ can be achieved in two ways: by increasing time $T$ or by increasing the number of Fourier coefficients $k$. In both cases the quality of the estimator is improved by decreasing its variance.

It is interesting to note that $\text{Var} \left( \hat{\theta}_{k,T} - \theta_0 \right) = \mu_k^2 T^{2H-2} / b_1^2 \nu_k^2$ also depends on Hurst parameter $H$. For $H > 1/2$ the constant $1/b_1$ is close to one, and increases as function of $H$ for $H \in (0, 1/2)$. The function $T^{2H-2}$ increases in $H$ for any $t > 1$. The constants $\mu_k$ and $\nu_k$, $k \geq 1$, do not depend on $H$. Overall, $T^{2H-2} / b_1^2$ increases in $H$ for any $t > 1$ and thus quality of the estimators is higher for smaller $H$.

A natural question is whether we can improve the quality of the estimators by considering several Fourier coefficients $u_k(t)$. The answer is that by statistical methods used above this is not possible. First, note that the measures associated to any two or more processes $u_k$ are singular, and thus MLE does not exist for such vector-valued functions. In other words, by considering two or more Fourier coefficients $u_k$, we get a singular model, a fact that will be explored in the next section. Also, since each process $u_k$ is driven by the same noise, each individual Fourier coefficient $u_k$ contains the same amount of information: the sigma-algebra generated by $u_k(t)$, $t \in [0, T]$ coincides with the sigma-algebra generated by $W^H(t)$, $t \in [0, T]$. However, the speed of convergence of the sequence $\hat{\theta}_{k,T}$ can be improved by using accelerating convergence techniques from numerical analysis. Two methods have been discussed in details in [6]: the weighted average method and Aitken’s $\Delta^2$ method. For the sake of completeness, we will state here the corresponding results applied to the sequence $\{\hat{\theta}_{k,T}\}_{k \geq 1}$.

**Weighted averaging.** Suppose that $\beta_k$, $k \geq 1$, is a sequence of non-negative numbers such that $\sum_{k \geq 1} \beta_k = +\infty$, and consider the weighted averaging estimator

$$ (19) \quad \hat{\theta}_{(N,T)} = \frac{\sum_{k=1}^N \beta_k \hat{\theta}_{k,T}}{\sum_{k=1}^N \beta_k}, \quad N \geq 1, \quad T > 0. $$

Then (a) $\hat{\theta}_{(N,T)}$ is an unbiased estimator of $\theta_0$ for every $N \geq 1$ and $T > 0$; (b) $\lim_{T \to \infty} \hat{\theta}_{(N,T)} = \theta_0$ a.e. for every $N \geq 1$ (consistency in $T$); (c) if in addition the consistency condition (18) is fulfilled, then $\lim_{N \to \infty} \hat{\theta}_{(N,T)} = \theta_0$ with probability one for every $T > 0$ (consistency in $N$).
Aitken’s $\Delta^2$ method. Define the following sequence of estimators

\begin{equation}
\hat{\theta}_k = \hat{\theta}_{k,T} - \frac{(\hat{\theta}_{k+1,T} - \hat{\theta}_{k,T})^2}{\hat{\theta}_{k+2,T} + 2\hat{\theta}_{k+1,T} - \hat{\theta}_{k,T}}.
\end{equation}

One can show that the new sequence $\hat{\theta}_k$ converges to the true parameter $\theta_0$ with probability one. Moreover, if $\mu_k/\nu_k \sim \alpha k^{-\delta}$ for some $\alpha, \delta > 0$, then

\begin{equation}
\frac{\mathbb{E}(\hat{\theta}_k - \theta_0)^2}{\mathbb{E}(\hat{\theta}_{k,T} - \theta_0)^2} \sim \frac{1}{(1 + \delta_1)^2},
\end{equation}

for some $\delta_1 > 0$. If $\mu_k/\nu_k = (-1)^k/k$, then

\begin{equation}
\frac{\mathbb{E}(\hat{\theta}_k - \theta_0)^2}{\mathbb{E}(\hat{\theta}_{k,T} - \theta_0)^2} \sim \frac{c}{k^2}, \quad c > 0.
\end{equation}

In both cases, the new sequence $\hat{\theta}_k$ converges faster than $\hat{\theta}_{k,T}$ to $\theta_0$.

The proofs of the above results follow from Theorem 5 and some direct computations, which will be omitted here.

4. Exact Estimators

In regular models the unknown parameter can be found only approximately, and the consistency is gained either in large sample or small noise regime. For singular models the parameter can be found exactly. For example, if all Fourier coefficients of the solution $u$ of equation (2) are known, according to the results from previous sections, one can find the value of $\theta_0$ exactly, on any interval of time $[0, T]$. The possibility to evaluate $\theta_0$ exactly is based on singularity of the measures generated by $u^\theta$ for different values of $\theta$. However, while theoretically it is possible to estimate the true parameter exactly, in practice we (or computer) can perform only a finite number of operations. Recall that the measures associated to an individual Fourier coefficient $u^\theta_k$ are regular, while a vector consisting of any two or more Fourier coefficients will produce measures that are singular. In this section we will explore this singularity, and show that in fact the true parameter can be estimated exactly from a finite number of Fourier coefficients. Moreover, the described method allow to find both parameters $\theta$ and $H$, either individually or simultaneously.

Following [6] we say that an estimator is closed-form exact if it produces the exact value of the parameter of interest after finite number of additions, substraction, multiplications, and divisions on the elementary functions of the observations.

Closed-form exact estimators exists for the model (2) if we assume that observations are $u_k(t), k \geq 1, \ t \in [0, T]$. For every non-zero Fourier coefficient $u_k$ of the form (14), set $v_k(t) = \ln(u_k(t)/u_k(0)), \ t \in [0, T]$. Then

\begin{equation}
v_k(t) = (\rho_k + \theta \nu_k)t - \frac{1}{2} \mu_k^2 H^2 + \mu_k W^H(t).
\end{equation}
Case 1. $\theta$ unknown, $H$ known. Assume that $\nu_k \mu_m \neq \nu_m \mu_k$ for some $k, m \in \mathbb{N}$. Then, taking (21) for these $k$ and $m$, by direct arithmetic evaluations, one gets the exact estimator of the parameter $\theta$

$$\theta = \frac{\mu_m \nu_k - \mu_k \nu_m + (\rho_m \mu_k - \rho_m \mu_m) t + \frac{1}{2}(\mu_k^2 \mu_m - \mu_m^2 \mu_k) t^{2H}}{t(\nu_k \mu_m - \nu_m \mu_k)},$$

for any $t > 0$ and $k, m \in \mathbb{N}$ for which $\nu_k \mu_m \neq \nu_m \mu_k$.

Note that if $\mu_k = \mu_m$ then the above exact estimator does not depend on $H$, and $\theta$ can be evaluated even if $H$ is unknown. This is the case, for example, if $\mathcal{M}$ is the identity operator (see Example 1 below).

Case 2. $H$ unknown, $\theta$ known. Assume now that the parameter of interest is the Hurst parameter $H$ and assume that $\theta$ is known. By the same arguments as above, one can solve for $H$ the system of two equations generated by (21) for some $k$ and $m$, and get the following exact estimator for $H$

$$H = \frac{1}{2 \ln t} \ln \left[\frac{(\rho_k + \theta \nu_k) t \mu_m - (\rho_m + \theta \nu_m) t \mu_k - \nu_k \mu_m + \nu_m \mu_k}{\frac{1}{2} \mu_k \mu_m (\mu_k - \mu_m)}\right],$$

for any $t > 0$, $k \neq m$, and under assumption that the expression under logarithm is positive and finite.

Case 3. Both $\theta$ and $H$ unknown. Denote by $\alpha_{k,m} := (\nu_k \mu_m - \nu_m \mu_k) t$, $\beta_{k,m} := 1/2(\mu_k^2 \mu_m - \mu_m^2 \mu_k)$ and $\delta_{k,m} := \nu_k \mu_m - \nu_m \mu_k - \rho_k \mu_m t - \rho_m \mu_k t$. Assume that for some $m, k, i, j \in \mathbb{N}$, $\alpha_{k,m} \beta_{i,j} \neq \alpha_{k,m} \beta_{i,j}$. Then the following exact estimator for $\theta$ holds true

$$\theta = \frac{\delta_{k,m} \beta_{i,j} - \delta_{i,j} \beta_{k,m}}{\alpha_{k,m} \beta_{i,j} - \alpha_{i,j} \beta_{k,m}}.$$

If in addition $\delta_{k,m} \alpha_{i,j} \neq \delta_{i,j} \alpha_{k,m}$, then there exists an exact estimator for Hurst parameter $H$ given by

$$H = \frac{1}{2} \log \frac{\delta_{k,m} \alpha_{i,j} - \delta_{i,j} \alpha_{k,m}}{\beta_{k,m} \alpha_{i,j} - \beta_{i,j} \alpha_{k,m}}.$$

Note that for this case, generally speaking, it is sufficient to know only three Fourier coefficients, i.e. some of the indices $k, m, i, j$ can coincide.

Remark 7.

(a) Applying the above idea, closed-form exact estimators can be obtained for equations driven by several fractional Brownian motions, even with different Hurst parameters. If we assume that the noise is driven by $n$ fBMs, and that one of the parameters $\theta$ or $H$ is known, then by considering $n + 1$ Fourier coefficients we can eliminate all noises and get a closed-form estimator as a solution, under some non-degeneracy assumptions. Respectively, if both parameters are unknown, then one can estimate them by considering $n + 2$ Fourier coefficients.
(b) Note that the construction of the exact estimators assumed only the existence of the solution and did not impose any additional assumptions on the order of the operators $A_0$, $A_1$, $M$, in contrast to MLE estimators where the consistency holds only under additional assumptions on order of corresponding operators.

(c) The MLE $\hat{\theta}_{k,T}$ depend on the whole trajectory of the Fourier coefficient $u_k(t)$, $t \in [0,T]$. All exact estimators depend only on initial and terminal value of $u_k$’s.

5. Examples

We conclude the paper with two practical examples where we explore some of the estimators proposed above.

**Example 1.** Stochastic heat equation. Let $\theta$ be a positive number, and consider the following equation

(26) \[ du(t, x) = \theta u_{xx}(t, x)dt + u(t, x)dW^H(t), \quad t > 0, \ x \in (0, \pi), \]

with zero boundary conditions and some nonzero initial value $u(0) \in L_2(0, \pi)$. In this case the operator $A_1$ is the Laplace operator on $(0, \pi)$ with zero boundary conditions that has the eigenfunctions $h_k(x) = \sqrt{2/\pi} \sin(kx)$, $k > 0$, and eigenvalues $\nu_k = -k^2$, $\rho_k = 0$, $\mu_k = 1$, $k > 0$. Assume that $u(t, x)$ is known for $x \in [0, \pi]$ and $t \in [0, T]$, hence $u_k(t) := \int_0^T h_k(x)u(t, x)dx$, $k \in \mathbb{N}$, is observable. Denote by $v_k(t) := \log(u_k(t)/u_k(0))$ for every $k \in \mathbb{N}$, and $u_k(0) \neq 0$. By Theorem 5, the MLE for $\theta$ has the form

\[ \hat{\theta}_k = -\frac{\int_0^T l(T, s)dv_k(s)}{k^2b_1T^{2-2H}} - \frac{Hb_2}{k^2b_1T^{1-2H}}, \quad k \in \mathbb{N}. \]

The exact estimators \[ \hat{\theta}_k \]

for $\theta$ are given by

\[ \theta = \frac{1}{T(m^2 - k^2)} \ln \frac{u_k(T)u_m(0)}{u_m(T)u_k(0)}, \]

for any $k \neq m$ and $T > 0$. Note that the exact estimators do not depend on $H$. However, since $M$ is the identity operator, and $\mu_k = 1$, $k \geq 1$, there are no exact-type estimators for $H$.

**Example 2.** Assume that $G$ is a bounded domain in $\mathbb{R}^d$, and let $\Delta$ be the Laplace operator on $G$ with zero boundary conditions. Then $\Delta$ has only point spectrum with countable many eigenvalues, call them $\sigma_k$, $k \in \mathbb{N}$. Moreover, the set of corresponding eigenvalues forms an orthonormal basis in $L_2(G)$; the eigenvalues can be arranged so that $0 < -\sigma_1 \leq -\sigma_2 \leq \ldots$; the eigenvalues have the asymptotic $\sigma_k \sim k^{2/d}$. In the space $\mathcal{H}^0(G)$ let us consider the following stochastic evolution equation

\[ du(t) = [\Delta u(t) + \theta u(t)]dt + (1 - \Delta)^r u(t)dW^H(t), \]

with some nonzero initial values in $\mathcal{H}^0(G)$, and some $r \in \mathbb{R}$. According to our notations we have the operators $A_0 = \Delta$, $A_1 = I$, $M = (1 - \Delta)^r$, with corresponding eigenvalues $\nu_k = 1$, $\rho_k = \sigma_k$, $\mu_k = (1 + \sigma_k)^r$. The equation
is diagonalizable, and by Theorem 3 it has a unique solution in the triple \((H^0, H^0, H^{-1})\) for any \(r \leq 1/2\).

The maximum likelihood estimator in this case has the form

\[
\hat{\theta}_{N, t} = \frac{\tilde{v}_k(t)}{b_1 \sigma_k t^{2-2H}} + \frac{H b_2 (1 - \sigma_k)^{2r}}{\sigma_k b_1 t^{1-2H}} - \frac{1}{\sigma_k}, \quad t > 0, \ k \in \mathbb{N},
\]

which is an unbiased estimator of the parameter \(\theta\).

**(2a)** *Large time asymptotics.* \(\lim_{t \to \infty} \hat{\theta}_{k, t} = \theta_0\) a.s. for all \(k \geq 1\); \(\lim_{t \to \infty} t^{1-H}(\hat{\theta}_{k, t} - \theta_0) \overset{d}{=} \xi\), where \(\xi \sim \mathcal{N}(0, (1 - \sigma_k)^2/b_1^2)\).

**(2b)** *Consistency in number of spatial Fourier coefficients.* Assume that \(r < 0\). Then \(\lim_{k \to \infty} \hat{\theta}_{k, t} = \theta_0\) a.s., for every \(t > 0\), and the sequence \((1 - \sigma_k)^{-1}(\hat{\theta}_{k, t} - \theta_0)\) converges in distribution to a Gaussian random variable with mean zero and variance \(t^{2H-2}/b_1^2\). If \(r \in [0, 1/2]\) the solution still exists in the space \(H^{0}(G)\), while the estimator \(\hat{\theta}_{k, t}\) is not consistent in \(k\).

**(2b)** *Exact estimators.* Let \(v_k(t) = \ln(u_k(t)/u_k(0))\). Assume that Hurst parameter \(H\) is known. Then we have the following exact estimator for \(\theta\)

\[
\theta = \frac{(1 - \sigma_m)^r v_k - (1 - \sigma_k)^r v_m}{t((1 - \sigma_m)^r - (1 - \sigma_k)^r)} + \frac{\sigma_m(1 - \sigma_k)^r - \sigma_k(1 - \sigma_m)^r}{(1 - \sigma_m)^r - (1 - \sigma_k)^r} + \frac{t^{2H-1} (1 - \sigma_m)^{2r}(1 - \sigma_m)^r - (1 - \sigma_k)^{2r}(1 - \sigma_k)^r}{2(1 - \sigma_m)^r - (1 - \sigma_k)^r},
\]

for any \(k \neq m\) and \(t > 0\).

If \(\theta\) is known, then the Hurst parameter \(H\) can be found by

\[
H = \frac{1}{2} \log_t \left[ \frac{(\sigma_k + \theta)(1 - \sigma_m)^r - (\sigma_m + \theta)(1 - \sigma_k)^r}{2(1 - \sigma_k)^{2r}(1 - \sigma_m)^r - (1 - \sigma_m)^{2r}(1 - \sigma_k)^r} \right],
\]

for any \(k \neq m, \ t > 0\).

Finally one can write the exact estimators (24) and (25) for the case when both parameters \(\theta\) and \(H\) are unknown. Note that the exact estimators exist for all \(r\) as long as the solution exists (maybe in a larger space) and the Fourier coefficients \(u_k(t)\) are computable.

**Acknowledgments**

The author acknowledges support from NSF grant DMS-0908099. The author is grateful to the anonymous referee for very helpful suggestions.

**References**


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