



HAL
open science

Maximum likelihood estimators and random walks in long memory models

Karine Bertin, Soledad Torres, Ciprian A. Tudor

► **To cite this version:**

Karine Bertin, Soledad Torres, Ciprian A. Tudor. Maximum likelihood estimators and random walks in long memory models. 2007. hal-00184942v1

HAL Id: hal-00184942

<https://hal.science/hal-00184942v1>

Preprint submitted on 4 Nov 2007 (v1), last revised 18 Dec 2009 (v2)

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

²SAMOS/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

November 4, 2007

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 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 is an invariance property after time scaling met in various topics of both theoretical and applied sciences. The long range dependence (or long memory) property means that in a model the correlations between the observations that are far apart decay to zero at a slower rate that one would expect from the independent data. These two properties, that appear many times together in practice, are crucial in applications such as network traffic analysis, mathematical finance, astrophysics, hydrology or image processing. We refer to the

monographs [1], [4] or [8] for complete expositions on theoretical and practical aspects of self-similar long-memory processes.

In the last years the study, from the stochastic calculus point of view, of such processes has been intensively developed. The most popular self-similar stochastic process that exhibits long-range dependence is of course the fractional Brownian motion (fBm). Recently, stochastic integrals of various types with respect to the fBm have been constructed and stochastic differential equations driven by the fBm have been considered (see e.g. [6]). Another example of a self-similar process 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], [11]). 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 very large utilization 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 Gaussian and to calculus for it is easier; but in concrete situations when the gaussianity hypothesis is not plausible for the model, one can use for example the Rosenblatt process.

The stochastic analysis of the fractional Brownian motion led to the statistical inference for diffusion processes with fBm as driving noise. We refer here to 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 $H \in (0, 1)$ and b is a deterministic function satisfying some regularity conditions and assume that the parameter $\theta \in \mathbb{R}$ has to be estimated. Such questions have been recently treated in several papers (see [5] for the case $H \in (\frac{1}{2}, 1)$ and b linear or [13] for the general case or [10] for the two-parameter case): 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 depends on the properties of the deterministic fractional operators related to the fBm. 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. [7], Section 3.4.

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 4 and Appendix for the definition) and $a \in \mathbb{R}$ is the parameter to be estimated. We mention that, since the process is not a semimartingale, it is not Gaussian and its density

function is not explicitly known, the techniques used 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 in [12], converges weakly in the Skorohod topology to Z^H as $n \rightarrow \infty$. 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 pseudo-MLE estimator) using an Euler scheme method (both in Gaussian and non-Gaussian cases) and we prove that this estimator is strongly 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.

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. In Section 3 we introduce a statistical model driven by a disturbed random walk that converges weakly to the fBm. We construct an estimator for the drift parameter and we prove its strong consistency (in the almost sure sense) and its L^2 consistency under the condition

$$\alpha > 2 - 2H$$

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 [7], paragraph 3.4). Section 4 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 the same assumptions as in the Gaussian case. Section 5 contains some numerical simulations to illustrate our main result and in 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 (Ω, \mathcal{F}, P) and let us consider the following simple model

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

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

$$N\Delta t \rightarrow \infty, \quad N \rightarrow \infty. \quad (2)$$

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 $H \in (\frac{1}{2}, 1)$. The model is now

$$Y_t = at + B_t^H, \quad t \in [0, T], \quad (3)$$

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, \dots, Y_{N^α} (the condition on α will be clarified later). We use the same Euler type method with $t_j = \frac{j}{N}$ and we denote $Y_{t_j} = Y_j$. Since $B_{t_{j+1}}^H - B_{t_j}^H \sim N(0, (\Delta t)^{2H})$ we have

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

and we obtain the expression for the MLE

$$\hat{a}_N = \frac{1}{N^\alpha \Delta t} \sum_{j=0}^{N^\alpha-1} (Y_{j+1} - Y_j) \quad (4)$$

with

$$\hat{a}_N - a = \frac{N}{N^\alpha} \sum_{j=0}^{N^\alpha-1} (B_{(j+1)/N}^H - B_{j/N}^H) \quad (5)$$

and thus

$$\begin{aligned} |\hat{a}_N - a|^2 &= N^{2-2\alpha} \sum_{i,j=0}^{N^\alpha-1} \mathbb{E} \left((B_{(i+1)/N}^H - B_{i/N}^H)(B_{(j+1)/N}^H - B_{j/N}^H) \right) \\ &= N^{2-2H-2\alpha} + 2N^{2-2\alpha} N^{-2H} \sum_{i < j} (|i-j+1|^{2H} + |i-j-1|^{2H} - 2|i-j|^{2H}) \\ &= N^{2-2H-2\alpha} + 2N^{2-2\alpha} N^{-2H} \sum_{k=1}^{N^\alpha} (N^\alpha - k) (|k+1|^{2H} + |k-1|^{2H} - 2k^{2H}). \end{aligned}$$

A simple study shows that the last sum $\sum_{k=1}^{N^\alpha} (N^\alpha - k) (|k+1|^{2H} + |k-1|^{2H} - 2k^{2H})$ behaves as $N \rightarrow \infty$ as $N^{2H\alpha}$ and therefore the estimator \hat{a}_N is L^2 consistent if and only if

$$\alpha > 1.$$

Remark 3 We would like to mention that, since the increments of the process B^H are not independent anymore, the almost sure convergent to zero of the difference (5) cannot be obtained directly by the strong law of large numbers. Moreover, since the process B^H is not a semimartingale, the martingale type techniques cannot be used to study the estimator \hat{a}_N . This problem will be avoided by the use of the random walks that approximate B^H .

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 $\mathbb{E} |Z_t^H - Z_s^H|^2 = |t - s|^{2H}$ for every $s, t \in [0, T]$ (see Section 4 for details). Therefore we will assume throughout this paper that the diffusion coefficient is equal to 1.

3 MLE based on random walks: the Gaussian case

As we said, the form of the estimator (4) does not allow to apply martingale approaches to treat this estimator. We therefore propose a new model: we replace in (3) the fractional Brownian motion B^H by the its associated disturbed random walk

$$B_t^{H,N} = \sum_{i=1}^{[Nt]} \sqrt{N} \left(\int_{\frac{i-1}{N}}^{\frac{i}{N}} K^H \left(\frac{[Nt]}{N}, s \right) ds \right) \xi_i, \quad t \in [0, T]$$

where the ξ_i are i.i.d. random variables with zero-mean and variance equal to 1 and K^H is defined in Appendix by (15). It has been proved in [9] that the sequence $B^{H,N}$ converges weakly as $N \rightarrow \infty$ in the Skorohod topology to the fBm. Our new model, with $B^{H,N}$ replacing B^H still keeps the main properties of the original process: it is a long range dependence model and the distribution of the noise is asymptotically selfsimilar. We have a non-semimartingale model but as we will see later, the martingales can be used to treat this model. Moreover, the model can be extended to non-Gaussian case when the driven process is the Rosenblatt process (Section 4).

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

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

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

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

Now, since

$$B_{\frac{j+1}{N}}^{H,N} - B_{\frac{j}{N}}^{H,N} = f_j(\xi_1, \dots, \xi_j) + F_j \xi_{j+1}$$

where

$$f_j(\xi_1, \dots, \xi_j) = f_j = \sqrt{N} \sum_{i=1}^j \int_{\frac{i-1}{N}}^{\frac{i}{N}} \left(K^H \left(\frac{j+1}{N}, s \right) - K^H \left(\frac{j}{N}, s \right) \right) ds \xi_i$$

and

$$F_j = \sqrt{N} \int_{\frac{j}{N}}^{\frac{j+1}{N}} K \left(\frac{j+1}{N}, s \right) ds,$$

we obtain

$$Y_{j+1} = Y_j + \frac{a}{N} + f_j(\xi_1, \dots, \xi_j) + F_j \xi_{j+1}. \quad (6)$$

Let us denote by \mathcal{F}_j the σ -algebra generated by the random variables Y_1, \dots, Y_j . From relation (6) it is clear that \mathcal{F}_j coincides with the σ -algebra generated by ξ_1, \dots, ξ_j for every j .

From now on, we will assume that the r.v. ξ_i follows a standard normal law $N(0, 1)$. Then, given ξ_1, \dots, ξ_j the random variable Y_{j+1} is conditionally Gaussian and the conditional density of Y_{j+1} given Y_1, \dots, Y_j can be written as

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

The likelihood function can be expressed as

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

This leads to the expression of the MLE

$$\hat{a}_N = N \frac{\sum_{j=0}^{N^\alpha-1} \frac{(Y_{j+1} - Y_j - f_j)}{F_j^2}}{\sum_{j=0}^{N^\alpha-1} \frac{1}{F_j^2}} \quad (7)$$

with

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

Remark 5 Clearly the estimator (7) is unbiased.

Let's study the consistency of (7).

Proposition 1 Assume that $\alpha > 2 - 2H$. Then the estimator \hat{a}_N is L^2 -consistent, that is,

$$\mathbb{E} |\hat{a}_N - a|^2 \xrightarrow{N \rightarrow \infty} 0.$$

Proof: By the independence of ξ_i , we can write

$$\begin{aligned} \mathbb{E} |\hat{a}_N - a|^2 &= N^2 \mathbb{E} \left(\frac{\sum_{j=0}^{N^\alpha-1} \frac{\xi_{j+1}}{F_j}}{\sum_{j=0}^{N^\alpha-1} \frac{1}{F_j^2}} \right)^2 \\ &= N^2 \frac{1}{\sum_{j=0}^{N^\alpha-1} \frac{1}{F_j^2}}. \end{aligned}$$

We will need the following bound:

$$\mathbb{E} \left| B_{\frac{i+1}{N}}^{H,N} - B_{\frac{j}{N}}^{H,N} \right|^2 \leq \frac{1}{N^{2H}}. \quad (8)$$

Indeed, we can write

$$\begin{aligned} B_{\frac{i+1}{N}}^{H,N} - B_{\frac{j}{N}}^{H,N} &= \sum_{i=1}^j \sqrt{N} \left(\int_{\frac{i-1}{N}}^{\frac{i}{N}} \left(K^H\left(\frac{j+1}{N}\right) - K^H\left(\frac{j}{N}\right) \right) ds \right) \xi_i \\ &\quad + \sqrt{N} \left(\int_{\frac{j}{N}}^{\frac{j+1}{N}} K^H\left(\frac{j+1}{N}, s\right) ds \right) \xi_{j+1} \end{aligned}$$

and thus

$$\begin{aligned} &\mathbb{E} \left| B_{\frac{i+1}{N}}^{H,N} - B_{\frac{j}{N}}^{H,N} \right|^2 \\ &= \sum_{i=1}^j N \left(\int_{\frac{i-1}{N}}^{\frac{i}{N}} \left(K^H\left(\frac{j+1}{N}\right) - K^H\left(\frac{j}{N}\right) \right) ds \right)^2 + N \left(\int_{\frac{j}{N}}^{\frac{j+1}{N}} K^H\left(\frac{j+1}{N}, s\right) ds \right)^2 \\ &\leq \sum_{i=1}^j \int_{\frac{i-1}{N}}^{\frac{i}{N}} \left(K^H\left(\frac{j+1}{N}\right) - K^H\left(\frac{j}{N}\right) \right)^2 ds + \int_{\frac{j}{N}}^{\frac{j+1}{N}} K^H\left(\frac{j+1}{N}, s\right)^2 ds \\ &= \mathbb{E} \left| B_{\frac{i+1}{N}}^H - B_{\frac{j}{N}}^H \right|^2 = \frac{1}{N^{2H}} \end{aligned}$$

where the second equality above can be seen by using the integral representation of the fBm (see Appendix).

On the other hand, note that

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

and then, for every $j = 0, \dots, N^\alpha - 1$

$$F_j^2 \leq \frac{1}{N^{2H}}. \quad (9)$$

By (9), we obtain that

$$\mathbb{E} |\hat{a}_N - a|^2 \leq N^{2-\alpha-2H}$$

and this converges to zero if $\alpha > 2 - 2H$. ■

Remark 6 *Note that for the use of the approximated model instead of the continuous time model (3) we need a smaller number of observations because $2 - 2H < 1$.*

We introduce some notation. Set

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

and note that $(A_M)_{M \geq 1}$ is a discrete martingale with respect to the filtration $\mathcal{F}_M = \sigma(\xi_1, \dots, \xi_M)$ whose bracket is

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

We can therefore apply martingale results to treat the convergence a.s. of \hat{a}_N to the parameter a . We recall for instance the Robbins-Siegmund criterium for the almost sure convergence (see e.g. [3], page 18): let $(V_N)_N, (B_N)_N$ be \mathcal{F}_N adapted, positive sequences such that

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

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

Lemma 1 *Let $\alpha > 2 - 2H$. Then the sequence*

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

converges to ∞ for some $0 < \gamma < 1$.

Proof: By (9) we get

$$\frac{1}{N^2} \left(\sum_{j=0}^{N^\alpha-1} \frac{1}{F_j^2} \right)^{1-\gamma} \geq N^{(\alpha+2H)(1-\gamma)-2}$$

and this converges to ∞ for $\alpha > \frac{2}{1-\gamma} - 2H$. Then one can find $0 < \gamma < 1$ close enough to zero to have this relation true. \blacksquare

We will regard now the almost sure convergence of the estimator (7).

Proposition 2 *Assume that $\alpha > 2 - 2H$. Then the estimator \hat{a}_N is strongly consistent, that is $\hat{a}_N \rightarrow a$ a.s. as $N \rightarrow \infty$.*

Proof: We denote for $0 < \gamma < 1$

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

We can write, for every $0 < \gamma < 1$ and if $M = N^\alpha$

$$(\hat{a}_N - a)^2 = \frac{V_M}{\frac{1}{N^2} \langle A \rangle_M^{1-\gamma}}. \quad (10)$$

Since

$$\mathbb{E}(A_{M+1}^2 / \mathcal{F}_M) = A_M^2 + \langle A \rangle_{M+1} - \langle A \rangle_M,$$

we have (note that $\langle A \rangle_{M+1}$ is \mathcal{F}_M measurable)

$$\mathbb{E}(V_{M+1} / \mathcal{F}_M) \leq V_M + B_M$$

where we denote by

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

Since $\langle A \rangle_M$ is a positive increasing sequence that converges to ∞ (cf. Lemma 1), we deduce that

$$\sum_{M \geq 1} B_M \leq C + \int_1^\infty x^{-1-\gamma} ds < \infty,$$

where C is a constant. The Robbins-Siegmund criterium will imply the almost sure convergence of V_M to a random variable V_∞ . Then applying Lemma 1 in (10), we deduce the almost sure convergence of \hat{a}_N to a . \blacksquare

4 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 [11]). It can be defined through its representation as double iterated integral with respect to a standard Wiener process given by equation (16) in the Appendix. Among its main properties, we recall

- it exhibits long-range dependence (the covariance function decay at a power function at zero)
- it is H -selfsimilar in the sense that for any $c > 0$, $(Z^H(ct)) \stackrel{(d)}{=} (c^H Z^H(t))$, where " $\stackrel{(d)}{=}$ " 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

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

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

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

- 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

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

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

$$Z_t^{H,N} = \sum_{i,k=1; i \neq k}^{[Nt]} N^2 \int_{\frac{i-1}{N}}^{\frac{i}{N}} \int_{\frac{k-1}{N}}^{\frac{k}{N}} F\left(\frac{[Nt]}{N}, u, v\right) dv du \frac{\xi_i}{\sqrt{N}} \frac{\xi_k}{\sqrt{N}}, \quad t \in [0, T] \quad (11)$$

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 (17). It has been proved in [12] that the random walk (11) converges weakly in the Skorohod topology to the Rosenblatt process Z^H .

We consider the following discretization of the Rosenblatt process

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

where for $j \neq 1$ $Z_{\frac{j}{N}}^{H,N}$ is given by (11) and we set $Z_{\frac{1}{N}}^N = \xi_1/N^H$. With this slight modification, the process $(Z_{\frac{j}{N}}^{H,N})$ 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

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

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

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

Now, we have

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

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

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

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

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

Finally we have the model

$$Y_{j+1} = Y_j + \frac{a}{N} + f_j + g_j \xi_{j+1}. \quad (12)$$

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

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

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

Remark 7 Note that on the event B , conditioning with respect to the ξ_1, \dots, ξ_j is the same as conditioning with respect to the Y_1, \dots, Y_j . In fact, we have

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

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

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

Then, given ξ_1, \dots, ξ_j the random variable Y_{j+1} is conditionally Gaussian and the conditional density of Y_{j+1} given Y_1, \dots, Y_j can be written as

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

The likelihood function can be expressed as

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

By standard calculations, we will obtain

$$\hat{a}_N = \frac{N \sum_{j=0}^{N^\alpha-1} \frac{(Y_{j+1} - Y_j - f_j - a/N)}{g_j^2}}{\sum_{j=0}^{N^\alpha-1} \frac{1}{g_j^2}} \quad (13)$$

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

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

where \mathcal{F}_M is the σ -algebra generated by ξ_1, \dots, ξ_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, $\mathbb{E}(A_M^2) = \sum_{j=0}^{M-1} \mathbb{E}\left(\frac{1}{g_j^2}\right)$ 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

(13) 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 (13).

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 2 *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 $T_N \xrightarrow{N \rightarrow \infty} \infty$ 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 $U_N \xrightarrow{N \rightarrow \infty} \infty$ almost surely.

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

$$\sum_{N \geq 1} \mathbb{P} \left(T_N \leq N^\delta \right) < \infty \tag{14}$$

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 the bound (8))

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

it holds that

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

Since $g_j = \sqrt{c_j} X_j$ where $X_j \sim N(0, 1)$, by (4)

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

and thus relation (14) 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

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

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.

Lemma 3 *The sequence V_M converges to a random variable almost surely when $M \rightarrow \infty$.*

Proof: As in the proof of Proposition 2 we make use of the Robbins-Siegmund criterium. It holds that

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

By Lemma 2 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 (13) 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 2 and 3. ■

5 Simulation

We consider the problem of estimating a from the observations Y_1, \dots, Y_N in the model (12) driven by the approximated Rosenblatt noise. We have implemented the estimator \hat{a}_N . We have simulated the observations Y_1, \dots, Y_N for different values of H : 0.55, 0.75 and 0.9 and for the parameter of discretization $N = 100$ and the values of a : 2 and 20. For each case, we calculate 200 estimation \hat{a}_N and we give in the following tables the mean and the standard deviation of these estimation.

$a = 2$	$H = 0.55$	$H = 0.75$	$H = 0.9$
mean	2.23	2.24	2.14
stand. deviation	0.36	0.39	0.21

$a = 20$	$H = 0.55$	$H = 0.75$	$H = 0.9$
mean	20.29	20.25	20.11
stand. deviation	0.43	0.40	0.20

We obtain the estimations for the parameter of discretization $N = 100$ because the computational cost is quite important for larger N . But although with $N = 100$, 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

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

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

$$K^H(t, s) = c_H s^{\frac{1}{2}-H} \int_s^t (u-s)^{H-\frac{3}{2}} u^{H-\frac{1}{2}} du \quad (15)$$

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 [14])

$$Z_t^H = \int_0^t \int_0^t F(t, y_1, y_2) dW_{y_1} dW_{y_2} \quad (16)$$

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

$$F(t, y_1, y_2) = d(H) 1_{[0, t]}(y_1) 1_{[0, t]}(y_2) \int_{y_1 \vee y_2}^t \frac{\partial K^{H'}}{\partial u}(u, y_1) \frac{\partial K^{H'}}{\partial u}(u, y_2) du, \quad (17)$$

$$H' = \frac{H+1}{2} \text{ and } d(H) = \frac{1}{H+1} \left(\frac{H}{2(2H-1)} \right)^{-\frac{1}{2}}.$$

Acknowledgement This work has been partially supported by project by the Laboratory ANESTOC ACT-13. Karine Bertin has been partially supported by Project DIPUV 33/05 and Project FONDECYT 1061184. Soledad Torres has been partially supported by Project 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] M. Kleptsyna and A. Le Breton (2002): *Statistical Analysis of the fractional Ornstein-Uhlenbeck process*. Statistical Inference for stochastic processes, 5, pp. 229-248.
- [6] D. Nualart (2003): *Stochastic calculus with respect to the fractional Brownian motion and applications*. Contemp. Math. 336, , pag. 3-39.

- [7] B.L.S. Prakasa Rao (1999): *Statistical Inference for Diffusion Type Processes*. Arnold, HodderHeadlone Group, London.
- [8] G. Samorodnitsky and M. Taqqu (1994): *Stable Non-Gaussian random variables*. Chapman and Hall, London.
- [9] T. Sottinen (2001): *Fractional Brownian motion, random walks and binary market models*. Finance and Stochastics, 5, pag. 343-355.
- [10] T. Sottinen and C.A. Tudor (2007): *Parameter estimation for stochastic equations with additive fractional Brownian sheet*. Statistical Inference for Stochastic Processes, to appear.
- [11] M. Taqqu (1975): *Weak convergence to the fractional Brownian motion and to the Rosenblatt process*. Z. Wahrscheinlichkeitstheorie verw. Gebiete, 31, pag. 287-302.
- [12] S. Torres and C.A. Tudor (2007): *Donsker type theorem for the Rosenblatt process and a binary market model*. Preprint.
- [13] C.A. Tudor and F. Viens (2007): *Statistical Aspects of the fractional stochastic calculus*. The Annals of Statistics, 35(3), pag. 1183-1212.
- [14] C.A. Tudor (2007): *Analysis of the Rosenblatt process*. ESAIM-Probability and Statistics, to appear.