KUWAIT UNIVERSITY

# BERRY-ESSEEN BOUNDS OF SECOND MOMENT ESTIMATORS FOR GAUSSIAN PROCESSES OBSERVED AT HIGH FREQUENCY

Khalifa ES-SEBAIY

**International Workshop "Statistics of Stochastic Processes in Discrete and Continuous Time"**
**Taras Shevchenko National University of Kyiv**
**October 11-12, 2022**

*Ref.: S. Douissi, K. Es-Sebaiy, G. Kerchev, I. Nourdin. Berry-Esseen bounds of second moment estimators for Gaussian processes observed at high frequency, Electron. J. Statist. 16(1): 636-670 (2022).*

# Outline

♣ Malliavin calculus and Motivation

# Outline

- Malliavin calculus and Motivation
- Berry-Esseen bounds of second moment estimators for stationary and asymptotically stationary Gaussian processes

# Outline

- Malliavin calculus and Motivation
- Berry-Esseen bounds of second moment estimators for stationary and asymptotically stationary Gaussian processes
- Applications to Gaussian Ornstein-Uhlenbeck processes

Initiated by Paul Malliavin in 1976, the *Malliavin calculus* also known as the stochastic calculus of variations is an infinite-dimensional differential calculus, whose operators act on functionals of Gaussian/Lévy processes.

# Malliavin calculus

- The main motivation for introducing Malliavin calculus was to study the regularity properties of the laws of Wiener functionals such as the solutions of SDEs but later several other applications of this theory emerged ; anticipative stochastic calculus, stochastic PDEs, stochastic filtering and also in mathematical finance.
- In this talk, we will see some applications of this calculus to statistical inference for stochastic processes.

Let $\mathcal{H}$ be a real separable Hilbert space with inner product $\langle ., . \rangle_{\mathcal{H}}$.

- An **isonormal Gaussian process** over $\mathcal{H}$ is a centered Gaussian family $G = \{G(\varphi), \varphi \in \mathcal{H}\}$ of random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ such that : $\forall \varphi, \psi \in \mathcal{H}$
$$\mathbf{E}[G(\varphi)G(\psi)] = \langle \varphi, \psi \rangle_{\mathcal{H}}.$$
- $\mathcal{H}^{\otimes p}$ and $\mathcal{H}^{\odot p}$ for any integer $p \geq 1$ denote the $p$-th tensor product and the $p$-th symmetric tensor product of $\mathcal{H}$ respectively.

- The **Wiener chaos of order** $p$ associated with $G$ which will be denoted by $\mathcal{H}_p$ is by definition the closed linear subspace of $L^2(\Omega)$ generated by the random variables $\{H_p(G(\varphi)) : \varphi \in \mathcal{H}, \|\varphi\|_{\mathcal{H}} = 1\}$, where $H_p$ is the $p$-th Hermite polynomial defined by

$$H_p(x) = (-1)^p e^{\frac{x^2}{2}} \frac{d^p}{dx^p} e^{-\frac{x^2}{2}}, \quad p \geq 1,$$

and $H_0(x) = 1$, $H_1(x) = x$, $H_2(x) = x^2 - 1$.

- The $p$**-th multiple integral** of $\varphi^{\otimes p} \in \mathcal{H}^{\odot p}$ is defined by the equality $I_p(\varphi^{\otimes p}) = H_p(G(\varphi))$ for any $\varphi \in \mathcal{H}$ with $\|\varphi\|_{\mathcal{H}} = 1$. Moreover, the map $I_p$ is a linear isometry between $\mathcal{H}^{\odot p}$ equipped with the norm $\sqrt{p!}\|.\|_{\mathcal{H}^{\otimes p}}$ and $\mathcal{H}_p$ under the $L^2(\Omega)$'s norm.

For $p = 1$ and $p = 2$ we have the following:

$$I_1(\varphi) = H_1(G(\varphi)) = G(\varphi) \quad (1)$$

$$I_2(\varphi^{\otimes 2}) = H_2(G(\varphi)) = G(\varphi)^2 - 1. \quad (2)$$

- **The Wiener chaos expansion**. For any $F \in L^2(\Omega)$, there exists a unique sequence of functions $f_p \in \mathcal{H}^{\odot p}$ such that

$$F = \mathbf{E}[F] + \sum_{p=1}^{\infty} I_p(f_p),$$

  where the terms are all mutually orthogonal in $L^2(\Omega)$ and $f_p \in \mathcal{H}^{\odot p}$ are uniquely determined by $F$.

- **Isometry Property**. For any integers $1 \leq q \leq p$ and $f \in \mathcal{H}^{\odot p}$ and $g \in \mathcal{H}^{\odot q}$ with $\|f\|_{\mathcal{H}} = \|g\|_{\mathcal{H}} = 1$, we have

$$\mathbf{E}[I_p(f)I_q(g)] = \begin{cases} p!\langle f, g\rangle_{\mathcal{H}^{\otimes p}} & \text{if } p = q \\ 0 & \text{otherwise.} \end{cases}$$

**Remark**

*Notice that when the Gaussian process $G$ is the standard Brownian motion $W$ over $\mathcal{H} = L^2([0,1])$ in this case for $f \in L^2([0,1]^p)$,*

$$I_p(f) = \int_{[0,1]^p} f(t_1, ..., t_p) dW_{t_1} ... dW_{t_p}$$

*is nothing but a multiple Wiener-Itô integral of order $p$.*

- The fOU process is one of the most studied and widely applied stochastic process. It represents **interesting model for stochastic dynamics with memory**, with applications to e.g. finance, telecommunication networks and physics. In the finance context, several researchers in recent years have been interested in studying statistical estimation problems for fOU processes.
- The **statistical analysis of equations driven by fBm** is obviously more recent. The development of stochastic calculus with respect to the fBm allowed to study such models.
- On the other hand, the **long-range dependence property** makes the fBm important driving noise in modeling several phenomena arising, for instance, from volatility modeling in finance.

# Statistical inference for stochastic processes

- **The problem:** We want to estimate some unknown parameters appearing in SDEs driven by fractional Gaussian noises.
- One of the most popular fractional Gaussian stochastic processes is the fractional Brownian motion (fBm for short) $B^H$ of Hurst parameter $H \in (0,1)$. It's a centered Gaussian process with covariance function :

$$R_H(t,s) := \mathbf{E}(B_t^H B_s^H) = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right) \quad s,t \in \mathbb{R}_+,$$

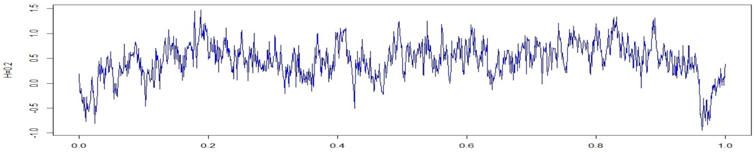


Figure: Sample path of a fBm for $H = 0.2$

# Fractional Ornstein-Uhlenbeck process

One of the estimation problems that has been actively studied in the last 10 years, is the estimation of the drift of the fOU which by definition solves the following linear SDE :

$$dX_t = -\theta X_t dt + dB_t^H, \quad t \geq 0, \qquad X_0 = 0, \tag{3}$$

where $B^H$ is a fBm of Hurst parameter $H \in (0,1)$, $\theta > 0$ is the drift assumed to be unknown. The SDE (3) has an explicit solution given by :

$$X_t = \int_0^t e^{-\theta(t-u)} dB_u^H = I_1^{B^H}(e^{-\theta(t-.)} \mathbf{1}_{[0,t]}(.)), \quad t \geq 0. \tag{4}$$

The main approaches used :

- **The maximum likelihood approach** by computing the MLE.
  Klepsyna and Le Breton 2002. The expression of MLEs are related to the fractional operators with respect to the Hurst parameter $H$.
  ▷ Hard to discretize it in practice, due to its complicated expression.
- **The least squares approach** by computing the LSE.
  Hu and Nualart 2010. LSE minimizes $\theta \mapsto \int_0^T |\dot{X}_t + \theta X_t|^2 dt$. Its expression is given by

$$\hat{\theta}_T = \theta - \frac{\int_0^T X_t \delta B_t^H}{\int_0^T X_t^2 dt},$$

  where $\int_0^T X_t \delta B_t^H$ is a Skorohod-type integral.
  ▷ However, this LSE depends on $\theta$, the parameter we want to estimate! Therefore, the LSE $\hat{\theta}_T$ is not an estimator at all!

- Hu and Nualart 2010 proposed an alternative estimator :

$$\left(\frac{1}{H\Gamma(2H)T}\int_0^T X_t^2 dt\right)^{-\frac{1}{2H}}. \tag{5}$$

▷ It's hard to observe a process continuously in time.

- Discretizing time is the only practical way of estimating parameters for continuous-time processes.
  For instance the natural Riemann-sum-type time discretization of (5) given by

$$\left(\frac{1}{nH\Gamma(2H)}\sum_{i=1}^{n} X_{t_i}^2\right)^{-\frac{1}{2H}} \tag{6}$$

with $t_i = i \times \Delta_n$. $\Delta_n$ denotes the mesh of the observations

- **The method of moments** Es-Sebaiy and Viens 2019.
  The estimators are based on the following polynomial variations

$$Q_{f_q,n} := \frac{1}{n}\sum_{i=1}^{n} f_q(Z_i)$$

  $\{Z_i = I_1^G(\varepsilon_i), i \geq 1\}$ is assumed to be stationary, $\varepsilon_i \in \mathcal{H}$, $f_q$ is a polynomial function of degree $q \geq 2$ ($q$ an even integer) given by

$$f_q(Z_i) = \sum_{k=1}^{q/2} d_{f_q,2k} I_{2k}^G(\varepsilon_i^{\otimes 2k}).$$

- The expression of the estimators rely directly on the vector of the observations $(Z_1, .., Z_n)$.
- Assume that the mesh of the observations ($\Delta_n = 1$).

When the mesh is assumed to be a fixed constant we solved many statistical parameter estimation problems when :

- Z is stationary or asymptotically stationary. Es-Sebaiy and Viens 2019.
- Z is not stationary. Douissi, Es-Sebaiy, Viens 2019 and Douissi, Alshahrani, Es-Sebaiy, Viens 2020

Now, we investigate another question: what happens if the mesh is no longer fixed? Namely, if the step size $\Delta_n$ depends on the sample size $n$ such that :

- $\Delta_n \to 0$ when $n \to +\infty$
- $T_n = n\Delta_n \to +\infty$ when $n \to +\infty$,

Can we solve the same parameter estimation problem even when the process is assumed to be continuously observed in time ? We answered this question in the recent publication Douissi, Es-Sebaiy, Kerchev, Nourdin 2022.

# Berry-Esséen bounds of second moment estimators for Gaussian processes observed at high frequency.

Let $Z = \{Z_t, t \geq 0\}$ be a continuous centered stationary Gaussian process and $f_Z := \mathbf{E}(Z_0^2) > 0$. We consider the following estimators of $f_Z$:

- When a complete path of the process over a large finite interval is observable, we use the estimator:

$$\widehat{f}_T(Z) := \frac{1}{T} \int_0^T Z_t^2 dt, \quad T > 0. \tag{7}$$

- A more practical assumption is that the process $Z$ is observed at discrete time instants $t_i = i\Delta_n$, where $i = 0, \ldots, n$ and $\Delta_n$ is the step size. Then we consider the following estimator over the *observation window* $T_n := n\Delta n$:

$$\widetilde{f}_n(Z) := \frac{1}{n} \sum_{i=1}^{n} Z_{t_i}^2, \quad n \geq 1. \tag{8}$$

Let $Z := \{Z_t, t \geq 0\}$ be a continuous centered stationary Gaussian process that can be represented as a Wiener-Itô (multiple) integral $Z_t = I_1(\mathbf{1}_{[0,t]})$ for every $t \geq 0$. Let $\rho(r) = E(Z_r Z_0)$ denote the covariance of $Z$ for every $r \geq 0$, and let $\rho(r) = \rho(-r)$ for all $r < 0$.

Our main assumption throughout the paper is that

$$\sigma_Z^2 := 4 \int_{\mathbb{R}} \rho^2(r) dr < \infty.$$

# Continuous-time observations

We estimate the variance $f_Z := E(Z_0^2)$, when the whole trajectory of $Z$ is observed up to time $T > 0$. We consider the estimator (7) given by

$$\widehat{f}_T(Z) = \frac{1}{T}\int_0^T Z_t^2 dt, \quad T > 0$$

as a statistic to estimate $f_Z$, based on the continuous-time observation of $Z$. First we show some simpler properties of $\widehat{f}_T(Z)$.

**Lemma**

*The estimator $\widehat{f}_T(Z)$ is unbiased and strongly consistent. In particular,*

$$\sqrt{T}\left\|\widehat{f}_T(Z) - f_Z\right\|_{L^2} \uparrow \sigma_Z \text{ as } T \to \infty. \tag{9}$$

## Theorem

*Assume* $\int_{\mathbb{R}} \rho^2(r)dr < \infty$*. Let* $\mathcal{N} \sim \mathcal{N}(0,1)$ *be the standard normal random variable. Then for all* $T > 0$*,*

$$d_{TV}\left(\frac{\widehat{f}_T(Z) - f_Z}{\sqrt{Var(\widehat{f}_T(Z) - f_Z)}}, \mathcal{N}\right) \leq \varphi_T(Z), \tag{10}$$

*where*

$$\varphi_T(Z) = C \max\left\{\frac{8}{\sqrt{T}}\left(\int_{-T}^{T} |\rho(t)|^{3/2} dt\right)^2, \frac{48}{T}\left(\int_{-T}^{T} |\rho(t)|^{4/3} dt\right)^3\right\},$$

*for some absolute constant* $C > 0$*. The same result holds for the Wasserstein distance.*

**Proof**. The random variable

$$\frac{\widehat{f}_T(Z) - f_Z}{\sqrt{Var(\widehat{f}_T(Z) - f_Z)}} = \frac{V_T(Z)}{\sqrt{E[V_T(Z)^2]}} = \frac{\frac{1}{\sqrt{T}} \int_0^T \left(Z_t^2 - E[Z_t^2]\right) dt}{\sqrt{E[V_T(Z)^2]}}$$

is centered and normalized. By the 4th moment theorem,

$$\begin{aligned} d_{TV}\left(\frac{\widehat{f}_T(Z) - f_Z}{\sqrt{Var(\widehat{f}_T(Z) - f_Z)}}, \mathcal{N}\right) &= \quad d_{TV}\left(\frac{V_T(Z)}{\sqrt{E[V_T(Z)^2]}}, \mathcal{N}\right) \\ \leq \quad & C \max\left\{\frac{\kappa_3(V_T(Z))}{E[V_T(Z)^2]^{3/2}}, \frac{\kappa_4(V_T(Z))}{E[V_T(Z)^2]^2}\right\}, \end{aligned}$$

where, for every $T > 0$,

$$\begin{aligned} |\kappa_3(V_T(Z))| &\leq \quad \frac{8}{\sqrt{T}} \left(\int_{-T}^{T} |\rho(t)|^{3/2} \, dt\right)^2, \\ |\kappa_4(V_T(Z))| &\leq \quad \frac{48}{T} \left(\int_{-T}^{T} |\rho(t)|^{4/3} \, dt\right)^3. \end{aligned}$$

## Corollary

*Assume there exists* $0 < \beta < \frac{3}{4}$ *such that,* $|\rho(t)| = \mathcal{O}(t^{2\beta-2})$*.*
*Then,*

$$d_{TV}\left(\frac{\widehat{f}_T(Z) - f_Z}{\sqrt{Var(\widehat{f}_T(Z) - f_Z)}}, \mathcal{N}\right) \leq C \begin{cases} T^{-1/2} & \text{if } 0 < \beta < \frac{2}{3}, \\ \log^2(T) T^{-1/2} & \text{if } \beta = \frac{2}{3}, \\ T^{6\beta - \frac{9}{2}} & \text{if } \frac{2}{3} < \beta < \frac{3}{4}. \end{cases} \quad (11)$$

## Corollary

*There exists a constant $C > 0$ such that, for all $T > 0$,*

$$d_{TV}\left(\frac{\sqrt{T}}{\sigma_Z}(\widehat{f}_T(Z) - f_Z), \mathcal{N}\right) \leq \varphi_T(Z) + 2\left|1 - \frac{\sigma_Z^2}{E(V_T(Z)^2)}\right|$$

*where $\varphi_T(Z)$ is as in the previous Theorem. Moreover, if there exists $0 < \beta < \frac{3}{4}$ such that, $|\rho(t)| = \mathcal{O}(t^{2\beta-2})$,*

$$d_{TV}\left(\frac{\sqrt{T}}{\sigma_Z}(\widehat{f}_T(Z) - f_Z), \mathcal{N}\right) \leq C \begin{cases} T^{-1/2} & \text{if } 0 < \beta \leq \frac{5}{8}, \\ T^{4\beta-3} & \text{if } \frac{5}{8} < \beta < \frac{3}{4}, \end{cases} \tag{12}$$

*using, see Es-Sebaiy and Viens (2019),*
$d_{TV}(\mu + \sigma F, \mathcal{N}) \leq d_{TV}(F, \mathcal{N}) + \sqrt{\frac{\pi}{2}}|\mu| + 2\left|1 - \frac{1}{\sigma^2}\right|$.

# Discrete-time observations

In this section we estimate the limiting variance $f_Z$ based on discrete high-frequency data in time of $Z$, by considering the discrete version $\widetilde{f}_n(Z)$ of the estimator $\widehat{f}_T(Z)$:

$$\widetilde{f}_n(Z) := \frac{1}{n} \sum_{i=1}^{n} Z_{t_i}^2,$$

where $t_i = i\Delta_n$, $i = 0, \ldots, n$, $\Delta_n \to 0$ and $T_n := n\Delta_n \to \infty$.

**Assumption**

*For all $s, t \in \mathbb{R}_+$ such that $|s-t|$ is small enough,*

$$E[|Z_t - Z_s|^2] \leq C|t-s|^{2\alpha}, \tag{13}$$

*for some constant $0 < \alpha < 1$.*

Similarly to the continuous alternative $\widehat{f}_T(Z)$, we first show the following:

**Lemma**

*The estimator $\widetilde{f}_n(Z)$ is unbiased. Assume that Assumption 1 holds. Then,*

$$E|\widehat{f}_T(Z) - \widetilde{f}_n(Z)|^2 \leq C_\alpha \Delta_n^{2\alpha}, \tag{14}$$

*where $C_\alpha > 0$ is a constant that depends only on $\alpha$. Moreover, if $n\Delta_n^\eta \to 0$, as $n \to \infty$ for some $\eta > 1$, then $\widetilde{f}_n(Z)$ is strongly consistent.*

## Theorem

*Assume* $\int_{\mathbb{R}} \rho^2(r)dr < \infty$ *and that*

$$E[|Z_t - Z_s|^2] \leq c|t-s|^{2\alpha}$$

*for some* $c > 0$ *and* $\alpha \in (0,1)$ *and when* $|t-s|$ *is small enough. Let* $\mathcal{N} \sim \mathcal{N}(0,1)$*. If* $\Delta_n \to 0$ *and* $n\Delta_n \to \infty$ *as* $n \to \infty$*, then there is* $C > 0$ *such that, for every* $n \geq 1$*,*

$$d_{TV}\left(\frac{\widetilde{f}_n(Z) - f_Z}{\sqrt{Var(\widetilde{f}_n(Z) - f_Z)}}, \mathcal{N}\right)$$

$$\leq \varphi_{T_n}(Z) + 2\left|1 - \frac{Var(\widehat{f}_{T_n}(Z) - f_Z)}{Var(\widetilde{f}_n(Z) - f_Z)}\right| + C\left[n\Delta_n^{2\alpha+1}\right]^{1/4},$$

*using* $d_{TV}(F,G) \leq C\left(\frac{E[(F-G)^2]}{E[F^2]}\right)^{1/4}$*, see Kosov 2019).*

## Corollary

*Let* $\sigma_Z^2 = 4\int_0^\infty \rho(u)^2 du$. *Under the same assumptions as in Theorem 7, for all* $n \geq 1$,

$$d_{TV}\left(\frac{\sqrt{T}}{\sigma_Z}(\widetilde{f}_n(Z) - f_Z), \mathcal{N}\right) \leq \varphi_{T_n}(Z) + 2\left|1 - \frac{\sigma_Z^2}{E(V_{T_n}(Z)^2)}\right| + C(n\Delta_n^{2\alpha+1})^{1/4}, \qquad (15)$$

*where* $C > 0$ *is an absolute constant depending on* $E[Z_0^4]$. *The same result holds for the Wasserstein distance.*

# Parameter estimation for non-stationary Gaussian processes

Let $Z$ be a centered stationary Gaussian process and let $Y$ be a stochastic process satisfying the following: there exists a constant $\gamma > 1$ such that for every $p \geq 1$ and for all $T > 0$,

$$\|Y_T\|_{L^p} = \mathcal{O}\left(T^{-\gamma}\right). \tag{16}$$

Then we consider second moment estimators for the process $X := Z + Y$. We consider the second moment estimators $\widehat{f}_T(X)$ and $\widetilde{f}_n(X)$, based on continuous-time and discrete-time observations of $X$:

$$\widehat{f}_T(X) := \frac{1}{T}\int_0^T X_t^2 dt, \quad T > 0, \tag{17}$$

$$\widetilde{f}_n(X) := \frac{1}{n}\sum_{i=1}^{n} X_{t_i}^2, \quad n \geq 1, \tag{18}$$

where $t_i = i\Delta_n$, $i = 0, \ldots, n$, $\Delta_n \to 0$ and $T_n := n\Delta_n \to \infty$.

### Proposition

*If* $\int_{\mathbb{R}} \rho^2(r)dr < \infty$ *and the condition (16) holds, then*

$$\left\| \widehat{f}_T(X) - f_X \right\|_{L^2} \leq CT^{-1/2}. \tag{19}$$

*Moreover,*

$$\widehat{f}_T(X) \longrightarrow f_X \quad \text{almost surely as } T \to \infty. \tag{20}$$

*In addition, if* $Z$ *satisfies the helix property (13), and* $n\Delta_n^{2\alpha+1} \to 0$*, as* $n \to \infty$*, then*

$$\widetilde{f}_n(X) \longrightarrow f_X \quad \text{almost surely as } T \to \infty. \tag{21}$$

## Theorem

*Assume that* $\int_{\mathbb{R}} \rho^2(r)dr < \infty$ *and that the condition (16) holds. Let* $\mathcal{N} \sim \mathcal{N}(0,1)$ *be the standard normal random variable. Then, there exists a constant* $C > 0$ *such that, for all* $T > 0$*,*

$$d_{TV}\left(\frac{\widehat{f}_T(X) - f_X - E[\widehat{f}_T(X) - f_X]}{\sqrt{Var(\widehat{f}_T(X) - f_X)}}, \mathcal{N}\right) \le \varphi_T(Z) + CT^{\frac{1-\gamma}{4}}, \tag{22}$$

*where* $\varphi_T(Z)$ *is defined in* (11)*. The same result holds for the Wasserstein distance.*

## Theorem

*Assume* $\int_{\mathbb{R}} \rho^2(r)dr < \infty$*, Z verifies (13), and the condition (16) holds. Let* $\mathcal{N} \sim \mathcal{N}(0,1)$*. Then, there is* $C > 0$ *such that, for every* $n \geq 1$*,*

$$d_{TV}\left(\frac{\sqrt{T_n}}{\sigma_Z}(\widetilde{f}_n(X) - f_X), N\right) \leq \varphi_{T_n}(Z) + C\left[n\Delta_n^{2\alpha+1}\right]^{1/4} + CT_n^{\frac{1-\gamma}{4}}.$$

# Application to Gaussian Ornstein-Uhlenbeck processes

We consider the Ornstein-Uhlenbeck process $X^\theta := \left\{X_t^\theta, t \geq 0\right\}$ driven by a fractional Brownian motion $\left\{B_t^H, t \geq 0\right\}$ of Hurst index $H \in (0,1)$. More precisely, $X^\theta$ is the solution of the following linear stochastic differential equation

$$X_0^\theta = 0; \quad dX_t^\theta = -\theta X_t^\theta dt + dB_t^H, \quad t \geq 0, \tag{23}$$

where $\theta > 0$ is an unknown parameter.
There is an explicit solution to (23) which is

$$X_t^\theta = \int_0^t e^{-\theta(t-s)} dB_s^H. \tag{24}$$

Moreover,

$$Z_t^\theta = \int_{-\infty}^t e^{-\theta(t-s)} dB_s^H. \tag{25}$$

is a stationary Gaussian process

We have

$$\begin{aligned}E[(Z_0^\theta)^2] =&E\left(\int_{-\infty}^0 e^{\theta s} dB_s^H\right)^2 = E\int_{-\infty}^0\int_{-\infty}^0 B_s^H B_r^H e^{\theta(s+r)} dsdr\\ =&\theta^2\int_{-\infty}^0\int_{-\infty}^0 \frac{1}{2}e^{\theta(s+r)}(|r|^{2H}+|s|^{2H}-|r-s|^{2H})drds\\ =&\theta^{-2H}H\Gamma(2H).\end{aligned}$$

Therefore, $\theta = g_{Z^\theta}(E[(Z_0^\theta)^2]$ where the invertible function $g_{Z^\theta}:\mathbb{R}^+\to\mathbb{R}^+$ is given by

$$g_{Z^\theta}(x) = \left(\frac{H\Gamma(2H)}{x}\right)^{\frac{1}{2H}}, \quad x>0. \tag{26}$$

## Proposition

*Let $H \in (0, 3/4)$. The estimators $\widehat{\theta}_T$ and $\widetilde{\theta}_n$ are strongly consistent. Denote*

$$\delta_H^2 := \frac{\theta}{(2H)^2} \times \begin{cases} (4H-1) + \frac{2\Gamma(2-4H)\Gamma(4H)}{\Gamma(2H)\Gamma(1-2H)} & \text{if } H \in \left(0, \frac{1}{2}\right), \\ (4H-1)\left(1 + \frac{\Gamma(3-4H)\Gamma(4H-1)}{\Gamma(2-2H)\Gamma(2H)}\right) & \text{if } H \in \left[\frac{1}{2}, \frac{3}{4}\right). \end{cases} \tag{27}$$

*The following limit theorems hold:*

1. $\sqrt{T}(\widehat{\theta}_T - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \delta_H^2)$ *as* $T \to \infty$.
2. *Assume there is* $p \in (1, \frac{3+2H}{1+2H} \wedge (1+2H))$ *such that* $n\Delta_n^p \to 0$. *Then* $\sqrt{T_n}(\widetilde{\theta} - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \delta_H^2)$ *as* $n \to \infty$.

### Theorem

*Let* $H \in (0, \frac{3}{4})$*, and* $\delta_H$ *be given by* (27)*. Then*

$$d_W\left(\frac{\sqrt{T}}{\delta_H}(\widehat{\theta}_T - \theta), \mathcal{N}\right) \leq C \begin{cases} \frac{1}{\sqrt{T}} & \text{if } 0 < H \leq \frac{5}{8}, \\ \frac{1}{T^{3-4H}} & \text{if } \frac{5}{8} < H < \frac{3}{4}. \end{cases} \tag{28}$$

*Moreover,*

$$d_W\left(\frac{\sqrt{T_n}}{\delta_H}(\widetilde{\theta}_n - \theta), \mathcal{N}\right)$$
$$\leq C\left[n\Delta_n^{2H+1}\right]^{1/2} + C \begin{cases} \frac{1}{\sqrt{n\Delta_n}} & \text{if } 0 < H \leq \frac{5}{8}, \\ \frac{1}{(n\Delta_n)3-4H} & \text{if } \frac{5}{8} < H < \frac{3}{4}. \end{cases} \tag{29}$$

# Fractional Ornstein-Uhlenbeck of the second kind

The so-called fractional Ornstein-Uhlenbeck process of the second kind, defined via the stochastic differential equation

$$S_0^\mu = 0, \text{ and } dS_t^\mu = -\mu S_t^\mu dt + dY_t^{(1)}, \quad t \geq 0, \tag{30}$$

where $Y_t^{(1)} = \int_0^t e^{-s} dB_{a_s}^H$ with $a_s = He^{\frac{s}{H}}$ and $\left\{B_t^H, t \geq 0\right\}$ is a fBm with Hurst parameter $H \in (\frac{1}{2}, 1)$, and where $\mu > 0$ is the unknown real parameter which we would like to estimate. The equation (30) admits an explicit solution

$$\begin{aligned} S_t^\mu &= e^{-\mu t} \int_0^t e^{\mu s} dY_s^{(1)} = e^{-\mu t} \int_0^t e^{(\mu-1)s} dB_{a_s}^H \\ &= H^{(1-\mu)H} e^{-\mu t} \int_{a_0}^{a_t} r^{(\mu-1)H} dB_r^H . \end{aligned}$$

Hence we can also write

$$S_t^\mu = Z_t^\mu - e^{-\mu t} Z_0^\mu,$$

where

$$Z_t^\mu = e^{-\mu t} \int_{-\infty}^{t} e^{(\mu-1)s} dB_{a_s}^H = H^{(1-\mu)H} e^{-\mu t} \int_0^{a_t} r^{(\mu-1)H} dB_r^H.$$

For every $H \in (\frac{1}{2}, 1)$,

$$g_{Z^\mu}^{-1}(\mu) = f_{X^\mu} = f_{Z^\mu} = E\left[(Z_0^\mu)^2\right] = \frac{(2H-1)H^{2H}}{\mu} \mathcal{B}(1-H+\mu H, 2H-1$$

where here $\mathcal{B}(.)$ is the usual beta function. Notice that the function $\mu \mapsto g_{Z^\mu}^{-1}(\mu)$ is monotone (decreasing) and convex from $\mathbb{R}_+$ to $\mathbb{R}_+$.

Now, the following Berry-Esseen result holds:

**Theorem**

*Assume $H \in (\frac{1}{2}, 1)$. Then*

$$d_W\left(\frac{\sqrt{T}}{\sigma_{Z^\mu} g'_{Z^\mu}(f_{X^\mu})}(\widehat{\mu}_T - \mu), \mathcal{N}\right) \le C/\sqrt{T}.$$

*Also,*

$$d_W\left(\frac{\sqrt{T_n}}{\sigma_{Z^\mu} g'_{Z^\mu}(f_{X^\mu})}(\widetilde{\mu}_n - \mu), \mathcal{N}\right) \le C\left[n\Delta_n^{2H+1}\right]^{1/2} + C/\sqrt{n\Delta_n}.$$

Thank you for your attention !

# References


- Douissi, S., Es-Sebaiy, K., Viens, F. (2019). Berry-Esseen bounds for parameter estimation of general Gaussian processes. *ALEA, Lat. Am. J. Probab. Math. Stat.*, **16**, 633-664.
- El Onsy, B., Es-Sebaiy, K., Viens, F. (2017). Parameter Estimation for a partially observed Ornstein-Uhlenbeck process with long-memory noise. Stochastics, 89(2), 431-468.
- Es-Sebaiy, K., Viens, F. (2019). Optimal rates for parameter estimation of stationary Gaussian processes. Stochastic Processes and their Applications, 129(9), 3018-3054.
- Douissi, S., Es-Sebaiy, K, Alshahrani, F and Viens, F. AR(1) processes driven by second-chaos white noise: Berry-Essen bounds for quadratic variation and parameter estimation. In press in Stochastic Processes and their Applications (2020).