Maximum likelihood estimators and random walks in long memory models

Karine Bertin, Soledad Torres, Ciprian Tudor

To cite this version:

Karine Bertin, Soledad Torres, Ciprian Tudor. Maximum likelihood estimators and random walks in long memory models. To appear in "Statistics". 2009. <hal-00184942v2>

HAL Id: hal-00184942

https://hal.archives-ouvertes.fr/hal-00184942v2

Submitted on 18 Dec 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Maximum likelihood estimators and random walks in long memory models

Karine Bertin [1] Soledad Torres [1] Ciprian A. Tudor [2] *

[1] Departamento de Estadística, CIMFAV Universidad de Valparaíso,
Casilla 123-V, 4059 Valparaiso, Chile.
soledad.torres@uv.cl karine.bertin@uv.cl

[2]SAMOS/MATISSE, Centre d'Economie de La Sorbonne,
Université de Panthéon-Sorbonne Paris 1,
90, rue de Tolbiac, 75634 Paris Cedex 13, France.
tudor@univ-paris1.fr

December 18, 2009

Abstract

We consider statistical models driven by Gaussian and non-Gaussian self-similar processes with long memory and we construct maximum likelihood estimators (MLE) for the drift parameter. Our approach is based in the non-Gaussian case on the approximation by random walks of the driving noise. We study the asymptotic behavior of the estimators and we give some numerical simulations to illustrate our results.

2000 AMS Classification Numbers: 60G18, 62M99.

Key words: Fractional Brownian motion, Maximum likelihood estimation, Rosenblatt process, Random walk.

1 Introduction

The self-similarity property for a stochastic process means that scaling of time is equivalent to an appropriate scaling of space. That is, a process $(Y_t)_{t\geq 0}$ is selfsimilar of order $H > 0$ if for all $c > 0$ the processes $(Y_{ct})_{t\geq 0}$ and $(c^H Y_t)_{t\geq 0}$ have the same finite dimensional distributions. This property is crucial in applications such as network traffic analysis, mathematical finance,

*New affiliation since September 2009: Laboratoire Paul Painlevé, Université de Lille 1, F-59655 Villeneuve d'Ascq, France, email: tudor@math.univ-lille1.fr.

astrophysics, hydrology or image processing. We refer to the monographs [1], [4] or [10] for complete expositions on theoretical and practical aspects of self-similar stochastic processes.

The most popular self-similar process is the fractional Brownian motion (fBm). Its practical applications are notorious. This process is defined as a centered Gaussian process $(B_t^H)_{t\geq 0}$ with covariance function

$$R^H(t,s) := \mathbb{E}(B_t^H B_s^H) = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right), \quad t,s \geq 0.$$

It can be also defined as the only Gaussian self-similar process with stationary increments. Recently, this stochastic process has been widely studied from the stochastic calculus point of view as well as from the statistical analysis point of view. Various types of stochastic integrals with respect to it have been introduced and several types of stochastic differential equations driven by fBm have been considered (see e.g. [8], Section 5). Another example of a self-similar process still with long memory (but non-Gaussian) is the so-called Rosenblatt process which appears as limit in limit theorems for stationary sequences with a certain correlation function (see [2], [13]). Although it received a less important attention than the fractional Brownian motion, this process is still of interest in practical applications because of its self-similarity, stationarity of increments and long-range dependence. Actually the numerous uses of the fractional Brownian motion in practice (hydrology, telecommunications) are due to these properties; one prefers in general fBm before other processes because it is a Gaussian process and the calculus for it is easier; but in concrete situations when the Gaussian hypothesis is not plausible for the model, the Rosenblatt process may be an interesting alternative model. We mention also the work [14] for examples of the utilisation of non-Gaussian self-similar processes in practice.

The stochastic analysis of the fractional Brownian motion naturally led to the statistical inference for diffusion processes with fBm as driving noise. We will study in this paper the problem of the estimation of the drift parameter. Assume that we have the model

$$dX_t = \theta b(X_t)dt + dB_t^H, \quad t \in [0,T]$$

where $(B_t^H)_{t\in[0,T]}$ is a fractional Brownian motion with Hurst index $H \in (0,1)$, b is a deterministic function satisfying some regularity conditions and the parameter $\theta \in \mathbb{R}$ has to be estimated. Such questions have been recently treated in several papers (see [6], [16] or [12]): in general the techniques used to construct maximum likelihood estimators (MLE) for the drift parameter θ are based on Girsanov transforms for fractional Brownian motion and depend on the properties of the deterministic fractional operators related to the fBm. Generally speaking, the authors of these papers assume that the whole trajectory of the process is continuously observed. Another possibility is to use Euler-type approximations for the solution of the above equation and to construct a MLE estimator based on the density of the observations given "the past" (as in e.g. [9], Section 3.4, for the case of stochastic equations driven by the Brownian motion).

In this work our purpose is to make a first step in the direction of statistical inference for diffusion processes with self-similar, long memory and non-Gaussian driving noise. As far as we know, there are not many result on statistical inference for stochastic differential equations driven by non-Gaussian processes which in addition are not semimartingales. The basic example of a such process is the Rosenblatt process. We consider here the simple model

$$X_t = at + Z_t^H,$$

where $(Z_t^H)_{t\in[0,T]}$ is a Rosenblatt process with known self-similarity index $H \in (\frac{1}{2}, 1)$ (see Sections 3 and Appendix for the definition) and $a \in \mathbb{R}$ is the parameter to be estimated. We mention that, since this process is not a semimartingale, it is not Gaussian and its density function is not explicitly known, the techniques considered in the Gaussian case cannot be applied here. We therefore use a different approach: we consider an approximated model in which we replace the noise Z^H by a two-dimensional disturbed random walk $Z^{H,n}$ that, from a result is [15], converges weakly in the Skorohod topology to Z^Has $n \to \infty$. Note that this approximated model still keeps the main properties of the original model since the noise is asymptotically self-similar and it exhibits long range dependence. We then construct a MLE estimator (called sometimes in the literature, see e.g. [9] "pseudo-MLE estimator") using an Euler scheme method and we prove that this estimator is consistent. Although we have not martingales in the model, this construction involving random walks allows to use martingale arguments to obtain the asymptotic behavior of the estimators. Of course, this does not solve the problem of estimating a in the standard model defined above but we think that our approach represents a step into the direction of developing models driven by non-semimartingales and non-Gaussian noises.

Our paper is organized as follows. In Section 2 we recall some facts on the pseudo MLE estimators for the drift parameter in models driven by the standard Wiener process and by the fBm. We construct, in each model, estimators for the drift parameter and we prove their strong consistency (in the almost sure sense) or their L^2 consistency under the condition $\alpha > 1$ where N^α is the number of observations at our disposal and the step of the Euler scheme is $\frac{1}{N}$. This condition extends the usual hypothesis in the standard Wiener case (see 2, see also [9], paragraph 3.4). Section 3 is devoted to the study of the situation when the noise is the approximated Rosenblatt process; we construct again the estimator through an inductive method and we study its asymptotic behavior. The strong consistency is obtained under similar assumptions as in the Gaussian case. Section 4 contains some numerical simulations and in the Appendix we recall the stochastic integral representations for the fBm and for the Rosenblatt process.

2 Preliminaries

Let us start by recalling some known facts on maximum likelihood estimation in simple standard cases. Let $(W_t)_{t\in[0,T]}$ be a Wiener process on a classical Wiener space $(\Omega, \mathcal{F}, P)$ and let us consider the following simple model

$$Y_t = at + W_t, \qquad t \in [0,T] \tag{1}$$

with $T>0$ and assume that the parameter $a \in \mathbb{R}$ has to be estimated. One can for example use the Euler type discretization of (1)

$$Y_{t_{j+1}}^{(n)} := Y_{t_{j+1}} = Y_{t_j} + a\Delta t + W_{t_{j+1}} - W_{t_j}, \quad j = 0, \dots, N-1,$$

with $Y_{t_0} = Y_0 = 0$ and $\Delta t = t_{j+1} - t_j$ the step size of the partition. In the following, we denote $Y_{t_j} = Y_j$. In the following f_Z denotes the density of the random variable Z.

The conditional density of Y_{j+1} with respect to $Y_1, \dots Y_j$ is, by the Markov property, the same as the conditional density of Y_{j+1} with respect to Y_j and since $W_{t_{j+1}} - W_{t_j}$ has the normal law $N(0, \Delta t)$, this density can be expressed by

$$f_{Y_{j+1}/Y_j}(y_{j+1}/y_j) = \frac{1}{\sqrt{2\pi(\Delta t)}} \exp\left(-\frac{1}{2}\frac{(y_{j+1} - y_j - a\Delta t)^2}{\Delta t}\right).$$

We easily obtain the likelihood function of the observations $Y_1, \dots, Y_N$

$$L(a, y_1, \dots, y_N) = f_{Y_1}(y_1) \prod_{j=1}^{N-1} f_{Y_{j+1}/Y_j}(y_{j+1}/y_j) = \frac{1}{(2\pi(\Delta t))^{N/2}} \exp\left(-\frac{1}{2}\sum_{j=0}^{N-1} \frac{(y_{j+1} - y_j - a\Delta t)^2}{\Delta t}\right),$$

and this gives a maximum likelihood estimation of the form

$$\hat{a}_N = \frac{1}{N\Delta t} \sum_{j=0}^{N-1} (Y_{j+1} - Y_j),$$

and the difference $\hat{a}_N - a$ can be written as

$$\hat{a}_N - a = \frac{1}{N\Delta t} \sum_{j=0}^{N-1} (W_{t_{j+1}} - W_{t_j}).$$

Then

$$\mathbb{E}\, |\hat{a}_N - a|^2 = \frac{1}{N\Delta t},$$

and this converges to zero (that is, the estimator is L^2-consistent) if and only if

$$N\Delta t \to \infty, \quad N \to \infty. \tag{2}$$

Note that the partition $t_j = \frac{j}{N}$ with $j = 0, \dots N$ does not satisfy (2).

Remark 1 *Under condition (2) we get, by the strong law of large numbers, the almost sure convergence of the estimator $\hat{a}_N$ to the parameter a.*

Remark 2 *We need in conclusion to consider an interval between observation of the order $\Delta t = \frac{1}{N^\alpha}$ with $\alpha < 1$ to have (2). Equivalently, if we dispose on N^α observations with $\alpha > 1$, i.e. $T > N^{\alpha-1}$, and the interval Δt is of order $\frac{1}{N}$, condition (2) still holds. Using this fact, we will denote in the sequel by N^α the number of observations and we will use discretization of order $\frac{1}{N}$ of the model.*

Next, let us take a look to the situation when the Brownian motion W is replaced by a fractional Brownian motion B^H with Hurst parameter $H \in (0,1)$. As far as we know, this situation has not been considered in the literature. The references [6] or [16] uses a different approach, based on the observation of the whole trajectory of the process Y_t below. The model is now

$$Y_t = at + B_t^H, \qquad t \in [0,T], \tag{3}$$

and as before we aim to estimate the drift parameter a by assuming that H is known and on the basis on discrete observations $Y_1, \ldots, Y_{N^\alpha}$ (the condition on α will be clarified later). We use the same Euler type method with $t_j = \frac{j}{N}$ and we denote $Y_{t_j} = Y_j$.

We can easily found the following expression for the observations $Y_j, j = 1, \ldots, N^\alpha$,

$$Y_j = j\frac{a}{N} + B^H_{\frac{j}{N}}.$$

We need to compute the density of the vector $(Y_1, \ldots, Y_{N^\alpha})$. Since the covariance matrix of this vector is given by $\Gamma = (\Gamma_{i,j})_{i,j=1,\ldots,N^\alpha}$ with

$$\Gamma_{i,j} = Cov\left(B^H_{\frac{i}{N}} B^H_{\frac{j}{N}}\right),$$

the density of $(Y_1, \ldots, Y_{N^\alpha})$ will be given by

$$(2\pi)^{-\frac{N^\alpha}{2}} \frac{1}{\sqrt{\det \Gamma}} \exp\left(-\frac{1}{2}\left(y_1 - \frac{a}{N}, \ldots, y_{N^\alpha} - N^\alpha \frac{a}{N}\right)^t \Gamma^{-1} \left(y_1 - \frac{a}{N}, \ldots, y_{N^\alpha} - N^\alpha \frac{a}{N}\right)\right)$$

and by maximizing the above expression with respect to the variable a we obtain the following MLE estimator

$$\hat{a}_N = N \frac{\sum_{i,j=1}^{N^\alpha} j \Gamma^{-1}_{i,j} Y_i}{\sum_{i,j=1}^{N^\alpha} ij \Gamma^{-1}_{i,j}} \tag{4}$$

and then

$$\hat{a}_N - a = N \frac{\sum_{i,j=1}^{N^\alpha} j \Gamma^{-1}_{i,j} B^H_{\frac{i}{N}}}{\sum_{i,j=1}^{N^\alpha} ij \Gamma^{-1}_{i,j}}, \tag{5}$$

where the $\Gamma^{-1}_{i,j}$ are the coordinates of the matrix Γ^{-1}.

Thus

$$\begin{aligned}\mathbb{E}\left|\hat{a}_N - a\right|^2 &= N^2 \frac{\sum_{i,j,k,l=1}^{N^\alpha} jl\Gamma_{i,j}^{-1}\Gamma_{k,l}^{-1} E\left(B^H_{\frac{i}{N}} B^H_{\frac{k}{N}}\right)}{\left(\sum_{i,j=1}^{N^\alpha} ij\Gamma_{i,j}^{-1}\right)^2} \\ &= N^2 \frac{\sum_{i,j,k,l=1}^{N^\alpha} jl\Gamma_{i,j}^{-1}\Gamma_{k,l}^{-1}\Gamma_{i,k}}{\left(\sum_{i,j=1}^{N^\alpha} ij\Gamma_{i,j}^{-1}\right)^2}.\end{aligned}$$

Note that

$$\sum_{i,j,k,l=1}^{N^\alpha} jl\Gamma_{i,j}^{-1}\Gamma_{k,l}^{-1} E\left(B^H_{\frac{i}{N}} B^H_{\frac{k}{N}}\right) = \sum_{j,k,l=1}^{N^\alpha} jl\Gamma_{k,l}^{-1}\left(\sum_{i=1}^{N^\alpha}\Gamma_{i,j}^{-1}\Gamma_{i,k}\right) = \sum_{j,k,l=1}^{N^\alpha} jl\Gamma_{k,l}^{-1}\delta_{jk} = \sum_{j,l=1}^{N^\alpha} jl\Gamma_{j,l}^{-1}$$

and consequently

$$E\left|\hat{a}_N - a\right|^2 = N^2 \frac{1}{\sum_{i,j=1}^{N^\alpha} ij\Gamma_{ij}^{-1}} = N^{2-2H}\frac{1}{\sum_{i,j=1}^{N^\alpha} ijm_{ij}^{-1}}$$

where m_{ij}^{-1} are the coefficients of the matrix M^{-1} with $M = (m_{ij})_{i,j=1,\dots,N^\alpha}$, $m_{ij} = \frac{1}{2}(i^{2H} + j^{2H} - |i-j|^{2H})$. Let x be the vector $(1, 2, \dots, N^\alpha)$ in $\mathbb{R}^{N^\alpha}$. We use the inequality

$$x^t M^{-1} x \geq \frac{\|x\|_2^2}{\lambda}$$

where λ is the largest eigenvalue of the matrix M. Then we will have

$$\mathbb{E}\left|\hat{a}_N - a\right|^2 \leq N^{2-2H}\frac{\lambda}{\|x\|_2^2}.$$

Since $1^2 + 2^2 + \dots p^2 = \frac{p(p+1)(2p+1)}{6}$ we see that $\|x\|_2^2$ behaves as $N^{3\alpha}$ and on the other hand (by the Gersghorin Circle Theorem (see [5], Theorem 8.1.3, pag. 395)

$$\lambda \leq \max_{i=1,\dots,N^\alpha}\sum_{l=1}^{N^\alpha}|m_{il}| \leq CN^{\alpha(2H+1)},$$

with C a positive constant. Finally,

$$\mathbb{E}\left|\hat{a}_N - a\right|^2 \leq CN^{2-2H-3\alpha+\alpha(2H-1)} = CN^{(2-2H)(1-\alpha)} \tag{6}$$

and this goes to zero if and only if $\alpha > 1$.

Let us summarize the above discussion.

Proposition 1 *Let $(B_t^H)_{t\in[0,T]}$ be a fractional Brownian motion with Hurst parameter $H \in (0,1)$ and let $\alpha > 1$. Then the estimator (4) is L^p consistent for any $p \geq 1$.*

Proof: Since for every n the random variable $\hat{a}_n - a$ is a centered random variable, it holds that, for some positive constant depending on p

$$\mathbb{E}\left|\hat{a}_N - a\right|^p \leq c_p \left(\mathbb{E}\left|\hat{a}_N - a\right|^2\right)^{\frac{p}{2}} \leq c_p N^{p(1-H)(1-\alpha)}$$

and this converges to zero as zero since $\alpha > 1$. ■

It is also possible to obtain the almost sure convergence of the estimator to the the true parameter from the the estimate (6).

Proposition 2 *Let $(B^H_t)_{t\in[0,T]}$ be a fractional Brownian motion with Hurst parameter $H \in (0,1)$ and let $\alpha > 1$. Then the estimator (4) is strongly consistent, that is, $\hat{a}_n - a \to_{n\to\infty} p$ almost surely.*

Proof: Using Chebyshev's inequality,

$$P\left(\left|\hat{a}_n - a\right| > N^{-\beta}\right) \leq c_p N^{\beta p} N^{p(1-H)(1-\alpha)}.$$

In order to apply the Borel-Cantelli lemma, we need to find a strictly positive β such that $\sum_{N\geq 1} N^{\beta p} N^{p(1-H)(1-\alpha)} < \infty$. One needs $p\beta + (1-H)(1-\alpha)p < -1$ and this is possible if and only if $\alpha > 1$. ■

Remark 3 *The fact that the restriction $\alpha > 1$ is interesting because it justifies looking for another type of estimator for which no such restriction holds. But for the approach using the Euler scheme discretization, this restriction is somehow expected because it appear also also in the Wiener case (see [9]).*

Remark 4 *Let us also comment on the problem of estimation of the diffusion parameter in the model (3). Assume that the fractional Brownian motion B^H is replaced by σB^H in (3), with $\sigma \in \mathbb{R}$. In this case it is known that the sequence*

$$N^{2H-1} \sum_{i=0}^{N-1} \left(Y_{\frac{i+1}{N}} - Y_{\frac{i}{N}}\right)^2$$

converges (in L^2 and almost surely) to σ^2. Thus we easily obtain an estimator for the diffusion parameter by using such quadratic variations. The above sequence has the same behavior if we replace the fBm B^H by the Rosenblatt process Z^H because the Rosenblatt process is also self-similar with stationary increments and it still satisfies $\mathbb{E}\left|Z^H_t - Z^H_s\right|^2 = |t-s|^{2H}$ for every $s,t \in [0,T]$ (see Section 4 for details). For this reason we assume throughout this paper that the diffusion coefficient is equal to 1.

3 MLE and random walks in the non-Gaussian case

We study in this section a non-Gaussian long-memory model. The driving process is now a Rosenblatt process with selfsimilarity order $H \in (\frac{1}{2}, 1)$. This process appears as a limit in the so-called *Non Central Limit Theorem* (see [2] or [13]). It can be defined through its representation as double iterated integral with respect to a standard Wiener process given by equation (13) in the Appendix. Among its main properties, we recall

- it exhibits long-range dependence (the covariance function decays at a power function at zero)
- it is H-selfsimilar in the sense that for any $c > 0$, $(Z^H(ct)) =^{(d)} (c^H Z^H(t))$, where " $=^{(d)}$ " means equivalence of all finite dimensional distributions ; moreover, it has stationary increments, that is, the joint distribution of $(Z^H(t+h) - Z^H(h), t \in [0,T])$ is independent of $h > 0$.
- the covariance function is

$$\mathbb{E}(Z_t^H Z_s^H) = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right), \qquad s,t \in [0,T]$$

  and consequently, for every $s,t \in [0,T]$

$$\mathbb{E}\left|Z_t^H - Z_s^H\right|^2 = |t-s|^{2H}$$

- the Rosenblatt process is Hölder continuous of order $\delta < H$
- it is not a Gaussian process; in fact, it can be written as a double stochastic integral of a two-variable deterministic function with respect to the Wiener process.

Assume that we want to construct a MLE estimator for the drift parameter a in the model

$$Y_t = at + Z_t^H, \qquad t \in [0,T].$$

We first note that the Rosenblatt process has only been defined for the self-similarity order $H > \frac{1}{2}$ and consequently we will have in the sequel always $H \in (\frac{1}{2}, 1)$. The approaches used previously do not work anymore because the Rosenblatt process is still not a semimartingale and moreover, in contrast to the fBm model, its density function is not explicitly known anymore. The method based on random walks approximation offers a solution to the problem of estimating the drift parameter a. We will use this direction; that is, we will replace the process Z_H by its associated random walk

$$Z_t^{H,N} = \sum_{i,k=1;i\neq k}^{[Nt]} N^2 \int_{\frac{i-1}{N}}^{\frac{i}{N}} \int_{\frac{k-1}{N}}^{\frac{k}{N}} F\left(\frac{[Nt]}{N}, u, v\right) dvdu \frac{\xi_i}{\sqrt{N}} \frac{\xi_k}{\sqrt{N}}, \qquad t \in [0,T] \tag{7}$$

where the ξ_i are i.i.d. variables of mean zero and variance one and the deterministic kernel F is defined in Appendix by (14). It has been proved in [15] that the random walk (7) converges weakly in the Skorohod topology to the Rosenblatt process Z^H.

We consider the following discretization of the Rosenblatt process

$$(Z^{N,H}_{\frac{j}{N}}), \quad j = 0, \dots N^{\alpha},$$

where for $j \neq 1$ $Z^{H,N}_{\frac{j}{N}}$ is given by (7) and we set $Z^N_{\frac{1}{N}} = \xi_1 / N^H$. With this slight modification, the process $(Z^{H,N}_{\frac{j}{N}})$ still converges weakly in the Skorohod topology to the Rosenblatt process Z^H. We will assume as above that the variables ξ_i follows a standard normal law $N(0,1)$.

Concretely, we want to estimate the drift parameter a on the basis of the observations

$$Y_{t_{j+1}} = Y_{t_j} + a(t_{j+1} - t_j) + \left(Z^{H,N}_{t_{j+1}} - Z^{H,N}_{t_j} \right)$$

where $t_j = \frac{j}{N}$, $j = 0, \dots, N^{\alpha}$ and $Y_0 = 0$. We will assume again that we have at our disposal a number N^{α} of observations and we use a discretization of order $\frac{1}{N}$ of the model. Denoting $Y_j := Y_{t_j}$, we can write

$$Y_{j+1} = Y_j + \frac{a}{N} + \left(Z^{H,N}_{\frac{j+1}{N}} - Z^{H,N}_{\frac{j}{N}} \right), \qquad j = 0, \dots, N^{\alpha} - 1.$$

Now, we have

$$Z^{H,N}_{\frac{j+1}{N}} - Z^{H,N}_{\frac{j}{N}} \quad = \quad f_j(\xi_1, \dots, \xi_j) + g_j(\xi_1, \dots, \xi_j)\xi_{j+1}$$

where $f_0 = 0$, $f_1 = f_1(\xi_1) = -\xi_1 / N^H$ and for $j \geq 2$

$$f_j = f_j(\xi_1, \dots, \xi_j) = N \sum_{i,k=1; i \neq k}^{j} \left(\int_{\frac{i-1}{N}}^{\frac{i}{N}} \int_{\frac{k-1}{N}}^{\frac{k}{N}} \left(F\left(\frac{j+1}{N}, u, v \right) - F\left(\frac{j}{N}, u, v \right) \right) dvdu \right) \xi_i \xi_k,$$

and $g_0 = 1/N^H$ for $j \geq 1$

$$g_j = g_j(\xi_1, \dots, \xi_j) = 2N \sum_{i=1}^{j} \left(\int_{\frac{i-1}{N}}^{\frac{i}{N}} \int_{\frac{j}{N}}^{\frac{j+1}{N}} F\left(\frac{j+1}{N}, u, v \right) dvdu \right) \xi_i.$$

Finally we have the model

$$Y_{j+1} = Y_j + \frac{a}{N} + f_j + g_j \xi_{j+1}. \tag{8}$$

In the following, we assume that $(\xi_1, \dots, \xi_n, \dots) \in B$ where

$$B = \cap_{j \geq 1} \left\{ g_j(\xi_1, \dots, \xi_j) \neq 0 \right\}.$$

The event B satisfy $\mathbb{P}(B) = 1$.

Remark 5 *Note that on the event B, conditioning with respect to the $\xi_1, \cdots, \xi_j$ is the same as conditioning with respect to the $Y_1, \cdots, Y_j$. In fact, we have*

$$\begin{aligned} Y_1 &= a/N + \xi_1/N^H \\ Y_j &= Y_{j-1} + a/N + f_{-1}(\xi_1, \ldots, \xi_{j-1}) + \xi_j g_{j-1}(\xi_1, \ldots, \xi_{j-1}), \quad j \geq 2. \end{aligned}$$

Since on B, for all $j \geq 2$, $g_{j-1}(\xi_1, \ldots, \xi_{j-1}) \neq 0$, the two σ-algebra $\sigma(\xi_1, \ldots, \xi_{j-1})$ and $\sigma(Y_1, \ldots, Y_{j-1})$ satisfy

$$\sigma(\xi_1, \ldots, \xi_{j-1}) \cap B = \sigma(Y_1, \ldots, Y_{j-1}) \cap B.$$

Then, given $\xi_1, \ldots, \xi_j$ the random variable Y_{j+1} is conditionally Gaussian and the conditional density of Y_{j+1} given $Y_1, \ldots, Y_j$ can be written as

$$f_{Y_{j+1}/Y_1,\ldots,Y_j}(y_{j+1}/y_1, \ldots, y_j) = \frac{1}{\sqrt{2\pi g_j^2}} \exp\left(-\frac{1}{2} \frac{(y_{j+1} - y_j - a/N - f_j)^2}{g_j^2} \right).$$

The likelihood function can be expressed as

$$\begin{aligned} L(a, y_1, \ldots, y_{N^\alpha}) &= f_{Y_1}(y_1) f_{Y_2/Y_1}(y_2/y_1) \ldots f_{Y_{N^\alpha}/Y_1,\ldots,Y_{N^\alpha-1}}(y_{N^\alpha}/y_1, \ldots, y_{N^\alpha-1}) \\ &= \prod_{j=0}^{N^\alpha-1} \frac{1}{\sqrt{2\pi g_j^2}} \exp\left(-\frac{1}{2} \frac{(y_{j+1} - y_j - a/N - f_j)^2}{g_j^2} \right). \end{aligned}$$

By standard calculations, we will obtain

$$\hat{a}_N = \frac{N \sum_{j=0}^{N^\alpha-1} \frac{(Y_{j+1} - Y_j - f_j - a/N)}{g_j^2}}{\sum_{j=0}^{N^\alpha-1} \frac{1}{g_j^2}} \tag{9}$$

and since $Y_{j+1} - Y_j - f_j = \frac{a}{N} + g_j \xi_{j+1}$ we obtain

$$\hat{a}_N - a = \frac{N \sum_{j=0}^{N^\alpha-1} \frac{\xi_{j+1}}{g_j}}{\sum_{j=0}^{N^\alpha-1} \frac{1}{g_j^2}}.$$

Let us comment on the above expression. Firstly, note that since $\sum_{j=0}^{N^\alpha-1} \frac{1}{g_j^2}$ is not deterministic anymore, the square mean of the difference $\hat{a}_N - a$ cannot be directly computed. Secondly, if we denote again by

$$A_M = \sum_{j=0}^{M-1} \frac{\xi_{j+1}}{g_j}$$

this discrete process still satisfies

$$\mathbb{E}\left(A_{M+1}/\mathcal{F}_M\right) = A_M, \quad \forall M \geq 1$$

where $\mathcal{F}_M$ is the σ-algebra generated by $\xi_1, \ldots, \xi_M$. However we cannot speak about square integrable martingales brackets because A_M^2 is not integrable (recall that the bracket is defined in general for square integrable martingales). In fact, $\mathbb{E}(A_M^2) = \sum_{j=0}^M \mathbb{E}\left(\frac{1}{g_j^2}\right)$ and this is not finite in general because g_j is a normal random variable. We also mention that, in contrast to the Gaussian case, the expectation of the estimator is not easily calculable anymore to decide if (9) is unbiased. On the other hand, from the numerical simulation it seems that the estimator is biased.

Nevertheless, martingale type methods can be employed to deal with the estimator (9). We use again the notation

$$\langle A \rangle_M = \sum_{j=0}^{M-1} \frac{1}{g_j^2}.$$

The following lemma is crucial.

Lemma 1 *Assume that* $\alpha > 2 - 2H$ *and let us denote by*

$$T_N := \frac{1}{N^2} \sum_{j=0}^{N^\alpha - 1} \frac{1}{g_j^2}.$$

Then $T_N \xrightarrow{N\to\infty} \infty$ *almost surely. Denote*

$$U_N = \frac{1}{N^2} \left(\sum_{j=0}^{N^\alpha - 1} \frac{1}{g_j^2} \right)^{1-\gamma}.$$

Then there exists $0 < \gamma < 1$*, such that* $U_N \xrightarrow{N\to\infty} \infty$ *almost surely.*

Proof: Let prove the convergence for T_N. We will use a Borel-Cantelli argument. To this end, we will show that

$$\sum_{N \geq 1} \mathbb{P}\left(T_N \leq N^\delta\right) < \infty \tag{10}$$

for some $\delta > 0$.

Fix $0 < \delta < \alpha - (2 - 2H)$. We have

$$\begin{aligned} \mathbb{P}\left(T_N \leq N^\delta\right) &= \mathbb{P}\left(\sum_{j=0}^{N^\alpha - 1} \frac{1}{g_j^2} \leq N^{2+\delta} \right) \\ &= \mathbb{P}\left(\frac{1}{\sum_{j=0}^{N^\alpha - 1} \frac{1}{g_j^2}} \geq N^{-2-\delta} \right). \end{aligned}$$

We now choose a p integer large enough such that $\frac{1}{p} < \alpha - (2-2H) - \delta$ and we apply the Markov inequality. We can bound $\mathbb{P}\left(T_N \leq N^\delta\right)$ by

$$\mathbb{P}\left(T_N \leq N^\delta\right) \leq N^{(2+\delta)p} \mathbb{E}\left|\frac{1}{\sum_{j=0}^{N^\alpha-1} \frac{1}{g_j^2}}\right|^p$$

and using the inequality "harmonic mean is less than the arithmetic mean" we obtain

$$\mathbb{P}\left(T_N \leq N^\delta\right) \leq N^{(2+\delta)p} N^{-2\alpha p} \mathbb{E}\left(\sum_{j=0}^{N^\alpha-1} g_j^2\right)^p$$

Note that (the first inequality below can be argued similarly as in [11])

$$\frac{1}{N^{2H}} \geq \mathbb{E}\left|Z^{H,N}_{\frac{j+1}{N}} - Z^{H,N}_{\frac{j}{N}}\right|^2 = \mathbb{E}\left(f_j^2\right) + \mathbb{E}\left(g_j^2\right),$$

it holds that

$$\mathbb{E}\left(g_j^2\right) := c_j \leq \frac{1}{N^{2H}}.$$

Since $g_j = \sqrt{c_j} X_j$ where $X_j \sim N(0,1)$, by (3)

$$\begin{aligned}
\mathbb{P}\left(T_N \leq N^\delta\right) &\leq N^{(2+\delta-2\alpha-2H)p} \mathbb{E}\left(\sum_{j=0}^{N^\alpha-1} X_j^2\right)^p \\
&\leq N^{(2+\delta-2\alpha-2H)p} N^{\alpha p} = N^{p(2+\delta-\alpha-2H)}
\end{aligned}$$

and thus relation (10) is valid.

The convergence for U_N can be obtained in a similar way with $0 < \gamma < 1$ such that $\alpha(2-\gamma) - 2 + 2H\gamma > 0$, $0 < \delta < \alpha(2-\gamma) - 2 + 2H\gamma$ and p such that $1/p < \alpha(2-\gamma) - (2 - 2H\alpha) - \delta$. ■

We use the notation

$$V_M := \frac{A_M^2}{\langle A \rangle_M^{1+\gamma}}, \qquad M \geq 1$$

and

$$B_M := \frac{\langle A \rangle_{M+1} - \langle A \rangle_M}{\langle A \rangle_{M+1}^{1+\gamma}}.$$

Note that V_M and B_M are $\mathcal{F}_M$ adapted.

We recall the Robbins-Siegmund criterium for the almost sure convergence which will play an important role in the sequel (see e.g. [3], page 18): let $(V_N)_N, (B_N)_N$ be $\mathcal{F}_N$ adapted, positive sequences such that

$$\mathbb{E}\left(V_{N+1}/\mathcal{F}_N\right) \leq V_N + B_N \qquad \text{a.s. } .$$

Then the sequence of random variables $(V_N)_N$ converges as $N \to \infty$ to a random variable V_∞ almost surely on the set $\{\sum_{N\geq 1} B_N < \infty\}$.

Lemma 2 *The sequence V_M converges to a random variable almost surely when $M \to \infty$.*

Proof: We make use of the Robbins-Siegmund criterium. It holds that

$$\begin{aligned}
&\mathbb{E}\left(V_{M+1}/\mathcal{F}_M\right) = \mathbb{E}\left(\frac{A_{M+1}^2}{\langle A\rangle_{M+1}^{1+\gamma}}/\mathcal{F}_M\right)\\
&\leq \frac{1}{\langle A\rangle_{M+1}^{1+\gamma}}\mathbb{E}\left(A_{M+1}^2/\mathcal{F}_M\right) = \frac{1}{\langle A\rangle_{M+1}^{1+\gamma}}\left(A_M^2 + \langle A\rangle_{M+1} - \langle A\rangle_M\right)\\
&\leq V_M + B_M.
\end{aligned}$$

By Lemma 1 the sequence $\langle A\rangle_M$ converges to ∞ and therefore

$$\sum_N B_{N^\alpha} \leq C + \int_1^\infty x^{-1-\gamma}ds < \infty.$$

Once can conclude by applying the Robbins-Siegmund criterium. ■

We state now the main result of this section.

Theorem 1 *The estimator (9) is strongly consistent.*

Proof: We have

$$(\hat{a}_N - a)^2 = \frac{V_{N^\alpha}}{\frac{1}{N^2}\langle A\rangle_{N^\alpha}^{1-\gamma}} = \frac{V_{N^\alpha}}{U_N}$$

and we conclude by Lemmas 1 and 2. ■

Comment: The L^1 (or L^p) consistence of the estimator (9) is on open problem. Note that its expression is a fraction whose numerator and denominator are non-integrable involving inverses of Gaussian random variables. The natural approaches to deal with do not work. A first basic idea when is to try to apply the Hölder inequality for the product FG with $F = \sum_{j=0}^{N^\alpha-1}\frac{\xi_{j+1}}{g_j}$ and $G = \sum_{j=0}^{N^\alpha-1}\frac{1}{g_j^2}$. But when we do this, we are confronted with the moments of the random variable $|F|$ and it has not moments. Indeed, the second moment is $\mathbb{E}\sum_{j=0}^{N^\alpha-1}\frac{1}{g_j^2}$ which is obviously infinite. Even $\mathbb{E}|F|$ is infinite. Another way is to use the inequality $\left(\sum_{i=1}^{n}|a_i b_i|\right)^2 \leq \left(\sum_{i=1}^{n}|a_i|^2\right)\left(\sum_{i=1}^{n}|b_i|^2\right)$. In this way we avoid the non-integrability of the random variable F defined above. We get

$$\sum_{j=0}^{N^\alpha-1}\frac{\xi_{j+1}}{g_j} \leq \left(\sum_{j=0}^{N^\alpha-1}\xi_{j+1}^2\right)^{\frac{1}{2}}\left(\sum_{j=0}^{N^\alpha-1}\frac{1}{g_j^2}\right)^{\frac{1}{2}}$$

and then

$$|\hat{a}_N - a| \leq N\frac{\left(\sum_{j=0}^{N^\alpha-1}\xi_{j+1}^2\right)^{\frac{1}{2}}}{\left(\sum_{j=0}^{N^\alpha-1}\frac{1}{g_j^2}\right)^{\frac{1}{2}}}$$

and by saying that the harmonic mean is less that the arithmetic mean (we have the feeling that it is rather optimal in this case)

$$|\hat{a}_N - a| \leq N^{1-\alpha}\left(\sum_{j=0}^{N^\alpha-1}\xi_{j+1}^2\right)^{\frac{1}{2}}\left(\sum_{j=0}^{N^\alpha-1}g_j^2\right)^{\frac{1}{2}}. \tag{11}$$

We tried now to apply Hölder for $\mathbb{E}|\hat{a}_N - a|$ and this gives

$$\mathbb{E}|\hat{a}_N - a| \leq N^{1-\alpha}\left(\mathbb{E}\left(\sum_{j=0}^{N^\alpha-1}\xi_{j+1}^2\right)^{\frac{p}{2}}\right)^{1/p}\left(\mathbb{E}\left(\sum_{j=0}^{N^\alpha-1}g_j^2\right)^{\frac{q}{2}}\right)^{1/q}$$

But now $\sum_{j=0}^{N^\alpha-1}\xi_{j+1}^2$ is a chi-square variable with N^α degree of freedom, its p moment behaves as $N^{\alpha p}$ and it can be also proved that $\left(\mathbb{E}\left(\sum_{j=0}^{N^\alpha-1}g_j^2\right)^{\frac{q}{2}}\right)^{1/q}$ is of order less than $N^{-\alpha}2N^{-H}$ (in the paper it is proved in the case $q=2$) which unfortunately gives $\sum_{j=0}^{N^\alpha-1}\xi_{j+1}^2 \leq N^{1-H} \to \infty$. One needs to compute directly the norm of $|\hat{a}_N - a|$ in (11) by using the distribution of the vector $(g_1,\ldots,g_n)$. But it turns out that this joint distribution is complicated and not tractable because of the presence of the kernel of the Rosenblatt process.

4 Simulation

We consider the problem of estimating a from the observations $Y_1,\ldots,Y_N$ (8) driven by the approximated Rosenblatt process respectively. In all of the cases, we use $\alpha = 2$. We have implemented the estimator $\hat{a}_N$. We have simulated the observations $Y_1,\ldots,Y_{N^2}$ for different values of $H: 0.55, 0.75$ and 0.9 and the values of a: 2 and 20. We consider the cases $N=100$ and $N=200$, in others words in the first case we have 10000 observations and in the second case 40000. For each case, we calculate 100 estimation $\hat{a}_N$ and we give in the following tables the mean and the standard deviation of these estimation.

Finally in the model driven by the approximated Rosenblatt process (Section 3), we only construct an estimator that depends on the observations Y_j and of the ξ_j.

The results for $N = 100$ are the following.

$a = 2$	$H = 0.55$	$H = 0.75$	$H = 0.9$
mean	2.1860	2.1324	2.1416
stand. deviation	0.3757	0.3702	0.3393

$a = 20$	$H = 0.55$	$H = 0.75$	$H = 0.9$
mean	20.2643	20.2013	20.1506
stand. deviation	0.4272	0.4982	0.4507

The results for $N = 200$ are the following.

$a = 2$	$H = 0.55$	$H = 0.75$	$H = 0.9$
mean	2.0433	2.1048	2.0559
stand. deviation	0.2910	0.2795	0.2768

$a = 20$	$H = 0.55$	$H = 0.75$	$H = 0.9$
mean	20.1895	20.1080	20.0943
stand. deviation	0.1952	0.2131	0.2162

We can observe that the quality of estimation increases as N increases. We obtain these last estimations for the parameter of discretization $N = 200$. Since the computational cost is quite important, we do not implement the estimator for larger N. But although with $N = 200$, the estimator is quite good.

Appendix: Representation of fBm and Rosenblatt process as stochastic integral with respect to a Wiener process

The fractional Brownian process $(B_t^H)_{t\in[0,T]}$ with Hurst parameter $H \in (0,1)$ can be written (see e.g. [8], Chapter 5 or [7])

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

where $(W_t, t \in [0,T])$ is a standard Wiener process,

$$K^H(t,s) = c_H s^{\frac{1}{2}-H} \int_s^t (u-s)^{H-\frac{3}{2}} u^{H-\frac{1}{2}} du \tag{12}$$

where $t > s$ and $c_H = \left(\frac{H(2H-1)}{\beta(2-2H, H-\frac{1}{2})}\right)^{\frac{1}{2}}$ and $\beta(\cdot,\cdot)$ is the beta function. For $t > s$, we have

$$\frac{\partial K^H}{\partial t}(t,s) = c_H \left(\frac{s}{t}\right)^{\frac{1}{2}-H} (t-s)^{H-\frac{3}{2}}.$$

An analogous representation for the Rosenblatt process $(Z_t^H)_{t\in[0,T]}$ is (see [17])

$$Z_t^H = \int_0^t \int_0^t F(t, y_1, y_2) dW_{y_1} dW_{y_2} \tag{13}$$

where $(W_t, t \in [0,T])$ is a Brownian motion,

$$F(t, y_1, y_2) = d(H) 1_{[0,t]}(y_1) 1_{[0,t]}(y_2) \int_{y_1 \vee y_2}^t \frac{\partial K^{H'}}{\partial u}(u, y_1) \frac{\partial K^{H'}}{\partial u}(u, y_2) du, \tag{14}$$

$H' = \frac{H+1}{2}$ and $d(H) = \frac{1}{H+1} \left(\frac{H}{2(2H-1)} \right)^{-\frac{1}{2}}$. Actually the representation (13) should be understood as a multiple integral of order 2 with respect to the Wiener process W. We refer to [8] for the construction and the basic properties of multiple stochastic integrals.

Aknowledgement This work has been partially supported by project by the Laboratory ANESTOC PBCT ACT-13. Karine Bertin has been partially supported by Project FONDECYT 1090285. Soledad Torres has been partially supported by Project FONDECYT 1070919.

References

[1] J. Beran (1994): *Statistics for Long-Memory Processes.* Chapman and Hall.

[2] R.L. Dobrushin and P. Major (1979): *Non-central limit theorems for non-linear functionals of Gaussian fields.* Z. Wahrscheinlichkeitstheorie verw. Gebiete, 50, pag. 27-52.

[3] M. Duflo (1997): *Random iterative models.* Springer.

[4] P. Embrechts and M. Maejima (2002): *Selfsimilar processes.* Princeton University Press, Princeton, New York.

[5] G.H. Golub and C.F. van Loan (1989): *Matrix computations.* Hopkins University Press.

[6] M. Kleptsyna and A. Le Breton (2002): *Statistical Analysis of the fractional Ornstein-Uhlenbeck process.* Statistical Inference for Stochastic Processes, 5, pag. 229-248.

[7] D. Nualart (2003): *Stochastic calculus with respect to the fractional Brownian motion and applications.* Contemp. Math. 336, , pag. 3-39.

[8] D. Nualart (2006): Malliavin Calculus and Related Topics. Second Edition. Springer.

[9] B.L.S. Prakasa Rao (1999): *Statistical Inference for Diffusion Type Processes.* Arnold, HodderHeadlone Group, London.

[10] G. Samorodnitsky and M. Taqqu (1994): *Stable Non-Gaussian random variables.* Chapman and Hall, London.

[11] T. Sottinen (2001): *Fractional Brownian motion, random walks and binary market models.* Finance and Stochastics, 5, pag. 343-355.

[12] T. Sottinen and C.A. Tudor (2007): *Parameter estimation for stochastic equations with additive fractional Brownian sheet.* Statistical Inference for Stochastic Processes, 11(3), 221–236.

[13] M. Taqqu (1975): *Weak convergence to the fractional Brownian motion and to the Rosenblatt process.* Z. Wahrscheinlichkeitstheorie verw. Gebiete, 31, pag. 287-302.

[14] M. Taqqu (1978): *A representation for self-similar processes.* Stochastic Process. Appl. 7, no. 1, pag. 55–64.

[15] S. Torres and C.A. Tudor (2007): *Donsker type theorem for the Rosenblatt process and a binary market model.* Stochastic Analysis and Applications, 27(3), pag. 555-573.

[16] C.A. Tudor and F. Viens (2007): *Statistical Aspects of the fractional stochastic calculus.* The Annals of Statistics, 35(3), pag. 1183-1212.

[17] C.A. Tudor (2008): *Analysis of the Rosenblatt process.* ESAIM-Probability and Statistics, 12, pag. 230-257.