# On estimation for the fractional Ornstein-Uhlembeck process observed at discrete time

Stefano M. Iacus

Department of Economics, Management and Quantitative Methods
University of Milan
& R Core Team

joint work with A. Brouste (Univ. Le Mans, FR)

ISI Tokyo Satellite Meeting, Tokyo, 2-9-2013

# Fractional Brownian Motion

Let $W^H = \left(W_t^H, t \geq 0\right)$ be a normalized fractional Brownian motion (fBM), *i.e.* the zero mean Gaussian processes with covariance function

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

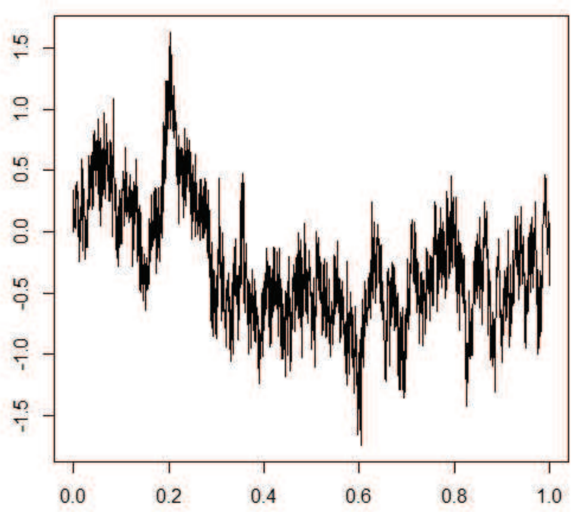
with Hurst exponent $H \in (0,1)$.

- the process is self-similar ($W_{at}^H \sim a^H W_t^H$)
- presents long range dependence (persistency - antipersistency)
- has dependent increments (apart for $H = \frac{1}{2}$).

For $H = \frac{1}{2}$, $W_t^H = W_t$ is a standard Brownian motion, i.e. independent increments.

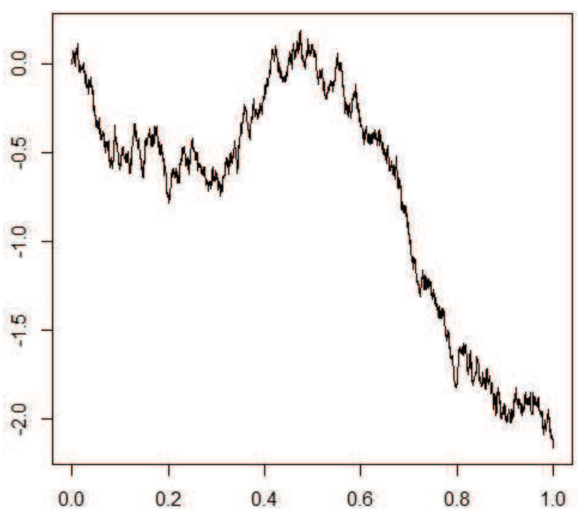
# Fractional Brownian Motion

$0 < H < \frac{1}{2}$
antipersistent



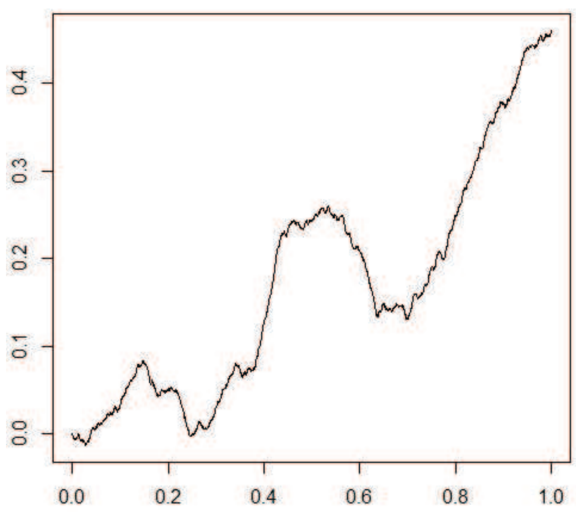

negative
correlated

$H = \frac{1}{2}$
independence




Brownian
motion

$\frac{1}{2} < H < 1$
persistent




positive
correlated

# fractional Ornstein-Uhlenbeck (fOU)

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

$$Y_t = y_0 - \lambda \int_0^t Y_s ds + \sigma W_t^H, \quad t > 0, \quad Y_0 = y_0, \tag{1}$$

where unknown parameter $\vartheta = (\lambda, \sigma, H)$ belongs to an open subset $\Theta$ of $(0, \Lambda) \times [\underline{\sigma}, \overline{\sigma}] \times (0, 1)$, $0 < \Lambda < +\infty$, $0 < \underline{\sigma} < \overline{\sigma} < +\infty$ and $W^H = (W_t^H, t \geq 0)$ is a standard fractional Brownian motion [10, 12] of Hurst parameter $H \in (0, 1)$,

The fOU process is not Markovian nor a semimartingale for $H \neq \frac{1}{2}$ but nevertheless Gaussian and ergodic. ([2])

We denote discrete observations of $Y_t$ by $X_j = Y_{t_j} = Y(t_j)$, where $0 = t_0 < t_1 < \cdots < t_N = T$ is a grid of deterministic times.

# Auxiliary known facts about fBm

Let $H \in (0,1)$, $t > s$, $\{W_t, t \in [0,T]\}$ a standard Brownian motion and

$$K^H(t,s) = c_H s^{\frac{1}{2}-H} \int_s^t (u-s)^{H-\frac{3}{2}} u^{H-\frac{1}{2}} du$$

with $c_H = \left( \frac{H(2H-1)}{Beta\left(2-2H, H-\frac{1}{2}\right)} \right)^{\frac{1}{2}}$. The fBM can be written as follows:

$$W_t^H = \int_0^t K^H(t,s) dW_s, \quad t \in [0,T]$$

The following is called *random walk approximation* to fBM, $t \in [0,T]$,

$$B_t^{H,N} = \sum_{i=1}^{[Nt]} \sqrt{N} \left\{ \int_{\frac{i-1}{N}}^{\frac{i}{N}} K^H\left(\frac{[Nt]}{N}, s\right) ds \right\} \xi_i \qquad \xi_i\text{'s i.i.d., } \mathbf{E}(\xi_i) = 0,\ \mathbf{Var}(\xi_i) = 1.$$

Then, as $N \to \infty$, $B_t^{H,N} \stackrel{w}{\to} W_t^H$ in Skorohod topology [14].

# fOU estimation: QMLE

Bertin *et al.* (2011) [1] considered the following statistical problem

$$dY_t = \lambda dt + dW_t^H$$

with $\lambda \in \mathbb{R}$ unknown and $H \in (\frac{1}{2}, 1)$ known.
Using random walk approximation & Euler scheme for the fOU, with $X_j = Y_{t_j}$, $t_j = j\Delta$, $j = 0, 1, \ldots, N$, $N\Delta = T$

$$X_{j+1} = X_j + a\Delta + \left( B_{t_{j+1}}^{H,N} - B_{t_j}^{H,N} \right)$$

the following QMLE estimator of $\lambda$

$$\hat{\lambda}_N = N \frac{\sum_{j=0}^{N^\alpha - 1} \frac{(1+\alpha_j)(X_{j+1} - X_j - h_j(X_1, \ldots, X_j))}{F_j^2}}{\sum_{j=0}^{N^\alpha - 1} \frac{(1+\alpha_j)^2}{F_j^2}}$$

where $\alpha_j$, $F_j$ and $h_j(\cdots)$ are explicit functions on the data.

# fOU estimation: true MLE

Bertin *et al.* (2011) proved that, under the asymptotic: $N \to \infty$, $T = N\Delta \to \infty$ and $\Delta = \frac{1}{N^\alpha}$ with $\alpha < 1$ the QMLE estimator $\hat{\lambda}_N$ is unbiased and consistent for $\lambda$ given the known $H \in (\frac{1}{2}, 1)$.

Let $S_N = \sum_{i=0}^{N-1} \left( X_{\frac{i+1}{N}} - X_{\frac{i}{N}} \right)^2$ and given that [15] $N^{2H-1} S_N \sim 1$, for large $N$ if $H$ is estimated from the data with

$$\hat{H}_N = 1 + \frac{\log S_N}{\log N}$$

by simulation results only it has been shown that the estimator $\hat{\lambda}_N$ is consistent and its variance is an increasing function of $H$.

# fOU estimation: true MLE

Hu *et al.* (2011) [6] considered the following statistical problem

$$dY_t = \lambda dt + \sigma dW_t^H$$

with $\lambda \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$ unknown and $H \in (0,1)$ known.

Let $\mathbf{t} = (\Delta, 2\Delta, \ldots, N\Delta)'$, $\mathbf{X} = (X_1, X_2, \ldots, X_N)'$, and

$$\Gamma_H = [\mathbf{Cov}(W_{i\Delta}^H, W_{j\Delta}^H)]_{i,j=1,2,\ldots,N}$$

Then, by Malliavin calculus, it is possible to prove that the true MLE estimators

$$\hat{\mu}_N = \frac{\mathbf{t}'\Gamma_H^{-1}\mathbf{X}}{\mathbf{t}'\Gamma_H^{-1}\mathbf{t}} \qquad \text{and} \qquad \hat{\sigma}_N = \frac{1}{N}\frac{(\mathbf{X}'\Gamma_H^{-1}\mathbf{X})(\mathbf{t}'\Gamma_H^{-1}\mathbf{t}) - (\mathbf{t}'\Gamma_H^{-1}\mathbf{X})^2}{\mathbf{t}'\Gamma_H^{-1}\mathbf{t}}$$

are strongly consistent as $N \to \infty$ [ though $\mathbf{E}(\hat{\sigma}_N^2) = \frac{N-1}{N}\sigma^2$ ]

# fOU estimation: true MLE

Central Limit Theorems. Further, it is possible to prove that

$$\sqrt{\mathbf{t}'\Gamma_H^{-1}\mathbf{t}}(\hat{\mu}_N - \mu) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2\right)$$

and

$$\frac{1}{\sigma^2}\sqrt{\frac{N}{2}}\left(\hat{\sigma}_N^2 - \sigma^2\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2\right)$$

as $N \to \infty$.

Simulations show that the empirical variance of $\hat{\mu}_N$ increases with $H$.

# fOU estimation: contrast functions

Ludena (2004) [11] Let $H \in (\frac{1}{2}, \frac{3}{4})$ be known and consider the "vanishing drift" fOU process

$$Y_t = y_0 + \int_0^t \sigma(\theta, Y_s) \mathrm{d}B_s^H$$

Let $U_N(\theta) = \frac{1}{N} \sum_{k=1}^{N} h(\theta, X_{k\Delta}, \Delta X_k n^H)$ with $h = h(\theta, x, y)$ at most of polynomial growth in $x$ and $y$. The minimum contrast estimator

$$\hat{\theta}_N = \arg\min_{\theta \in \Theta} U_N(\theta)$$

is asymptotically Gaussian, i.e. $\sqrt{N}(\hat{\theta}_N - \theta) \stackrel{d}{\to} \mathcal{N}$.

Result extends to the following fOU model

$$dY_t = -\lambda Y_t dt + \sigma(\theta) dW_t^H$$

with $\lambda > 0$ known.

# fOU estimation: contrast functions

Neuenkirch and Tindel (2011) [13] Let $H \in (\frac{1}{2}, 1)$ be known and consider the fSDE

$$Y_t = y_0 + \int_0^t b(Y_s; \theta)\mathrm{d}s + \sum_{j=1}^{m} \sigma_j W_t^{(j)H}$$

where $\mathbf{W}_t^H = (W_t^{(1)H}, W_t^{(2)H}, \ldots, W_t^{(m)H})'$ is an $m$-dimensional fBM, $\sigma_j$, $j = 1, 2, \ldots, m$ and $b(y, \theta)$ are known (up to $\theta$). Let $\Delta = \kappa N^{-\alpha}$, $\alpha \in (0, 1)$ and $\kappa > 0$. Let

$$Q_N(\theta) = \frac{1}{N\Delta^2} \sum_{k=0}^{N-1} \left( |\Delta X_k - b(X_k; \theta)\Delta|^2 - \sum_{j=1}^{m} |\sigma_j|^2 \Delta^{2H} \right)$$

then, the least squares estimator $\hat{\theta}_N = \arg\min_{\theta \in \Theta} |Q_N(\theta)|$ is strongly consistent.

For the special case $dY_t = \theta Y_t dt + dW_t^H$, $\hat{\theta}_N$ is explicit.

# fOU estimation: plug-in

Xiao *et al.* (2011) [16] Let $H \in (\frac{1}{2}, 1)$ be known and consider the fOU process

$$dY_t = -\lambda Y_t dt + \sigma dW_t^H$$

The estimator

$$\hat{\sigma}_N^2 = \frac{\Gamma(3-2H)}{2H\Gamma^3(\frac{3}{2}-H)\Gamma(H+\frac{1}{2})(N\Delta)^{2-2H}} \times$$
$$\sum_{j=1}^{N}\left(\sum_{i=1}^{j}(i\Delta)^{\frac{1}{2}-H}(j\Delta-\Delta-i\Delta)^{\frac{1}{2}-H}\Delta X_i - \sum_{i=1}^{j}(i\Delta)^{\frac{1}{2}-H}(j\Delta-i\Delta)^{\frac{1}{2}-H}\Delta X_i\right)^2$$

is strongly consistent for $\sigma^2$.

Moreover, for $H \in (\frac{1}{2}, \frac{3}{4})$ the estimator $\hat{\lambda}_N$ (with $\sigma$ known)

$$\hat{\lambda}_N = \left(\frac{1}{\sigma^2 H\Gamma(2H)N}\sum_{i=0}^{N} X_i^2\right)^{-\frac{1}{2H}}$$

is also strongly consistent for $\lambda$.

# Our proposal

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

$$dY_t = \lambda Y_t dt + \sigma dW_t^H, \quad t \in [0, T]$$

from discrete observations

$$(X_n := Y_{n\Delta_N}, n = 0, 1, \ldots, N) \ ,$$

where $T = T_N = N\Delta_N \longrightarrow +\infty$ and $\Delta_N \longrightarrow 0$ as $N \longrightarrow +\infty$.

Goal: estimate all three elements of $\vartheta$. As $H$ and $\sigma$ can be efficiently estimated without the knowledge of $\lambda$ we propose a two stage procedure.

# Quadratic generalized variations

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

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

Let it be normalized with

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

In the following, we will also consider dilatated filter $\mathbf{a}^2$ associated to $\mathbf{a}$ defined by

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

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

# Quadratic generalized variations estimators of $H$ and $\sigma$

Denote by

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

the generalized quadratic variations associated to the filter $\mathbf{a}$ (see for instance [7]). Then, the estimators of $H$ and $\sigma$ are as follows

$$\widehat{H}_N = \frac{1}{2} \log_2 \frac{V_{N,\mathbf{a}^2}}{V_{N,\mathbf{a}}}$$

and

$$\widehat{\sigma}_N = \left( -2 \cdot \frac{V_{N,\mathbf{a}}}{\sum_{k,\ell} a_k a_\ell |k-\ell|^{2\widehat{H}_N} \Delta_N^{2\widehat{H}_N}} \right)^{\frac{1}{2}}$$

# Properties of estimators of $H$ and $\sigma$

**Theorem 1.** *Let $\mathbf{a}$ be a filter of order $L \geq 2$. Then, both estimators $\widehat{H}_N$ and $\widehat{\sigma}_N$ are strongly consistent, i.e.*

$$(\widehat{H}_N, \widehat{\sigma}_N) \xrightarrow{a.s.} (H, \sigma) \quad \text{as } N \longrightarrow +\infty.$$

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

$$\sqrt{N}(\widehat{H}_N - H) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma_1(\vartheta, \mathbf{a}))$$

*and*

$$\frac{\sqrt{N}}{\log N}(\widehat{\sigma}_N - \sigma) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma_2(\vartheta, \mathbf{a}))$$

*where $\Gamma_1(\vartheta, \mathbf{a})$ and $\Gamma_2(\vartheta, \mathbf{a})$ symmetric definite positive matrices depending on $\sigma$, $H$, $\lambda$ and the filter $\mathbf{a}$ (see next slide).*

Proof: based on an application of [7, Theorem 3(i)].

# Asymptotic variances of $\hat{H}_N$ and $\hat{\sigma}_N$

Let

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

$$\Gamma_1(\vartheta, \mathbf{a}) = \frac{1}{2\log(2)^2} \sum_{j\in\mathbb{Z}} \left( \rho_H^{\mathbf{a},\mathbf{a}}(i)^2 + \rho_H^{\mathbf{a}^2,\mathbf{a}^2}(i)^2 - 2\rho_H^{\mathbf{a},\mathbf{a}^2}(i)^2 \right)$$

and

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

see also [3].

# Properties of estimators of $H$ and $\sigma$

**Remark 1.** *Classical filters of order $L \geq 1$ are defined by*

$$a_k = c_{L,k} = \frac{(-1)^{1-k}}{2^K}\binom{K}{k} = \frac{(-1)^{1-k}}{2^K}\frac{K!}{k!(K-k)!} \quad \textit{for} \quad 0 \leq k \leq K.$$

*Daubechies filters of even order can also be considered (see [4]), for instance the order 2 Daubechies' filter:*

$$\frac{1}{\sqrt{2}}(.4829629131445341, -.8365163037378077, .2241438680420134, .1294095225512603)$$

**Remark 2.** *For classical order 1 quadratic variations ($L = 1$) and $\mathbf{a} = \left(-\frac{1}{2}, \frac{1}{2}\right)$ we can also obtain consistency for any value of $H$, but the central limit theorem holds only for $H < \frac{3}{4}$ (see [7]).*

# Estimator of $\lambda$

From [5], we know the following result

$$\lim_{t \longrightarrow \infty} \mathbf{Var}(Y_t) = \lim_{t \longrightarrow \infty} \frac{1}{t} \int_0^t Y_t^2 dt = \frac{\sigma^2 \Gamma\left(2H+1\right)}{2\lambda^{2H}} =: \mu_2 \, .$$

This gives a natural plug-in estimator of $\lambda$, namely

$$\widehat{\lambda}_N = \left( \frac{2\, \widehat{\mu}_{2,N}}{\widehat{\sigma}_N^2 \Gamma\left(2\widehat{H}_N + 1\right)} \right)^{-\frac{1}{2\widehat{H}_N}}$$

where $\widehat{\mu}_{2,N}$ is the empirical moment of order 2, *i.e*

$$\widehat{\mu}_{2,N} = \frac{1}{N} \sum_{n=1}^{N} X_n^2.$$

# Properties of the estimator of $\lambda$

**Theorem 2.** *Let* $H \in \left(\frac{1}{2}, \frac{3}{4}\right)$ *and a mesh satisfying the condition* $N\Delta_N^p \longrightarrow 0$*,* $p > 1$*, as* $N \longrightarrow +\infty$*. Then, as* $N \longrightarrow +\infty$*,*

$$\widehat{\lambda}_N \xrightarrow{a.s.} \lambda$$

*and*

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

*where* $\Gamma_3(\vartheta) = \lambda\left(\frac{\sigma_H}{2H}\right)^2$ *and*

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

Proof. based on [5], [9, Lemma 8], [8] and [2].

# Properties of the estimator of $\lambda$

**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* $H$*.*

# Monte Carlo Analysis

# Performance of $\hat{H}_N$ and $\hat{\sigma}_N$

Let $\lambda = 2$ and let $\sigma = 1, 2$ for $H = 0.5, 0.7, 0.9$ in $\mathrm{d}Y_t = -\lambda Y_t \mathrm{d}t + \sigma \mathrm{d}W_t^H$.

| $\widehat{H}_N$ | $H = 0.5$ | $H = 0.7$ | $H = 0.9$ |
|---|---|---|---|
| $\sigma = 1$ | 0.499 (0.035) | 0.697 (0.033) | 0.898 (0.031) |
| $\sigma = 2$ | 0.498 (0.033) | 0.700 (0.034) | 0.898 (0.033) |

| $\widehat{\sigma}_N$ | $H = 0.5$ | $H = 0.7$ | $H = 0.9$ |
|---|---|---|---|
| $\sigma = 1$ | 1.024 (0.262) | 1.016 (0.282) | 1.081 (0.437) |
| $\sigma = 2$ | 2.035 (0.510) | 2.073 (0.564) | 2.213 (1.110) |

Table 1: Mean average (sd parenthesis) of 500 Monte-Carlo simulations for the estimation of $H$ (left) and $\sigma$ (right) for different cases. Here $T = 100$, $N = 1000$ and $\lambda = 2$.

| $\widehat{H}_N$ | $H = 0.5$ | $H = 0.7$ | $H = 0.9$ |
|---|---|---|---|
| $\sigma = 1$ | 0.500 (0.003) | 0.700 (0.003) | 0.900 (0.003) |
| $\sigma = 2$ | 0.500 (0.004) | 0.700 (0.003) | 0.900 (0.003) |

| $\widehat{\sigma}_N$ | $H = 0.5$ | $H = 0.7$ | $H = 0.9$ |
|---|---|---|---|
| $\sigma = 1$ | 1.000 (0.025) | 1.001 (0.026) | 0.999 (0.036) |
| $\sigma = 2$ | 2.001 (0.053) | 2.002 (0.053) | 1.997 (0.073) |

Table 2: Mean average (sd in parenthesis) of 500 Monte-Carlo simulations for the estimation of $H$ (left) and $\sigma$ (right) for different cases, and for $T_N = 100$, $N = 100000$ and $\lambda = 2$.

# Performance of $\hat{\lambda}_N$

Let $\lambda = 0.5, 1$, $H = 0.5, 0.6, 0.7$ with $\sigma = 1$ in $\mathrm{d}Y_t = -\lambda Y_t \mathrm{d}t + \sigma \mathrm{d}W_t^H$.

| $\hat{\lambda}_N$ | $H = 0.5$ | $H = 0.6$ | $H = 0.7$ |
|---|---|---|---|
| $\lambda = 0.5$ | 0.093 (0.037) | 0.214 (0.057) | 0.353 (0.069) |
| $\lambda = 1$ | 0.138 (0.052) | 0.276 (0.068) | 0.432 (0.078) |

| $\hat{\lambda}_N$ | $H = 0.5$ | $H = 0.6$ | $H = 0.7$ |
|---|---|---|---|
| $\lambda = 0.5$ | 0.476 (0.148) | 0.514 (0.166) | 0.605 (0.298) |
| $\lambda = 1$ | 0.906 (0.227) | 0.940 (0.238) | 1.005 (0.412) |

Table 3: Mean average (and standard deviation in parenthesis) of 500 Monte-Carlo simulation for the estimation of $\lambda$ for different values of $H$ and $\lambda$. Here $\sigma = 1$ and $T_N = 1$ and $N = 100000$ (left) and $T_N = 100$ and $N = 1000$ (right).

The value of $T_N$ is important for the estimation of the drift. 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.

# Asymptotic distribution of $\hat{\lambda}$

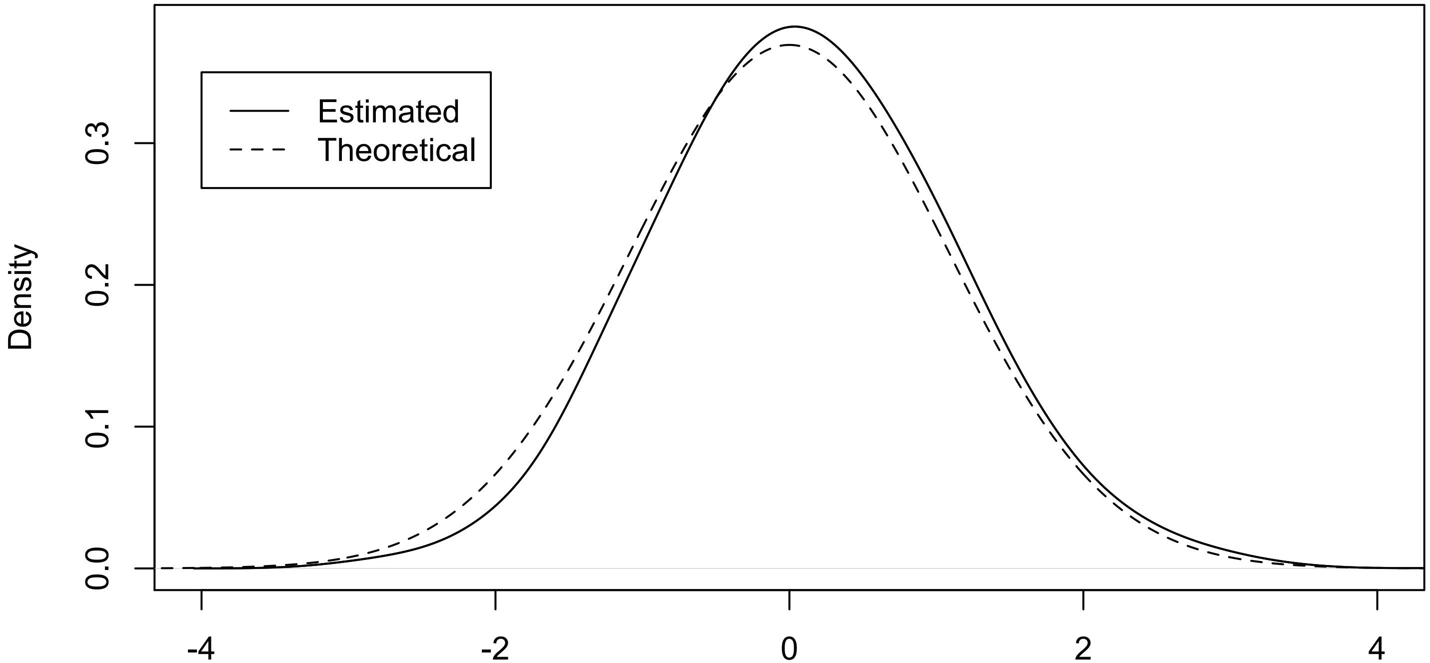



Figure 1: Kernel estimation for the density of $\left(\sqrt{T_N}\left(\widehat{\lambda}_N^{(m)} - \lambda\right)\right)_{m=1\ldots M}$, $M = 5000$, for $T_N = 1000$ and $T_N = 100000$ (fill line) and the theoretical Gaussian density $\mathcal{N}(0, \Gamma_3(\vartheta))$ (dashed line) for $\vartheta = (\lambda, \sigma, H) = (0.3, 1, 0.7)$ (for the value of $\Gamma_3(\vartheta)$ see Theorem 2).

# The YUIMA package

# The YUIMA R package

The Yuima Project aims at implementing, via the `yuima` package, a very abstract framework to describe probabilistic and statistical properties of stochastic processes in a way which is the closest as possible to their mathematical counterparts but also computationally efficient.

- it is an R package, using `S4` classes and methods, where the basic class extends to SDE's with jumps (simple Poisson, Lévy), SDE's driven by fBM, Markov switching regime processes, HMM, etc.
- separates the data description from the inference tools and simulation schemes
- the design allows for multidimensional, multi-noise processes specification
- it includes a variety of tools useful in finance, like asymptotic expansion of functionals of stochastic processes via Malliavin calculus

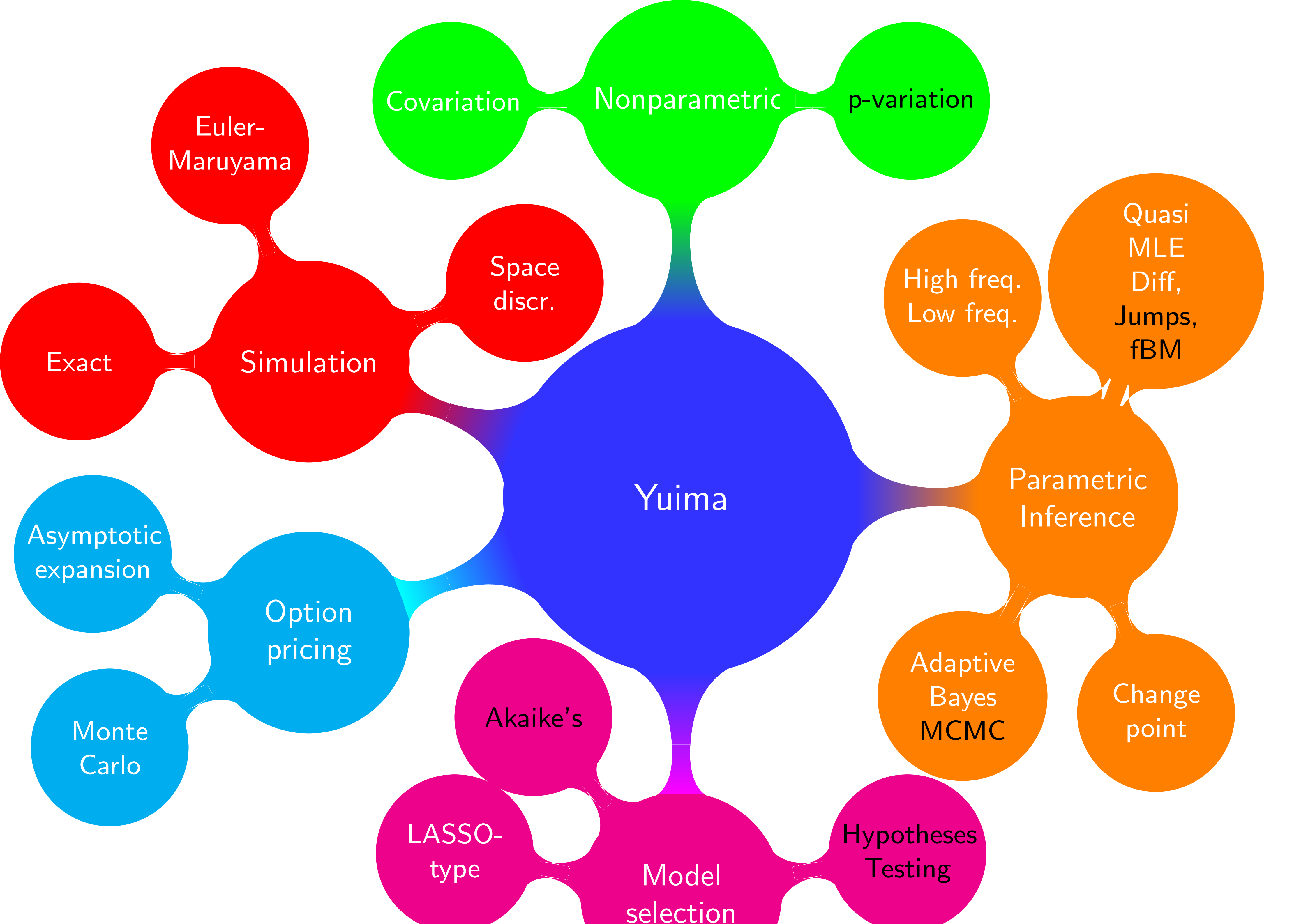
# The `yuima` object

The main object is the `yuima` object which allows to describe the model in a mathematically sound way.

Then the data and the sampling structure can be included as well or, just the sampling scheme from which data can be generated according to the model.

The package exposes very few generic functions like `simulate`, `qmle`, `plot`, etc. and some other specific functions for special tasks.

Before looking at the details, let us see an overview of the main object.


Covariation
Nonparametric
p-variation
Euler-Maruyama
Space discr.
Exact
Simulation
High freq. Low freq.
Quasi MLE Diff, Jumps, fBM
Yuima
Parametric Inference
Asymptotic expansion
Option pricing
Monte Carlo
Akaike's
Adaptive Bayes MCMC
Change point
LASSO-type
Model selection
Hypotheses Testing

# The model specification

We consider here the three main classes of SDE's which can be easily specified. All multidimensional and eventually parametric models.

- Diffusions $\mathrm{d}X_t = a(t, X_t)dt + b(t, X_t)\mathrm{d}W_t$
- Fractional Gaussian Noise, with $H$ the Hurst parameter

$$\mathrm{d}X_t = a(t, X_t)dt + b(t, X_t)\mathrm{d}W_t^H$$

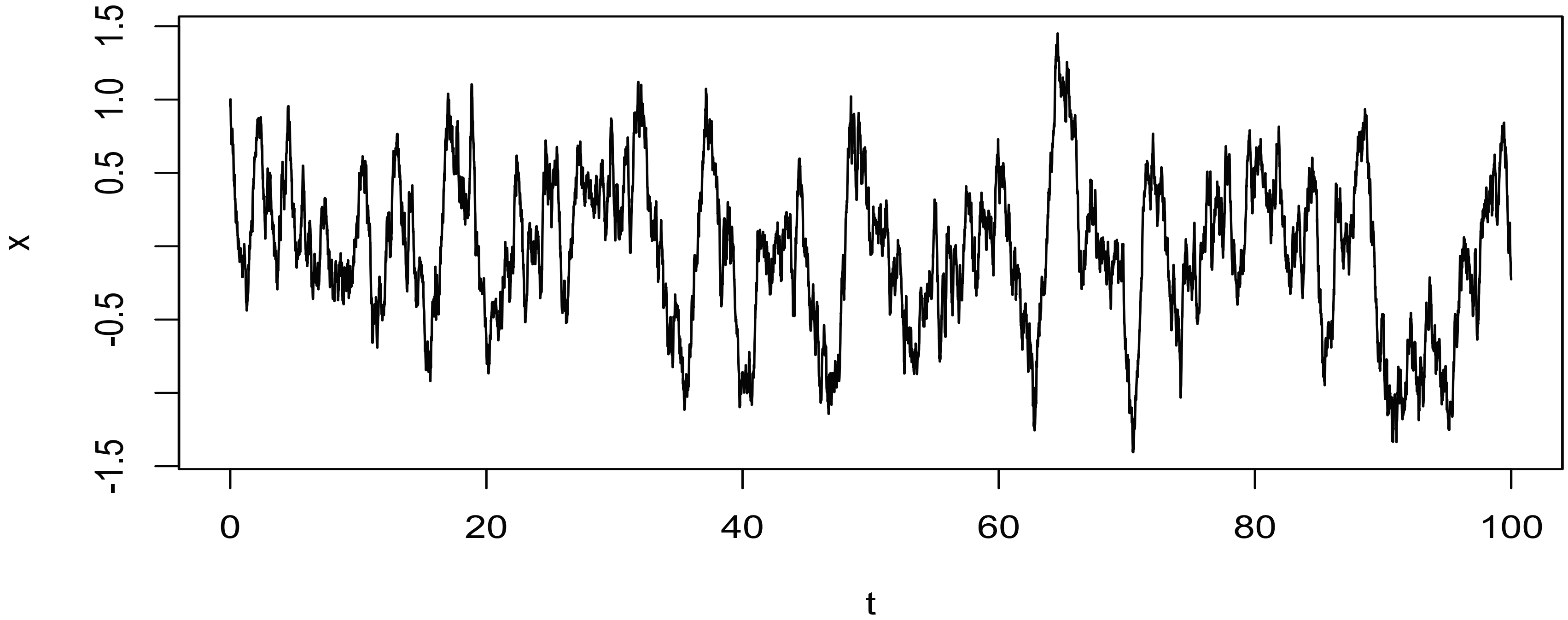
- Diffusions with jumps, Lévy

$$\begin{aligned}\mathrm{d}X_t = \; & a(X_t)dt + b(X_t)\mathrm{d}W_t + \int_{|z|>1} c(X_{t-}, z)\mu(\mathrm{d}t, \mathrm{d}z) \\ & + \int_{0<|z|\leq 1} c(X_{t-}, z)\{\mu(\mathrm{d}t, \mathrm{d}z) - \nu(\mathrm{d}z)\mathrm{d}t\}\end{aligned}$$

# Fractional Gaussian Noise

$$\mathrm{d}X_t = -2X_t\mathrm{d}t + \mathrm{d}W_t^H$$

```
> samp <- setSampling(Terminal=100, n=10000)
> mod <- setModel(drift="-2*x", diffusion="1",hurst=0.7)
> ou <- setYuima(model=mod, sampling=samp)
> fou <- simulate(ou, xinit=1)
```



# Estimation

The parameters $H$ and $\sigma$ can be estimates via the function `qgv` (quadratic generalized variations)

```
> qgv(fou)
```

and the parameter $\lambda$ using the least squares estimator `lse`

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

For more informations and software see

http://R-Forge.R-Project.org/projects/yuima

THANKS

[1] K. Bertin, S. Torres, and C. Tudor. Drift parameter estimation in fractional diffusions driven by perturbed random walk. *Statistic and Probability Letters*, 81(2):243–249, 2011.

[2] P. Cheridito, H. Kawaguchi, and M. Maejima. Fractional Ornstein-Uhlenbeck processes. *Electronic Journal of Probability*, 8(3):1–14, 2003.

[3] J.-F. Cœurjolly. Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths. *Stat. Infer. Stoch. Process.*, 4(2):199–227, 2001.

[4] I. Daubechies. *Ten Lectures on Wavelets*. SIAM, 1992.

[5] Y. Hu and D. Nualart. Parameter estimation for fractional Ornstein-Uhlenbeck processes. *Statistics and Probability Letters*, 80(11-12):1030–1038, 2010.

[6] Y. Hu, D. Nualart, W. Xiao, and W. Zhang. Exact maximum likelihood estimator for drift fractional brownian motion at discrete time. *Acta Mathematica Scientia*, 31(5):1851–1859, 2011.

[7] J. Istas and G. Lang. Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Annales de l'Institut Henri Poincaré*, 23(4):407–436, 1997.

[8] T. Kaarakka and P. Salminem. On fractional Ornstein-Uhlenbeck processes. *Communications on Stochastics Analysis*, 5(1):121–133, 2011.

[9] M. Kessler. Estimation of an ergodic diffusion from discrete observations. *Scandinavian Journal of Statistics*, 24:211–229, 1997.

[10] A. Kolmogorov. Winersche Spiralen und einige andere interessante Kurven in Hilbertschen Raum. *Acad. Sci. USSR*, 26:115–118, 1940.

[11] C. Ludena. Minimum contrast estimation for fractional diffusion. *Scandinavian Journal of Statistics*, 31:613–628, 2004.

[12] B. Mandelbrot and J. Van Ness. Fractional Brownian motions, fractional noises and application. *SIAM Review*, 10:422–437, 1968.

[13] A. Neuenkirch and S. Tindel. A Least Square-type procedure for parameter estimation in stochastic differential equations with additive fractional noise. *preprint*, 2011.

[14] T. Sottinen. Fractional brownian motion, random walks and binary market models, journal = Finance and Stochastics, volume = 5, pages = 343–355, year = 2001,.

[15] F. Viens and C. Tudor. Variations and estimators for the selfsimilarity order through malliavin calculus. *Annals of Probability*, 6:2093–2134, 2009.

[16] W. Xiao, W. Zhang, and W. Xu. Parameter estimation for fractional Ornstein-Uhlenbeck processes at discrete observation. *Applied Mathematical Modelling*, 35:4196–4207, 2011.