Parameter estimation for the discretely observed fractional Ornstein–Uhlenbeck process and the Yuima R package

Alexandre Brouste · Stefano M. Iacus

Received: 16 December 2011 / Accepted: 31 August 2012 / Published online: 14 September 2012 © Springer-Verlag 2012

Abstract This paper proposes consistent and asymptotically Gaussian estimators for the parameters $\lambda$, $\sigma$, and $H$ of the discretely observed fractional Ornstein–Uhlenbeck process solution of the stochastic differential equation $dY_t = -\lambda Y_t dt + \sigma dW_t^H$, where $(W_t^H, t \geq 0)$ is the fractional Brownian motion. For the estimation of the drift $\lambda$, the results are obtained only in the case when $\frac{1}{2} < H < \frac{3}{4}$. This paper also provides ready-to-use software for the R statistical environment based on the YUIMA package.

1 Introduction

Statistical inference for parameters of ergodic diffusion processes observed on discrete increasing grid have been much studied. Local asymptotic normality (LAN) property of the likelihoods have been shown in Gobet (2002) for elliptic ergodic diffusion, under proper conditions for the drift and the diffusion coefficient, and a mesh $\Delta N$ satisfying

$$\Delta N \to 0 \quad \text{and} \quad N \Delta N \to +\infty$$

when the size of the sample $N$ grows to infinity. Estimation procedure has been studied by many authors, mainly in the one-dimensional case (see, for instance,
Florens-Zmirou 1989; Kessler 1997; Yoshida 1992 in the multidimensional setting). All estimators in the previous works are based on contrasts (for contrasts framework, see Genon-Catalot 1990), assuming in the general case, that for some $p > 1$, as $N \to +\infty$, $N \Delta_N^P \to 0$. In particular, for the Ornstein–Uhlenbeck process, transitions densities are known, and all have been treated, as remarked in Jacod (2001).

In the fractional case, we consider the fractional Ornstein–Uhlenbeck process (fOU), the solution of

$$dY_t = -\lambda Y_t dt + \sigma dW_H^t, \quad t \geq 0,$$

where $W_H^t = (W_H^t, t \geq 0)$ is a normalized fractional Brownian motion (fBM), i.e. the zero mean Gaussian process with covariance function

$$E W_s^H W_t^H = \frac{1}{2} \left( |s|^{2H} + |t|^{2H} - |t-s|^{2H} \right)$$

with Hurst exponent $H \in (0, 1)$. The fOU process is neither Markovian nor a semi-martingale for $H \neq \frac{1}{2}$ but remains Gaussian and ergodic for $\lambda > 0$ (see Cheridito et al. 2003). For $H > \frac{1}{2}$, it even presents the long-range dependence property that makes it useful for different applications in biology and physics (with the Fractional Langevin Equation), ethernet traffic (Bregni and Erangoli 2005; Willinger et al. 1995) or finance (Xiao et al. 2011).

Statistical large sample properties of maximum likelihood estimator of the drift parameter in the continuous observations case have been treated in Bercu et al. (2010), Brouste and Kleptsyna (2010), Cialenco et al. (2009), Kleptsyna and Breton (2002) for different applications. Further, asymptotic properties of the Least Squares Estimator have been studied in Hu and Nualart (2010).

In the discrete and fractional case, we can cite few works on the topic. On the one hand, very recent works give methods to estimate the drift $\lambda$ by contrast procedure (Bertin et al. 2011; Hu et al. 2011; Ludena 2004; Neuenkirch and Tindel 2011) or the drift $\lambda$ and the diffusion coefficient $\sigma$ with discretization procedure of integral transform (Xiao et al. 2011). In these papers, the Hurst exponent is supposed to be known and only consistency is obtained. On the other hand, methods to estimate the Hurst exponent $H$ and the diffusion coefficient are presented in Berzin and Leon (2008) with classical order 2 variations convolution filters.

To the best of our knowledge, nothing have been done, to have a complete estimation procedure that could estimate all Hurst exponent, diffusion coefficient and drift parameter with central limit theorems and this is the gap we fill in this paper. Moreover, estimates of $H$, $\sigma$ and $\lambda$ presented in this paper slightly differ from all those studied previously.

In Sect. 2 we review the basic facts of stochastic differential equations driven by the fBM and we introduce the basic notations and assumptions. Sect. 3 presents consistent and asymptotically Gaussian estimators of the parameters of the fractional Ornstein–Uhlenbeck process from discrete time observations. In Sect. 4 we present ready-to-use software for the R statistical environment which allows the user to simulate and esti-
mate the parameters of the fOU process. We further present Monte-Carlo experiments to test the performance of the estimators under different sampling conditions.

2 Model specification

Let \( X = (Y_t, t \geq 0) \) be a fractional Ornstein–Uhlenbeck process (fOU), i.e. the solution of

\[
Y_t = y_0 - \lambda \int_0^t Y_s ds + \sigma W^H_t, \quad t \geq 0,
\]

where unknown parameter \( \vartheta = (\lambda, \sigma, H) \) belongs to an open subset \( \Theta \) of \((0, \Lambda) \times [\sigma, \overline{\sigma}] \times (0, 1), 0 < \Lambda < +\infty, 0 < \sigma < \overline{\sigma} < +\infty \) and \( W^H = (W^H_t, t \geq 0) \) is a standard fBM (Kolmogorov 1940; Mandelbrot 1968) of Hurst parameter \( H \in (0, 1) \), i.e. a Gaussian centered process of covariance function

\[
E W^H_t W^H_s = \frac{1}{2} \left( t^{2H} + s^{2H} - |t - s|^{2H} \right).
\]

It is worth emphasizing that in the case \( H = \frac{1}{2} \), \( W^\frac{1}{2} \) is the classical Wiener process. The fOU process is neither Markovian nor a semimartingale for \( H \neq \frac{1}{2} \) but remains Gaussian and ergodic. For \( H > \frac{1}{2} \), it even presents the long-range dependance property (see Cheridito et al. 2003).

The present work exposes an estimation procedure for estimating all three components of \( \vartheta \) given the regular discretization of the sample path \( Y^T = (Y_t, 0 \leq t \leq T) \), precisely

\[
\left( X_n := Y_{n \Delta_N}, n = 0 \ldots N \right),
\]

where \( T = T_N = N \Delta_N \rightarrow +\infty \) and \( \Delta_N \rightarrow 0 \) as \( N \rightarrow +\infty \).

In the following, convergences \( \xrightarrow{L}, \xrightarrow{p} \) and \( \xrightarrow{a.s.} \) stand respectively for the convergence in law, the convergence in probability and the almost-sure convergence.

3 Estimation procedure

Contrary to the previous works on the subject, we consider here the problem of estimation of \( H, \sigma \) and \( \lambda \) when all parameters are unknown, using discrete observations from the fractional Ornstein–Uhlenbeck process. Due to the fact that one can estimate \( H \) and \( \sigma \) without the knowledge of \( \lambda \), our approach consists naturally in a two step procedure.
3.1 Estimation of the Hurst exponent $H$ and the diffusion coefficient $\sigma$
with quadratic generalized variations

The key point of this paper is that the Hurst exponent $H$ and the diffusion coefficient $\sigma$ can be estimated without estimating $\lambda$.

Let $a = (a_0, \ldots, a_K)$ be a discrete filter of length $K + 1$, $K \in \mathbb{N}$, and of order $L \geq 1$, $K \geq L$, i.e.

$$\sum_{k=0}^{K} a_k k^\ell = 0 \quad \text{for} \quad 0 \leq \ell \leq L - 1 \quad \text{and} \quad \sum_{k=0}^{K} a_k k^L \neq 0. \quad (2)$$

Let it be normalized with

$$\sum_{k=0}^{K} (-1)^{1-k} a_k = 1. \quad (3)$$

In the following, we will also consider dilated filter $a^2$ associated to $a$ defined by

$$a^2_k = \begin{cases} a_{k'} & \text{if} \quad k = 2k' \\ 0 & \text{otherwise} \end{cases} \quad \text{for} \quad 0 \leq k \leq 2K.$$ 

Since $\sum_{k=0}^{2K} a^2_k k^r = 2^r \sum_{k=0}^{K} k^r a_k$, filter $a^2$ as the same order than $a$. We denote by

$$V_{N,a} = \sum_{i=0}^{N-K} \left( \sum_{k=0}^{K} a_k X_{i+k} \right)^2$$

the generalized quadratic variations associated to the filter $a$ (see for instance Istas and Lang 1997). Let us denote

$$\rho^{a^m,a^n}_H (i) = \frac{\sum_{k=0}^{mK} \sum_{\ell=0}^{nK} a^m_k a^n_\ell |mk - n\ell + i|^2H}{(mn)^H \sum_{k,\ell} a_k a_\ell |k - \ell|^2H}.$$ 

Finally,

$$\hat{H}_N = \frac{1}{2} \log_2 \frac{V_{N,a^2}}{V_{N,a}} \quad (4)$$

and

$$\hat{\sigma}_N = \left( -2 \frac{V_{N,a}}{\sum_{k,\ell} a_k a_\ell |k - \ell|^2 \hat{H}_N \Delta^2 N} \right)^{1/2}. \quad (5)$$
Theorem 1 Let \( a \) be a filter of order \( L \geq 2 \). Then, both estimators \( \hat{H}_N \) and \( \hat{\sigma}_N \) are strongly consistent, i.e.

\[
(\hat{H}_N, \hat{\sigma}_N) \xrightarrow{a.s.} (H, \sigma) \quad \text{as} \quad N \rightarrow +\infty.
\]

Moreover, we have asymptotical normality property, i.e. as \( N \rightarrow +\infty \), for all \( H \in (0, 1) \),

\[
\sqrt{N}(\hat{H}_N - H) \xrightarrow{L} \mathcal{N}(0, \Gamma_1(\vartheta, a))
\]

and

\[
\frac{\sqrt{N}}{\log N}(\hat{\sigma}_N - \sigma) \xrightarrow{L} \mathcal{N}(0, \Gamma_2(\vartheta, a)).
\]

Here \( \Gamma_1(\vartheta, a) \) and \( \Gamma_2(\vartheta, a) \) are symmetric definite positive matrices depending on \( \sigma, H \) and the filter \( a \) and defined by

\[
\Gamma_1(\vartheta, a) = \frac{1}{2\log(2)^2} \sum_{i \in \mathbb{Z}} \left( \rho_H a^2(i)^2 + \rho_H a^2(i)^2 - 2\rho_H a(i)^2 \right) \quad (6)
\]

and \( \Gamma_2(\vartheta, a) = \frac{\sigma^2}{4} \Gamma_1(\vartheta, a) \).

Proof The solution of (1) can be explicited

\[
Y_t = x_0 e^{-\lambda t} + \sigma \int_0^t e^{-\lambda(t-s)} dW^H_s,
\]

where the integral is defined as a Riemann-Stieltjes pathwise integral. Let us consider the stationary centered Gaussian solution

\[
Y^*_t = \sigma \int_{-\infty}^t e^{-\lambda(t-u)} dW^H_u.
\]

We have that

\[
Y^*_t = Y_t + e^{-\lambda t} \left( Y^*_0 - y_0 \right)
\]

which allows us to use only the stationary version in the following. As \( t \) is large, this is approximatively the same process, i.e.

\[
Y^*_t - Y_t = e^{-\lambda t} \left( Y^*_0 - y_0 \right) \xrightarrow{a.s.} 0 .
\]
At small values of $t$, the process $(Y_t, t \geq 0)$ can be also decomposed into the sum of its stationary version and a $C^\infty$ function $f$:

$$Y_t = Y_t^\dagger + e^{-\lambda t} (Y_0 - Y_0^\dagger).$$

It is worth emphasizing that the differentiable part $f$ has no influence on the behavior of the variogram (see Eq. (7)) that will be computed. In other words, the variogram of $Y_t$ has the same behavior at small values of $t$ as the variogram of $Y_t^\dagger$.

Let $v(t)$ denote the variogram of $Y_t^\dagger$. We now show that

$$v(t) = \mathbb{E} \left( Y_t^\dagger \right)^2 - \mathbb{E} Y_t^\dagger Y_0^\dagger = \frac{\sigma^2}{2} |t|^{2H} + r(t)$$  (7)

where $r(t) = o(|t|^{2H})$ as $t$ tends to zero. Indeed, it is known (see Cheridito et al. 2003, Lemma 2.1) that

$$\mathbb{E} Y_0^\dagger (Y_0^\dagger - Y_t^\dagger) = -\sigma^2 H (2H - 1) e^{-\lambda t} \int_{-\infty}^{0} e^{\lambda u} \left( \int_0^t e^{-\lambda (t-v)} (v-u)^{2H-2} dv \right) du.$$
Thus,

\[
\frac{dv}{dt}(t) = -\frac{1}{2} \sigma^2 H (2H - 1) \int_0^\infty e^{-\lambda w} \left( (t + w)^{2H-2} - |t - w|^{2H-2} \right) dw
\]

\[
= -\frac{1}{2} \sigma^2 H (2H - 1) t^{2H-1} \int_0^\infty e^{-\lambda t y} \left( (1 + y)^{2H-2} - |1 - y|^{2H-2} \right) dy
\]

\[
= -\frac{1}{2} \sigma^2 H (2H - 1) t^{2H-1} \int_0^\infty \left( (1 + y)^{2H-2} - |1 - y|^{2H-2} \right) dy + \tilde{r}(t)
\]

\[
= \sigma^2 H t^{2H-1} + \tilde{r}(t)
\]

with

\[
\tilde{r}(t) = -\frac{1}{2} \sigma^2 H (2H - 1) t^{2H} \sum_{i=0}^\infty c_i t^i
\]

where

\[
c_i = \int_0^\infty \frac{(-\lambda y)^i}{i!} \left( (1 + y)^{2H-2} - |1 - y|^{2H-2} \right) dy.
\]

Therefore, we proved that

\[
v(t) = \frac{\sigma^2}{2} |t|^{2H} + r(t).
\]

Now, applying results in Istas and Lang (1997, Theorem 3(i,iii)), the proof of Theorem 1 is complete because the following conditions are fulfilled:

– firstly, \( r(t) = o(|t|^{2H}) \) as \( t \) tends to zero,
– secondly, for classical generalized quadratic variations or order \( L \geq 2 \) (for instance \( L = 2 \)),

\[
|r^{(4)}(t)| \leq G |t|^{2H+1-\varepsilon-4}, \quad 0 < t \leq 1,
\]

with \( 2H + 1 - \varepsilon > 2H \) and \( 4 > 2H + 1 - \varepsilon + 1/2 \) for \( 0 < \varepsilon < 1 \) and any \( H \in (0, 1) \).

The values of \( \Gamma_1(\theta, \alpha) \) and \( \Gamma_2(\theta, \alpha) \) can be found in Cœurjolly (2001, Proposition 4).
Remark 1 We have two useful examples of filters. Classical filters of order $L \geq 1$ are defined by
\[ a_k = c_{L,k} = \frac{(-1)^{1-k}}{2^k} \binom{K}{k} = \frac{(-1)^{1-k}}{2^k \frac{K!}{k!(K-k)!}} \quad \text{pour} \quad 0 \leq k \leq K. \quad (9) \]

Daubechies filters of even order can also be considered (see Daubechies 1992), for instance the order 2 Daubechies’ filter:
\[ \frac{1}{\sqrt{2}} (.4829629131445341, -.8365163037378077, .2241438680420134, .1294095225512603). \quad (10) \]

Remark 2 For classical order 1 quadratic variations ($L = 1$) and $a = (\frac{-1}{2}, \frac{1}{2})$ we can also obtain consistency for any value of $H$, but the central limit theorem holds only for $H < \frac{3}{4}$ (see Istas and Lang 1997).

3.2 Estimation of the drift parameter $\lambda$ when both $H$ and $\sigma$ are unknown

From Hu and Nualart (2010), we know the following result
\[ \lim_{t \to \infty} \text{var}(Y_t) = \lim_{t \to \infty} \frac{1}{t} \int_0^t Y_t^2 dt = \frac{\sigma^2 \Gamma (2H + 1)}{2\lambda^2 H} =: \mu_2. \]

This gives a natural plug-in estimator of $\lambda$, namely
\[ \hat{\lambda}_N = \left( \frac{2 \hat{\mu}_{2,N}}{\hat{\sigma}_{N}^2 \Gamma (2\hat{H}_N + 1)} \right)^{-\frac{1}{2\hat{H}_N}} \quad (11) \]
where $\hat{\mu}_{2,N}$ is the empirical moment of order 2, i.e
\[ \hat{\mu}_{2,N} = \frac{1}{N} \sum_{n=1}^{N} X_n^2. \]

Theorem 2 Let $H \in \left( \frac{1}{2}, \frac{3}{4} \right)$ and a mesh satisfying the condition $N \Delta_N^p \to 0$, $p > 1$, and $\Delta_N (\log N)^2 \to 0$ as $N \to +\infty$. Then, as $N \to +\infty$,
\[ \hat{\lambda}_N \overset{a.s.}{\to} \lambda \]
and
\[ \sqrt{T_N} \left( \hat{\lambda}_N - \lambda \right) \overset{L}{\to} N(0, \Gamma_3(\vartheta)), \]
where $\Gamma_3(\vartheta) = \lambda \left(\frac{\sigma_H}{2H}\right)^2$ and

$$\sigma_H^2 = (4H - 1) \left( 1 + \frac{\Gamma(1 - 4H)\Gamma(4H - 1)}{\Gamma(2 - 2H)\Gamma(2H)} \right).$$

(12)

Proof Let us set $T_N = N\Delta N$. It had been shown in Hu and Nualart (2010) that, as $T_N \to +\infty$ (or as $N \to +\infty$),

$$\frac{1}{T_N} \int_0^{T_N} Y_t^2 \, dt \xrightarrow{a.s.} \kappa_H \lambda^{-2H}$$

(13)

and

$$\sqrt{T_N} \left( \frac{1}{T_N} \int_0^{T_N} Y_t^2 \, dt - \kappa_H \lambda^{-2H} \right) \xrightarrow{L} \mathcal{N}(0, (\sigma_H\kappa_H)^2 \lambda^{-4H-1})$$

(14)

where $\kappa_H = \frac{\sigma^2 \Gamma(2H+1)}{2}$ and $\sigma_H$ is defined by (12). Let us denote $\hat{\mu}_{2,N}$ the discretization of the integral

$$\hat{\mu}_{2,N} = \frac{1}{N} \sum_{n=1}^{N} X_n^2 \quad \text{and} \quad \mu_2 = \kappa_H \lambda^{-2H}.
$$

First, let us write

$$\hat{\mu}_{2,N} - \mu_2 = \left( \hat{\mu}_{2,N} - \frac{1}{T_N} \int_0^{T_N} Y_t^2 \, dt \right) + \left( \frac{1}{T_N} \int_0^{T_N} Y_t^2 \, dt - \mu_2 \right).$$

The first term is the discretization error and the second term is the “ergodic theorem” term. The second term is studied in Eq. (13). As $(Y_t, t \geq 0)$ is a Gaussian process and Hölder regular of order $\frac{1}{2} < H < \frac{3}{4}$, we have that the discretization error will tend to zero almost surely (Riemann–Stieljes path wise integral). Namely,

$$A_N = \frac{1}{T_N} \int_0^{T_N} Y_t^2 \, dt - \frac{1}{N} \sum_{n=1}^{N} X_n^2$$

$$= \frac{1}{N\Delta N} \sum_{n=1}^{N} \int_{n-1\Delta N}^{n\Delta N} (Y_t^2 - Y_{n\Delta N}^2) \, dt$$

$$= \frac{1}{N\Delta N} \sum_{n=1}^{N} \int_{n-1\Delta N}^{n\Delta N} (Y_t - Y_{n\Delta N}) (Y_t + Y_{n\Delta N}) \, dt.$$
Consequently,

\[ |A_N| \leq \frac{2}{N \Delta N} \sup_{t \geq 0} |Y_t| \sum_{n=1}^{N} \int_{(n-1)\Delta N}^{n\Delta N} |Y_t - Y_{n\Delta N}| \, dt. \]

As \((Y_t, t \geq 0)\) is Hölder continuous of order \(H - \varepsilon\) for all \(\varepsilon > 0\) (see for instance Kaarakka and Salminem 2011), we get that almost surely

\[ |Y_t - Y_{n\Delta N}| \leq C |t - n \Delta N|^{H-\varepsilon} = C \Delta_N^{H-\varepsilon} \text{ on } [(n-1)\Delta_N, n\Delta_N]. \]

Then

\[ |A_N| \leq 2C \Delta_N^{H-\varepsilon} \sup_{t \geq 0} |Y_t| \text{ a.s..} \]

Since the process \((Y_t, t \geq 0)\) is Gaussian and has ergodic properties when \(\lambda > 0\) (see Cheridito et al. 2003), we get the following results:

\[ A_N \xrightarrow{a.s.} 0 \text{ if } \Delta_N \xrightarrow{} 0 \]

and

\[ \sqrt{T_N} A_N \xrightarrow{p} 0 \text{ if } N \Delta_N^p \xrightarrow{} 0 \text{ } \forall \, p \in \mathbb{N}. \]

As \(N \xrightarrow{} +\infty\), we deduce

\[ \left( \hat{\mu}_{2,N} - \frac{1}{T_N} \int_{0}^{T_N} Y_t^2 \, dt \right) \xrightarrow{a.s.} 0 \]

and, using (13), we get

\[ \hat{\mu}_{2,N} \xrightarrow{a.s.} \mu_2. \quad (15) \]

In order to get the asymptotic normality property, we write

\[ \sqrt{T_N} (\hat{\mu}_{2,N} - \mu_2) = \sqrt{T_N} \left( \hat{\mu}_{2,N} - \frac{1}{T_N} \int_{0}^{T_N} Y_t^2 \, dt \right) + \sqrt{T_N} \left( \frac{1}{T_N} \int_{0}^{T_N} Y_t^2 \, dt - \mu_2 \right). \]
As we deduce also that

\[ \sqrt{T_N} \left( \hat{\mu}_{2,N} - \frac{1}{T_N} \int_0^{T_N} Y_t^2 \, dt \right) \xrightarrow{p} 0, \]

we can show, using (14), that

\[ \sqrt{T_N} \left( \hat{\mu}_{2,N} - \mu_2 \right) \xrightarrow{L} \mathcal{N} \left( 0, (\sigma H \kappa H)^2 \lambda^{-4H-1} \right). \]  

(16)

Let us introduce the following two quantities

\[ M_N = \begin{pmatrix} \hat{\mu}_{2,N} \\ \hat{H}_N \\ \hat{\sigma}_N \end{pmatrix} \quad \text{and} \quad m = \begin{pmatrix} \mu_2 \\ H \\ \sigma \end{pmatrix}. \]

Finally, results obtained in Theorem 1 and the convergence in (16) gives consistency of \( M_N \), i.e. \( M_N \xrightarrow{a.s.} m \) as \( N \xrightarrow{} +\infty \). Let us further define

\[ g(\mu_2, H, \sigma) = \left( \frac{2\mu_2}{\sigma^2 \Gamma(2H+1)} \right)^{-\frac{1}{2}} \]

The derivatives of \( g \) with respect to \( \sigma, H \) and \( \mu_2 \) are bounded when \( 0 < \Lambda < +\infty, 0 < \overline{\sigma} < +\infty \) and \( \frac{1}{2} < H < \frac{3}{4} \). Therefore, as \( \Delta_N (\log N)^2 \xrightarrow{} 0 \) as \( N \xrightarrow{} +\infty \), we can obtain by Taylor expansion that

\[ \sqrt{T_N} \left( g(M_N) - g(m) \right) \xrightarrow{L} \mathcal{N} \left( 0, g'_\mu_2 (m)^2 (\sigma H \kappa H)^2 \lambda^{-4H-1} \right) \]

or

\[ \sqrt{T_N} \left( \hat{\lambda}_N - \lambda \right) \xrightarrow{L} \mathcal{N} \left( 0, \Gamma_3(\vartheta) \right) \]

where \( \Gamma_3(\vartheta) = g'_\mu_2 (m)^2 (\sigma H \kappa H)^2 \lambda^{-4H-1} = \lambda \left( \frac{\sigma \mu H}{2} \right)^2, g'_\mu_2 (.) \) is the derivative of \( g \) with respect to \( \mu_2 \). Moreover, continuous mapping theorem gives

\[ \hat{\lambda}_N \xrightarrow{a.s.} \lambda \]

as \( N \xrightarrow{} +\infty \).

**Remark 3** The different conditions on \( \Delta_N \) raise the question of whether such a rate actually exists. One possible mesh is \( \Delta_N = \frac{\log N}{N} \).

**Remark 4** As in the classical case \( H = \frac{1}{2} \), the limit variance \( \Gamma_3(\vartheta) \) does not depend on the diffusion coefficient \( \sigma \). Let us also notice that the quantity \( \sigma_H^2 \) appearing in \( \Gamma_3(\vartheta) \) is an increasing function of \( H \).
4 Statistical software and Monte-Carlo analysis

In Sect. 4.1 we present a brief introduction to the yuima package for R statistical environment (R Development Core Team 2010) and its usefulness in simulating and estimating also fractional diffusion processes. Later, in Sect. 4.2, in order to test the performance of the estimators proposed in Sect. 3 under finite samples conditions, we run a Monte-Carlo analysis. We consider different setups for the filter and the parameters even outside the region $\frac{1}{2} < H < \frac{3}{4}$ and different sample size with large and small values of $T$ in order to test the performance of the estimator of the drift parameter when the stationarity is not reached by the process. All numerical experiments presented in the following have been done with the yuima package (Yuima Project Team 2011).

4.1 Overview of the yuima package

The yuima package is a comprehensive framework, based on the S4 system of classes and methods, which allows for the description of solutions of stochastic differential equations. Although we can only give few details here, the user can specify a stochastic differential equation of the form

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t^H + c(t, X_t)Z_t$$

where the coefficients $b(\cdot, \cdot), \sigma(\cdot, \cdot)$ and $c(\cdot, \cdot)$ are entirely specified by the user, even in parametric form. Here $(Z_t, t \geq 0)$ is a Lévy process (for more information on Lévy processes, see Bertoin (1998); Sato (1999) and $(W_t^H, t \geq 0)$ is a fBM (recall that $(W_{\frac{1}{2}}^1, t \geq 0)$ is the standard Brownian motion). The Lévy process $(Z_t, t \geq 0)$ and the fBM $(W_t^H, t \geq 0)$ can be present at the same time only when $H = \frac{1}{2}$, but all other combinations are possible.

4.1.1 Diffusion processes

Assume that we want to describe the following stochastic differential equation

$$dX_t = -3X_t dt + \frac{1}{1 + X_t^2}dW_t$$

This is done in yuima specifying the drift and diffusion coefficients as plain mathematical expressions

```r
R> mod1 <- setModel(drift = "-3*x", diffusion = "1/(1+x^2)")
```

At this point, the package fills the proper slots of the yuima object

```r
R> str(mod1)
Formal class ‘yuima.model’ [package ‘yuima’] with 16 slots
 ..@ drift : expression((-3 * x))
 ..@ diffusion : List of 1
```
For the model specification, it is possible to see that the jump coefficient is void and the Hurst parameter is set to 0.5, because this corresponds to the standard Brownian motion. Now, with \texttt{mod1} in hands, it is very easy to simulate a trajectory of the process and eventually plot it on screen as follows.

\begin{verbatim}
R> set.seed(123)
R> X <- simulate(mod1)
R> plot(X)
\end{verbatim}

4.1.2 Specification of parametric models

When a parametric model like

\[ dX_t = -\theta X_t \, dt + \frac{1}{1 + X_t^\gamma} \, dW_t \]

is specified, \textit{yuima} attempts to distinguish the parameters’ names from the ones of the state and time variables.

\begin{verbatim}
R> mod2 <- setModel(drift = "-theta*x", diffusion = "1/(1+x^gamma)"")
R> str(mod2)
\end{verbatim}

Next code shows how parametric models are handled internally

\begin{verbatim}
Formal class ‘yuima.model’ [package ‘yuima’] with 16 slots
 .. ..@ drift : expression((-theta * x))
 .. ..@ diffusion : List of 1
 .. ..$ : expression(1/(1 + x^gamma))
 ..@ hurst : num 0.5
 ..@ jump.coeff : expression()
\end{verbatim}
In order to simulate the parametric model it is necessary to specify the values of the parameters as the next code shows how to

R> set.seed(123)
R> X <- simulate(mod2, true.param = list(theta = 1, gamma = 3))
R> plot(X)

The yuima package provides the user, not only the simulation part, but also several parametric and non-parametric estimation procedures. In the next section we limit the exposition to a single example which is of specific interest to this paper. We show how to use yuima for simulation and estimation of the fractional Ornstein–Uhlenbeck process considered in this paper.

4.1.3 Example of numerical simulation and estimation of the fOU process

With the yuima package the fractional Gaussian noise is simulated with the Wood and Chan method (Wood and Chan 1994) or other techniques. We present below how to simulate one sample path of the fractional Ornstein–Uhlenbeck process with Euler–Maruyama method (Neuenkirch and Nourdin 2007). For instance, loading the package with

library(yuima)

we can simulate a regularly sampled path of the following model

\[ X_t = 1 - 2 \int_0^t X_t dt + dW_t^H, \quad H = 0.7, \]

with

samp <- setSampling(Terminal=100, n=10000)
mod <- setModel(drift=''-2*x'', diffusion='''1''', hurst=0.7)
Parameter estimation for the discretely observed fOU process

Table 1  Mean average (and standard deviation in parenthesis) of 500 Monte-Carlo simulations for the estimation of $H$ (left) and $\sigma$ (right) for different cases

| $\hat{H}$ | $H = 0.5$ | $H = 0.7$ | $H = 0.9$ | $\sigma = 1$ | $H = 0.5$ | $H = 0.7$ | $H = 0.9$ |
|-----------|-----------|-----------|-----------|--------------|-----------|-----------|-----------|
| $\sigma = 1$ | 0.499     | 0.697     | 0.898     | 1.024        | 1.016     | 1.081     |
|            | (0.035)   | (0.033)   | (0.031)   | (0.262)      | (0.282)   | (0.437)   |
| $\sigma = 2$ | 0.498     | 0.700     | 0.898     | 2.035        | 2.073     | 2.213     |
|            | (0.033)   | (0.034)   | (0.033)   | (0.510)      | (0.564)   | (1.110)   |

$T_N = 100$, $N = 1,000$ and $\lambda = 2$

```
ou <- setYuima(model=mod, sampling=samp)
fou <- simulate(ou, xinit=1)
```

The estimation procedure of the Hurst parameter have been implemented in `qgv` function. In order to estimate only the parameter $H$, one can use

```
qgv(fou)
```

that works also for non linear fractional diffusions (see Melichov 2011). The procedure for joint estimation of the Hurst exponent $H$, diffusion coefficient $\sigma$ and drift parameter $\lambda$ is called `lse(,frac=TRUE)`. So for example, in order to estimate the three different parameters $H$, $\lambda$ and $\sigma$, one can use

```
lse(fou,frac=TRUE)
```

which uses by default the order 2 Daubechies’ filter (see Remark 1) if the user does not specify the `filter` argument.

4.2 Performance of the Hurst parameter and diffusion coefficient estimation

4.2.1 For different setups of the parameters

In this first simulation part, we present mean average values and standard deviation values for both estimators $\hat{H}_N$ and $\hat{\sigma}_N$ (see Sect. 3.1 and the Eqs. (4) and (5) for the respective definitions) with 500 Monte-Carlo replications. This have been done for different Hurst exponents $H$ and different diffusion coefficients $\sigma$ in the model (1), the parameter $\lambda$ being fixed equal to 2. The results are presented in Tables 1 and 2 for different values of the horizon time $T_N$ and the sample size $N$.

Contrary to the estimation of the drift (see Sect. 4.3), we have consistent estimates of $H$ and $\sigma$ for any values of $T_N$. Only the size of the sample $N$ have influence on the performance of the estimate.

In order to illustrate the asymptotical normality for the estimators $\hat{H}_N$ and $\hat{\sigma}_N$ respectively of $H$ and $\sigma$, we present in Fig. 1 the kernel estimation of the density.

4.2.2 For different setups of the filter

The function `qgv` estimates with quadratic generalized variations method both parameter $H$ and $\sigma$. It uses by default the order 2 Daubechies filter (see Remark 1) if the user does not specify the `filter` argument.
Table 2 Mean average (and standard deviation in parenthesis) of 500 Monte-Carlo simulations for the estimation of \( H \) (left) and \( \sigma \) (right) for different cases, and for \( T_N = 100, N = 100,000 \) and \( \lambda = 2 \)

| \( H \) | \( \hat{H} \) | \( \hat{\sigma} \) | \( \hat{\sigma}_N \) |
|---|---|---|---|
| \( \sigma = 1 \) | 0.500 (0.003) | 1.000 (0.025) | 0.003 (0.003) |
| \( \sigma = 2 \) | 0.500 (0.004) | 2.001 (0.053) | 0.004 (0.004) |

**Fig. 1** Kernel estimation for the density of \( \left( \sqrt{N} \left( \hat{H}_{N}^{(m)} - H \right) \right)_{m=1...M} \) (on the left) and \( \left( \sqrt{\log N} \left( \hat{\sigma}_{N}^{(m)} - \sigma \right) \right)_{m=1...M} \) (on the right), \( M = 500 \), for \( T_N = 100 \) and \( N = 100,000 \) (fill line) and the fitted Gaussian density (dashed line) for \( \vartheta = (\lambda, \sigma, H) = (2, 2, 0.7) \).

We present in this section the comparison study of the asymptotical variance \( \Gamma_1(\vartheta, a) \) and \( \Gamma_2(\vartheta, a) \) of both estimators \( \hat{H}_N \) and \( \hat{\sigma}_N \) respectively in Theorem 1 for different filters \( a \), the parameters \( \vartheta = (\lambda, \sigma, H) \) being fixed. It is worth emphasizing that this study on the filters for the fractional Ornstein–Uhlenbeck process can be compared to the one initiated on the fBM (Cœurjolly 2001).

In the following, \( c_1, c_2 \) are the classical filters of order 1 and 2 respectively (see (9)). We denote by \( a_{Db4} \) the order 2 Daubechies filter presented in Remark 1.

In the sense of limit variance, we can see on Fig. 2 that the estimator \( \hat{H}_N \) based on the order 2 Daubechies’ filter \( a_{Db4} \) is better than the estimator based on the classical order 2 filter \( c_2 \) for any \( H \in (0, 1) \). Moreover, the estimator \( \hat{H}_N \) based on classical order 1 filter \( c_1 \) is better than the two others second order filters for \( 0 < H < \frac{3}{4} \) where the Theorem 1 is still valid (see Remark 2).

The typical trend of the fOU process (see Table 3) vanishes with the uses of the second order filters estimators. In this sense, the bias is much bigger with order 1 classical filter than for the two other second order filters at least when the parameter \( H \) increases to \( \frac{3}{4} \) as it were already noticed in Cœurjolly (2001), Istas and Lang (1997).

4.3 Plug-in for the estimation of drift parameter \( \lambda \)

In this second simulation part, we present mean average values and standard deviation values for the estimator \( \hat{\lambda}_N \) (see Sect. 3.2 for the definition) of the drift with 500 Monte-Carlo replications. This have been done for different values of \( \lambda \) and \( H \) in model (1), the diffusion coefficient \( \sigma \) being fixed to 1 (see Remark 4). The results are presented in Table 4 for different values of the horizon time \( T_N \) and the sample size \( N \).
Fig. 2  Computation of $\Gamma_1(\vartheta, a)$ for different values of $H$ depending on the filter used in $\hat{H}_N$

Table 3  Mean average (and standard deviation in parenthesis) of 500 Monte-Carlo simulations for the estimation of $H$ for different type of filters and different values of $H$

| $H$  | $a = c_1$ | $a = c_2$ | $a = a_{Db4}$ | $H$  | $a = c_1$ | $a = c_2$ | $a = a_{Db4}$ |
|------|-----------|-----------|---------------|------|-----------|-----------|---------------|
| $H = 0.1$ | 0.098     | 0.099     | 0.100         | $H = 0.1$ | 0.100     | 0.099     | 0.100         |
|         | (0.010)   | (0.015)   | (0.012)       |       | (0.010)   | (0.015)   | (0.012)       |
| $H = 0.7$ | 0.676     | 0.699     | 0.699         | $H = 0.7$ | 0.699     | 0.701     | 0.699         |
|         | (0.005)   | (0.012)   | (0.010)       |       | (0.006)   | (0.012)   | (0.010)       |

This have been computed with $T_N = 10$, $N = 10,000$ and $\sigma = 2$ for $\lambda = 2$ (fOU) on the left and $\lambda = 0$ (fBm) on the right.

Table 4  Mean average (and standard deviation in parenthesis) of 500 Monte-Carlo simulation for the estimation of $\lambda$ for different values of $H$ and $\lambda$

| $H = 0.5$ | $H = 0.6$ | $H = 0.7$ | $H = 0.5$ | $H = 0.6$ | $H = 0.7$ |
|-----------|-----------|-----------|-----------|-----------|-----------|
| $\lambda = 0.5$ | 0.093     | 0.214     | 0.353     | 0.476     | 0.514     | 0.605     |
|           | (0.037)   | (0.057)   | (0.069)   | (0.148)   | (0.166)   | (0.298)   |
| $\lambda = 1$ | 0.138     | 0.276     | 0.432     | 0.906     | 0.940     | 1.005     |
|           | (0.052)   | (0.068)   | (0.078)   | (0.227)   | (0.238)   | (0.412)   |

Here $\sigma = 1$ and $T_N = 1$ and $N = 100,000$ (left) and $T_N = 100$ and $N = 1,000$ (right).

We can see in Table 4 that the values of $T_N$ is important for the estimation of the drift. Actually, the consistency of the estimates are valid for increasing values of $T_N$ and decreasing values of the mesh size $\Delta_N$. Moreover, the bigger $H$, the harder the estimation of the drift parameter. This phenomena can be explained by the long-range dependence property of the fOU process. It is the same for $\lambda$; as $\lambda$ increases, its estimation is harder (see Remark 4). It can be explained by the fact that when $\lambda$ is bigger, the fOU process enters faster in its stationary behavior where it is more difficult to detect the trend.
Finally, in order to illustrate the asymptotical normality for the estimator $\hat{\lambda}_N$ of $\lambda$, we present in Fig. 3 the kernel estimation of the density.

**Acknowledgments** We would like to thank Marina Kleptsyna for the discussions and her interest for this work. Computing resources have been financed by Mostapad project in CNRS FR 2962. This work has been supported by the project PRIN 2009JW2STY, Ministero dell’Istruzione dell’Università e della Ricerca.

**References**

Bercu B, Coutin L, Savy N (2010) Sharp large deviations for the fractional Ornstein-Uhlenbeck process. Teoriya Veroyatnostei i ee Primeneniya
Bertin K, Torres S, Tudor C (2011) Drift parameter estimation in fractional diffusions driven by perturbed random walk. Stat Probab Lett 81(2):243–249
Bertoin J (1998) Lévy processes. Cambridge University Press, Cambridge
Berzin C, Leon J (2008) Estimation in models driven by fractional Brownian motion. Annales de l’Institut Henri Poincaré 44(2):191–213
Bregni S, Erangoli W (2005) Fractional noise in experimental measurements of IP traffic in a metropolitan area network. Proc IEEE GlobeCom 2005(2):781–785
Brouste A, Kleptsyna M (2010) Asymptotic properties of MLE for partially observed fractional diffusion system. Stat Inference Stoch Process 13(1):1–13
Cheridito P, Kawaguchi H, Maejima M (2003) Fractional Ornstein–Uhlenbeck processes. Electron J Probab 8(3):1–14
Cialenco I, Lototsky S, Pospisil J (2009) Asymptotic properties of the maximum likelihood estimator for stochastic parabolic equations with additive fractional Brownian motion. Stoch Dyn 9(2):169–185
Cérou JF (2001) Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths. Stat Inference Stoch Process 4(2):199–227
Daubechies I (1992) Ten lectures on wavelets. SIAM, Philadelphia
Florens-Zmirou D (1989) Approximate discrete time schemes for statistics of diffusion processes. Statistics 20:263–284
Genon-Catalot V (1990) Maximum constraint for diffusion processes from discrete observation. Statistics 21:99–116
Gobet E (2002) LAN property for ergodic diffusions with discrete observations. Annales de l’Institut Henri Poincaré 38(5):711–737
Hu Y, Nualart D (2010) Parameter estimation for fractional Ornstein–Uhlenbeck processes. Stat Probab Lett 80(11–12):1030–1038
Hu Y, Nualart D, Xiao W, Zhang W (2011) Exact maximum likelihood estimator for drift fractional brownian motion at discrete time. Acta Mathematica Scientia 31(5):1851–1859
Istas J, Lang G (1997) Quadratic variations and estimation of the local Hölder index of a Gaussian process. Annales de l’Institut Henri Poincaré 23(4):407–436
Jacod J (2001) Inference for stochastic processes. Statistics 683
Kaarakka T, Salminem P (2011) On fractional Ornstein–Uhlenbeck processes. Commun Stoch Anal 5(1):121–133
Kessler M (1997) Estimation of an ergodic diffusion from discrete observations. Scand J Stat 24:211–229
Kleptsyna M, Breton AL (2002) Statistical analysis of the fractional Ornstein–Uhlenbeck type process. Stat Inference Stoch Process 5:229–241
Kolmogorov A (1940) Winersche Spiralen und einige andere interessante Kurven in Hilbertschen Raum. Acad Sci USSR 26:115–118
Ludena C (2004) Minimum contrast estimation for fractional diffusion. Scand J Stat 31:613–628
Mandelbrot B, Ness JV (1968) Fractional Brownian motions, fractional noises and application. SIAM Rev 10:422–437
Melichov D (2011) On estimation of the Hurst index of solutions of stochastic equations. Ph.D. thesis, Vilnius Gediminas Technical University
Neuenkirch A, Nourdin I (2007) Exact rate of convergence of some approximation schemes associated to SDEs driven by a fractional Brownian motion. J Theor Prob 20(4):871–899
Neuenkirch A, Tindel S (2011) A least square-type procedure for parameter estimation in stochastic differential equations with additive fractional noise. preprint. http://hal.inria.fr/hal-00639030/PS/zero-squares-nl.ps
R Development Core Team: R (2010) A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org. ISBN 3-900051-07-0
Sato K (1999) Lévy processes and infinitely divisible distributions. Cambridge University Press, Cambridge
Willinger W, Taqqu M, Leland W, Wilson D (1995) Self-similarity in high-speed packet traffic: analysis and modeling of ethernet traffic measurements. Stat Sci 10:67–85
Wood A, Chan G (1994) Simulation of stationary Gaussian processes. J Comput Graph Stat 3(4):409–432
Xiao W, Zhang W, Xu W (2011) Parameter estimation for fractional Ornstein–Uhlenbeck processes at discrete observation. Appl Math Model 35:4196–4207
Yoshida N (1992) Estimation for diffusion processes from discrete observations. J Multivar Anal 41:220–242
Yuima Project Team: yuima (2011) The YUIMA project package (unstable version). http://R-Forge.R-project.org/projects/yuima/. R package version 0.1.1936