A wavelet based estimator for the parameter of self-similarity of fractional Brownian motion

Lieve Delbeke and Walter Van Assche
Katholieke Universiteit Leuven
Celestijnenlaan 200 B
B-3001 Heverlee (BELGIUM)

January 25, 1996; revised July 15, 1996

Abstract

Let $X = \{X(t), t \in \mathbb{R}\}$ be a continuous fractional Brownian motion with parameter of self-similarity H $(0 < H < 1)$. Let ψ be a wavelet function with compact support and let $\psi_{j,k}$ be rescaled versions of the function ψ at position k, where $k \in \mathbb{N}_0$. We investigate the basic properties of the estimator for H

$$\hat{H}_n^k = \sum_{j=0}^{n-1} w_{j,n} \log_2 |D_{j,k}| - \frac{1}{2},$$

where $D_{j,k} = \int_R X(t)\psi_{j,k}(t)\,dt$ are wavelet coefficients, $w_{j,n}$ are certain weights, and n is the number of scales used. This estimator is a weighted sum of random variables $(\log_2 |D_{j,k}|)_{j\in\mathbb{Z}}$ $(k \in \mathbb{N}_0$ fixed), which form a stationary, strongly mixing sequence. We show that the estimator is unbiased, consistent and has asymptotically a normal distribution.

Keywords and phrases: self-similar processes, fractional Brownian motion, wavelets.

1 Introduction

Self-similar processes are stochastic processes which are invariant in distribution under judicious scaling of time and scale. They are used in probability theory because of their connection to limit theorems and they are frequently used in stochastic modeling. Some aspects of self-similarity appear for example in economics, geophysics, turbulence and astronomy. A recent example of a time series from astronomy which contains a structure of self-similarity can be found in [18]. The best known self-similar process is Brownian motion.

The parameter of self-similarity, or scaling coefficient, is a positive number denoted by H. A stochastic process $X = \{X(t), t \in I\!R\}$ is self-similar with index H if, for any $c > 0$, the finite-dimensional distributions of $\{X(ct), t \in I\!R\}$ are the same as those of $\{c^H X(t), t \in I\!R\}$. For a given $H \in (0,1)$, there is basically one single Gaussian process that is self-similar with scaling coefficient H and has stationary increments, namely fractional Brownian motion (see [17, p. 318]). Brownian motion itself has $H = 1/2$.

Fix $0 < H < 1$, then fractional Brownian motion is defined as the Gaussian process with mean zero and autocovariance function

$$\begin{aligned} R_H(t_1, t_2) &= \mathrm{E}[X(t_1)X(t_2)] \\ &= \frac{1}{2}(|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H})\mathrm{Var}X(1). \end{aligned}$$

This process has the property $X(0) = 0$ almost surely.

To get knowledge of the scaling coefficient of a time series which has self-similarity structure, we split up the process in its scale components. To do this, we use a wavelet function ψ and its rescaled versions $\psi_{j,k}$ at position k, where $k \in I\!N_0$. The functions $\psi_{j,k} : I\!R \to I\!R$ $(j, k \in Z\!\!\!Z)$ are defined as follows:

$$\psi_{j,k}(t) = 2^{-\frac{j}{2}}\psi(2^{-j}t - k).$$

(See, e.g. [6] and [7].) We suppose that the wavelet ψ has compact support. Define

$$D_{j,k} = \int_{I\!R} X(t)\psi_{j,k}(t)\,dt \qquad j, k \in Z\!\!\!Z,$$

and the weights

$$w_{j,n} = \frac{j - \frac{n-1}{2}}{c_n}, \tag{1}$$

where $c_n = \frac{n(n^2-1)}{12}$. We estimate the value of the scaling coefficient H by the statistic

$$\hat{H}_n^k = \sum_{j=0}^{n-1} w_{j,n} \log_2 |D_{j,k}| - \frac{1}{2}, \tag{2}$$

where n is the number of scales used. The estimation is based on the fact that

$$D_{j,k} \stackrel{d}{=} 2^{j(H+\frac{1}{2})} D_{0,k}, \tag{3}$$

i.e., $D_{j,k}$ has the same statistical distribution as $2^{j(H+\frac{1}{2})} D_{0,k}$, for all j and k. In the first paragraph, we prove the unbiasedness of the estimator. The unbiasedness of the estimator $\hat{H}_n^k$ means that the mean value of the estimator is H, whatever the number of used scales. In practice, if we take different realizations of the stochastic process with scaling index H of length 2^n, compute the in-products $d_{j,k}$ (which we denote here by small letters because they are numbers and not anymore stochastic variables) and the estimated values $\hat{h}_n^k$ (according to (2)), then the sample mean of the values will be around H; more exactly, according to the weak law of large numbers, the sample mean of the values $\hat{h}_n^k$ tends to H if the the sample of realizations of the process becomes larger. In the second and third paragraph, we show the basic asymptotic properties of the estimator. If we increase the number n of scales on which we observe the process, then the random variable $\hat{H}_n^k$ converges in distribution (or in probability) to the non random number H. The variance of the estimator decreases and the rate of convergence is $O(n^{-3})$. As a result the estimator $\hat{H}_n^k$ has asymptotically a normal distribution with mean H: the distribution of the error $(\hat{H}_n^k - H)n^{\frac{3}{2}}$ converges to a normal distribution with mean 0.

1.1 The unbiasedness of the estimator $\hat{H}_n^k$

We formulate some properties of the weights $w_{j,n}$, defined in (1).

Lemma 1

$$\sum_{j=0}^{n-1} w_{j,n} = 0;$$

$$\sum_{j=0}^{n-1} w_{j,n}^2 = \frac{1}{c_n};$$

$$\sum_{j=0}^{n-1} w_{j,n}.j = 1.$$

The properties of the weights are used to prove the following theorem:

Theorem 1 *The estimator $\hat{H}_n^k$ is unbiased:* $\mathrm{E}(\hat{H}_n^k) = H$.

Proof: Let $Y_{j,k} = \log_2 |D_{j,k}|$. Since the sequence $(D_{j,k})_{j\in\mathbb{Z}}$ is stationary, the sequence $(Y_{j,k})_{j\in\mathbb{Z}}$ is also stationary. Straightforward calculus gives

$$\mathrm{E}[\hat{H}_n^k] \quad = \quad \mathrm{E}[\sum_{j=0}^{n-1} w_{j,n} Y_{j,k}] - \frac{1}{2}$$

$$
\begin{aligned}
&= \sum_{j=0}^{n-1} w_{j,n}\mathrm{E}[Y_{j,k}] - \frac{1}{2} \\
&= \sum_{j=0}^{n-1} w_{j,n}(j(H+\frac{1}{2}) + \mathrm{E}[Y_{0,k}]) - \frac{1}{2} \qquad (4) \\
&= (H+\frac{1}{2})\sum_{j=0}^{n-1} w_{j,n} j + (\sum_{j=0}^{n-1} w_{j,n})\mathrm{E}[Y_{0,k}] - \frac{1}{2} \qquad (5) \\
&= H,
\end{aligned}
$$

In step (4), we used relation (3) and in step (5), we used lemma 1. Therefore, the estimator $\hat{H}_n^k$ is unbiased. ■

2 Asymptotic properties of the estimator $\hat{H}_n^k$

2.1 The consistency of the estimator $\hat{H}_n^k$

The estimator $\hat{H}_n^k$ is consistent if $\hat{H}_n^k$ converges to H in distribution (or in probability). The main problem in proving the consistency is that the wavelet coefficients $D_{j,k}$ and $D_{j',k}$ are not independent. Nevertheless, the consistency can be proved by looking at the normalised wavelet coefficients $\tilde{D}_{j,k} = D_{j,k}2^{-j(H+\frac{1}{2})}$ (see [11, p. 112]) and by proving that they form a strongly mixing stationary sequence. This means that the autocorrelation function of the logarithm of the normalised wavelet coefficients $\tilde{D}_{j,k}$ decays sufficiently fast as the number of scales n increases. The covariance between two wavelet coefficients $D_{j,k}$ and $D_{j',k}$ corresponding to the scales j and j', depends on the wavelet function ψ. The rescaled wavelet coefficients $\tilde{D}_{j,k}$ form a stationary sequence in the sense formulated in the following lemma (See [11, p. 112]).

Lemma 2 *Define* $\tilde{D}_{j,k} = \frac{D_{j,k}}{2^{j(H+\frac{1}{2})}}$. *The random sequence* $(\tilde{D}_{j,k})_{j\in\mathbb{Z}}$ *(where* $k \in \mathbb{Z}$*) is stationary:*

1. $\mathrm{E}[\tilde{D}_{j,k}] = \mathrm{E}[\tilde{D}_{0,k}]$.
2. $\mathrm{Covar}[\tilde{D}_{j,k}, \tilde{D}_{j',k}]$ *only depends on the difference of* j *and* j'*:*

$$
\mathrm{Covar}[\tilde{D}_{j,k}, \tilde{D}_{j',k}] = C\left(-\int_{-\infty}^{+\infty} A_\psi(2^{j-j'}, \tau - (2^{(j-j')}-1)k)|\tau|^{2H}\,d\tau\right) 2^{-(j-j')(H+\frac{1}{2})} \qquad (6)
$$

where $C = \frac{\mathrm{E}[X(1)]^2}{2}$ *and* A_ψ *is the function*

$$
A_\psi(\alpha,\tau) = \sqrt{\alpha}\int_{-\infty}^{+\infty} \psi(t)\psi(\alpha t - \tau)\,dt. \qquad (7)
$$

The autocorrelation function $\tilde{\rho}_k(j)$ of the sequence $(\tilde{D}_{j,k})_{j\in\mathbb{Z}}$ can be expressed as follows:

Lemma 3 *The correlation of $\tilde{D}_{0,k}$ and $\tilde{D}_{j,k}$ is given by*

$$\tilde{\rho}_k(j) = \frac{\int_{-\infty}^{+\infty} A_\psi(2^j, \tau - (2^j-1)k)|\tau|^{2H}\,d\tau}{\int_{-\infty}^{+\infty} A_\psi(1,\tau)|\tau|^{2H}\,d\tau}\; 2^{-j(H+\frac{1}{2})},$$

where A_ψ is the function given in (7).

The autocorrelation function of the normalised wavelet coefficients decays exponentially fast. The convergence rate depends on the number of null moments of the wavelet function: the more null moments the wavelet function has, the faster the decay of the autocorrelation function. This is the content of the following theorem. In this theorem, we will suppose that the wavelet function ψ has support $[0,1]$. This means that we use a rescaled version of a mother wavelet that has support $[0,b]$ $(b \geq 1)$.

Theorem 2 *Let ψ a wavelet function with support $[0,1]$ and suppose that ψ has I null moments ($I \geq 1$):*

$$\int_0^1 \psi(t)t^i\,dt = 0, \qquad i = 0,\ldots,I-1.$$

Let $k \geq 1$. Then we have

$$\tilde{\rho}_k(j) = O(2^{-(I-H)j}).$$

Proof: Let

$$f_k(2^j) = \mathrm{Covar}[\tilde{D}_{j,k}, \tilde{D}_{0,k}]$$

and $u = 2^j$. Then by the use of the expression (6),

$$f_k(u) = -\frac{\sigma^2}{2}\left(u^{-H}\int_0^1 ds\psi(s)\int_{us+(u-1)k-1}^{us+(u-1)k} dt\psi(us-t+(u-1)k)|t|^{2H}\right).$$

Since for all points (s,t) of the integration area $A = \{(s,t)|0 \leq s \leq 1, us+(u-1)k-1 \leq t \leq us+(u-1)k\}$ we have $t \geq 0$, we get

$$f_k(u) = -\frac{\sigma^2}{2}\left(u^{-H}\int_0^1 ds\psi(s)\int_{us+(u-1)k-1}^{us+(u-1)k} dt\psi(us-t+(u-1)k)t^{2H}\right).$$

We do the substitution $t \to v = us - t + (u-1)k$:

$$\begin{aligned} f_k(u) &= -\frac{\sigma^2}{2}u^{-H}\int_0^1 ds\psi(s)\int_0^1 dv\psi(v)(us+(u-1)k-v)^{2H} \\ &= -\frac{\sigma^2}{2}u^{H}\int_0^1 ds\psi(s)\int_0^1 dv\psi(v)(s+\frac{u-1}{u}k-\frac{v}{u})^{2H} \end{aligned}$$

Since for all $0 \leq s, v \leq 1$, $\frac{v}{u} < s + \frac{u-1}{u}k$ (we use that $k \geq 1$ and that $j > 1$), we can write

$$f_k(u) = -\frac{\sigma^2}{2}u^H$$
$$\times \int_0^1 ds\psi(s) \int_0^1 dv\psi(v) \sum_{i=0}^{+\infty} \left(\begin{array}{c} 2H \\ i \end{array} \right) (-1)^i (s + \frac{u-1}{u}k)^{2H-i} (\frac{v}{u})^i$$

Since $k \geq 1$, we can apply the theorem of Fubini to change the order of the integration and the summation; the condition to be satisfied is

$$\sum_{i=0}^{+\infty} | \left(\begin{array}{c} 2H \\ i \end{array} \right) | \int_0^1 ds \int_0^1 dv |\psi(s)||\psi(v)| (\frac{u-1}{u}k + s)^{2H-i} (\frac{v}{u})^i < \infty.$$

This sum is finite, since

$$\begin{aligned}
&\sum_{i=0}^{+\infty} | \left(\begin{array}{c} 2H \\ i \end{array} \right) | \int_0^1 ds \int_0^1 dv |\psi(s)||\psi(v)| (\frac{u-1}{u}k + s)^{2H-i} (\frac{v}{u})^i \\
\leq\ & M^2 (\frac{u-1}{u}k + 1)^{2H} \sum_{i=0}^{+\infty} | \left(\begin{array}{c} 2H \\ i \end{array} \right) | \int_0^1 ds (\frac{u-1}{u}k + s)^{-i} u^{-i} \\
\leq\ & M^2 (\frac{u-1}{u}k + 1)^{2H} (1 + \sum_{i=1}^{+\infty} | \left(\begin{array}{c} 2H \\ i \end{array} \right) | (\frac{u-1}{u}k)^{-i} u^{-i}) \\
\leq\ & M^2 (\frac{u-1}{u}k + 1)^{2H} \sum_{i=0}^{+\infty} | \left(\begin{array}{c} 2H \\ i \end{array} \right) | (u-1)^{-i} k^{-i}) && (8) \\
\leq\ & M^2 (k+1)^{2H} \sum_{i=0}^{+\infty} | \left(\begin{array}{c} 2H \\ i \end{array} \right) | (2k)^{-i}. && (9)
\end{aligned}$$

In step (8), we used that $u - 1 \geq \frac{1}{2}$. The last series is convergent for all $k \geq 1$, since a power series is absolute convergent within its convergence circle. So $f_k(u)$ becomes

$$f_k(u) = -\frac{\sigma^2}{2}u^H \sum_{i=0}^{+\infty} \left(\begin{array}{c} 2H \\ i \end{array} \right) (-1)^i u^{-i} \int_0^1 ds\psi(s)(s+\frac{u-1}{u}k)^{2H-i} \int_0^1 dv\psi(v)v^i.$$

We use that the wavelet function ψ has I null moments to get

$$\begin{aligned}
f_k(u) &= -\frac{\sigma^2}{2}u^H \sum_{i=I}^{+\infty} \left(\begin{array}{c} 2H \\ i \end{array} \right) (-1)^i u^{-i} \int_0^1 ds\psi(s)(s + \frac{u-1}{u}k)^{2H-i} \int_0^1 dv\psi(v)v^i \\
&= -\frac{\sigma^2}{2}u^{H-I} \sum_{l=0}^{+\infty} \left(\begin{array}{c} 2H \\ l+I \end{array} \right) (-1)^{l+I} u^{-l} \\
&\quad \times \int_0^1 ds\psi(s)(s + \frac{u-1}{u}k)^{2H-l-I} \int_0^1 dv\psi(v)v^{l+I}.
\end{aligned}$$

So the upper bound for $f_k(u)$ becomes

$$|f_k(u)| \leq \frac{\sigma^2}{2} M^2 u^{H-I} \sum_{l=0}^{+\infty} | \left(\begin{array}{c} 2H \\ l+I \end{array} \right) | u^{-l} \int_0^1 ds (s + \frac{u-1}{u} k)^{2H-l-I}.$$

Let

$$g_k(u) = \sum_{l=0}^{+\infty} | \left(\begin{array}{c} 2H \\ l+I \end{array} \right) | u^{-l} \int_0^1 ds (s + \frac{u-1}{u} k)^{2H-l-I}.$$

Let $L = \max\{\lfloor 2H - I \rfloor, 0\}$.We split up this series in two parts:

$$g_k(u) = g_{1,k}(u) + g_{2,k}(u),$$

where

$$g_{k,1}(u) = \sum_{l=0}^{L} | \left(\begin{array}{c} 2H \\ l+I \end{array} \right) | u^{-l} \int_0^1 ds (s + \frac{u-1}{u} k)^{2H-l-I}$$

and

$$g_{k,2}(u) = \sum_{l=L+1}^{+\infty} | \left(\begin{array}{c} 2H \\ l+I \end{array} \right) | u^{-l} \int_0^1 ds (s + \frac{u-1}{u} k)^{2H-l-I}.$$

We have

$$\begin{aligned} |g_{k,1}(u)| &= \sum_{l=0}^{L} | \left(\begin{array}{c} 2H \\ l+I \end{array} \right) | u^{-l} \int_0^1 ds (s + \frac{u-1}{u} k)^{2H-l-I} \qquad (10) \\ &\leq \sum_{l=0}^{L} | \left(\begin{array}{c} 2H \\ l+I \end{array} \right) | \int_0^1 ds (s + k)^{2H-l-I} \\ &\leq \sum_{l=0}^{L} | \left(\begin{array}{c} 2H \\ l+I \end{array} \right) | (1+k)^{2H-I-l}. \end{aligned}$$

In step (10), we used that $u \geq 1$ and that $\frac{u-1}{u} \leq 1$. Also we have

$$\begin{aligned} g_{k,2}(u) &= \sum_{l=L+1}^{+\infty} | \left(\begin{array}{c} 2H \\ l+I \end{array} \right) | u^{-l} \int_0^1 ds (s + \frac{u-1}{u} k)^{2H-l-I} \\ &\leq \sum_{l=L+1}^{+\infty} | \left(\begin{array}{c} 2H \\ l+I \end{array} \right) | 3^{-l} \int_0^1 ds (s + \frac{1}{2} k)^{2H-l-I} \qquad (11) \\ &\leq \sum_{l=L+1}^{+\infty} | \left(\begin{array}{c} 2H \\ l+I \end{array} \right) | 3^{-l} (\frac{1}{2} k)^{2H-l-I} \qquad (12) \\ &\leq 2^{I-2H} \sum_{l=L+1}^{+\infty} | \left(\begin{array}{c} 2H \\ l+I \end{array} \right) | (\frac{2}{3})^l. \end{aligned}$$

In step (11), we used that $\frac{u-1}{u} \geq \frac{1}{2}$ (since $u \geq 2$) and that $u \geq 3$. In step (12), we used that $k \geq 1$. We conclude that

$$\begin{aligned}
|f_k(u)| \quad &\leq \quad \frac{\sigma^2}{2} M^2 u^{H-I} \Big(\sum_{l=0}^{L} \left| \begin{pmatrix} 2H \\ l+I \end{pmatrix} \right| (1+k)^{2H-I-l} \\
&\quad + 2^{I-2H} \sum_{l=L+1}^{+\infty} \left| \begin{pmatrix} 2H \\ l+I \end{pmatrix} \right| (\frac{2}{3})^l \Big) .
\end{aligned}$$

So

$$|f_k(u)| \leq C^k u^{H-I}$$

for some constant $C_k > 0$. This ends the proof of the theorem. ■

Remark that in the last theorem we have to use a wavelet function translated over a position $k \geq 1$ to insure the convergence of the series (8) and (12). The proof is no longer valid for $k = 0$. The following result guarantees that in the case of normally distributed wavelet coefficients the correlation between the logarithm of the wavelet coefficients can be bounded by the correlation between the wavelet coefficients itself, up to a constant. We will not give the proof here.

Lemma 4 *Let* $(X, Y) = N(0, 0, 1, 1, \rho)$. *Then*

$$|\mathrm{Covar}(\log |X|, \log |Y|)| \leq K \frac{\sqrt{\rho}}{(1-\rho^2)^{\frac{1}{4}}},$$

for some $K > 0$.

The consistency of the estimator and the rate of convergence of the variance of the estimator are given in the following theorem. This theorem is proved by using theorem 2 and lemma 4.

Theorem 3 *For the estimator* $\hat{H}_n^k$ *one has:*

$$\lim_{n \to +\infty} \mathrm{Var}(\hat{H}_n^k) = 0,$$

so that $\hat{H}_n^k$ *is consistent. Furthermore*

$$\lim_{n \to \infty} n^3 \mathrm{Var}[\hat{H}_n^k] := \sigma_k^2 \quad = \quad 12 \mathrm{E}[Y_{1,k}]^2 + 24 \sum_{j=1}^{+\infty} \mathrm{E}[Y_{1,k} Y_{1+j,k}], \qquad (13)$$

where $Y_{j,k} = \log_2 |D_{j,k}| - \mathrm{E}(\log_2 |D_{j,k}|)$.

Proof: To simplify the notation, we write Y_j instead of $Y_{j,k}$. First, we prove the finiteness of the series in (13). By lemma (2) and (4), $\mathrm{Covar}(Y_1, Y_{1+j})$ is of order $2^{-\frac{1}{2}(I-H)j}$. Therefore

$$\begin{aligned} \left|\sum_{j=1}^{+\infty} \mathrm{Covar}(Y_1, Y_{1+j})\right| &\leq \sum_{j=1}^{+\infty} |\mathrm{Covar}(Y_1, Y_{1+j})| \\ &\leq C \sum_{j=1}^{+\infty} 2^{-\frac{1}{2}(I-H)j} < \infty. \end{aligned}$$

If $\rho_j = \mathrm{Covar}(Y_1, Y_{1+j})$, then

$$\begin{aligned} &\mathrm{Var}[\hat{H}_n^k] \\ &= \frac{1}{c_n^2}\mathrm{Var}[\sum_{p=0}^{n-1}(p - \frac{n-1}{2})Y_p] \\ &= \frac{1}{c_n^2}\sum_{p=0}^{n-1}\sum_{q=0}^{n-1}(p - \frac{n-1}{2})(q - \frac{n-1}{2})\mathrm{Covar}(Y_p, Y_q) \\ &= \frac{1}{c_n^2}\sum_{p=0}^{n-1}(p - \frac{n-1}{2})^2\mathrm{Var}[Y_p] + \frac{1}{c_n^2}\sum_{p\neq q}(p - \frac{n-1}{2})(q - \frac{n-1}{2})\mathrm{Covar}(Y_p, Y_q) \\ &= \frac{1}{c_n^2}\sum_{p=0}^{n-1}(p - \frac{n-1}{2})^2\mathrm{Var}[Y_p] \qquad (14) \\ &\quad + \frac{2}{c_n^2}\sum_{p<q}(p - \frac{n-1}{2})(q - \frac{n-1}{2})\mathrm{Covar}(Y_p, Y_q) \qquad (15) \\ &= \frac{1}{c_n}\mathrm{Var}[Y_1] + \frac{2}{c_n^2}\sum_{p<q}(p - \frac{n-1}{2})(q - \frac{n-1}{2})\mathrm{Covar}(Y_1, Y_{1+q-p}) \qquad (16) \\ &= \frac{1}{c_n}\mathrm{Var}[Y_1] + \frac{2}{c_n^2}\sum_{j=1}^{n-1}\sum_{p=0}^{n-1-j}(p - \frac{n-1}{2})(p + j - \frac{n-1}{2})\mathrm{Covar}(Y_1, Y_{1+j}). \end{aligned}$$

In step (15), we used the stationarity of the sequence $(Y_j)_{j\in\mathbb{Z}}$. In step (16), we made the substitution $j = q - p$ and changed the order of the sums. Since for fixed j $(1 \leq j \leq n-1)$

$$\sum_{p=0}^{n-1-j}(p - \frac{n-1}{2})(p + j - \frac{n-1}{2}) = \frac{1}{12}(j + 2j^3 - n - 3jn^2 + n^3),$$

the asymptotic behaviour of $c_n\mathrm{Var}[\hat{H}_n^k]$ becomes

$$c_n\mathrm{Var}[\hat{H}_n^k] \sim \mathrm{Var}[Y_1] + \frac{2}{c_n}\sum_{j=1}^{n-1}\frac{1}{12}(n^3 - n + 2j^3 - 3jn^2 + j)\mathrm{Covar}(Y_1, Y_{1+j}).$$

Since $c_n \sim \frac{1}{12}n^3$, we have

$$\lim_{n\to+\infty} n^3 \mathrm{Var}[\hat{H}_n^k] = 12\mathrm{Var}[Y_1]$$
$$+24 \lim_{n\to+\infty} \sum_{j=1}^{n-1} (1 - \frac{1}{n^2} + 2(\frac{j}{n})^3 - 3\frac{j}{n} + \frac{j}{n^3})\mathrm{Covar}(Y_1, Y_{1+j}).$$

Since

$$\begin{aligned}
|\sum_{j=1}^{n-1} \frac{j}{n^3}\mathrm{Covar}(Y_1, Y_{1+j})| &\leq \frac{1}{n^2}\sum_{j=1}^{n-1} \frac{j}{n}|\mathrm{Covar}(Y_1, Y_{1+j})| \\
&\leq \frac{1}{n^2}\sum_{j=1}^{n-1} |\mathrm{Covar}[Y_1, Y_{1+j}]| \\
&\leq \frac{1}{n^2}\sum_{j=1}^{+\infty} |\mathrm{Covar}[Y_1, Y_{1+j}]|,
\end{aligned}$$

we have

$$\lim_{n\to+\infty} \sum_{j=1}^{n-1} \frac{j}{n^3}\mathrm{Covar}[Y_1, Y_{1+j}] = 0. \tag{17}$$

In a similar way, we have

$$\lim_{n\to+\infty} \frac{1}{n^2}\sum_{j=1}^{n-1} \mathrm{Covar}(Y_1, Y_{1+j}) = 0. \tag{18}$$

Recall Kronecker's lemma which says the following: if $(x_n)_{n\leq 1}$ is a sequence of real numbers and $(a_n)_{n\geq 0}$ is a sequence of strictly positive numbers which tends to infinity. Then the convergence of $\sum_{n=1}^{+\infty} \frac{x_n}{a_n}$ implies that $\frac{1}{a_n}\sum_{j=1}^{n} x_j \to 0$. Take now $x_n = n\mathrm{Covar}[Y_1, Y_n]$ and $a_n = n$. Since $\sum_{n=1}^{+\infty} \mathrm{Covar}[Y_1, Y_n]$ is convergent, Kronecker's lemma gives that

$$\frac{1}{n}\sum_{j=1}^{n} j\mathrm{Covar}[Y_1 Y_j] \to 0.$$

If we take $x_n = n^3\mathrm{Covar}[Y_1, Y_n]$ and $a_n = n^3$, Kronecker's lemma also gives that

$$\frac{1}{n^3}\sum_{j=1}^{n} j^3\mathrm{Covar}[Y_1 Y_j] \to 0.$$

From (17) and (18), it follows that

$$\lim_{n\to+\infty} n^3\mathrm{Var}[\hat{H}_n^k] = 12\mathrm{Var}[Y_1] + 24\sum_{j=1}^{+\infty} \mathrm{Covar}(Y_1, Y_{1+j}).$$

This ends the proof of the theorem. ■

2.2 A Central Limit Theorem for the weighted sum

A central limit theorem is given for a particular weighted sum of random variables in which the random variables far apart from one another are nearly independent in the sense that they are α-mixing. For a sequence $\{X_n\}_{n\in\mathbb{N}}$ of random variables, let α_n be a number such that

$$|P(A\cap B)-P(A)P(B)|\leq\alpha_n$$

for $A\in\sigma(X_0,\ldots,X_k)$, $B\in\sigma(X_{k+n},X_{k+n+1},\ldots)$, and $k\geq 0$,$n\geq 1$.
If $\alpha_n\to 0$, then the sequence $\{X_n\}_{n\in N}$ is said to be α-mixing. (see [3, p. 315]).

Theorem 4 *Suppose that $X_0,X_1,\ldots$ is stationary and α-mixing with $\alpha_n=O(n^{-6-\epsilon})$ for some $\epsilon>0$ and that $\mathrm{E}[X_n]=0$ and $\mathrm{E}[X_n^{12}]<\infty$. If*

$$S_n=\sum_{j=0}^{n-1}w_{j,n}X_j,$$

where $w_{j,n}=\frac{j-\frac{n-1}{2}}{c_n}$ and $c_n=\frac{n(n^2-1)}{12}$, then

$$n^3Var[S_n]\to\sigma^2=12\mathrm{E}[X_1]^2+24\sum_{j=1}^{+\infty}\mathrm{E}[X_1X_{1+j}]. \tag{19}$$

If $\sigma>0$, then

$$\frac{n^{3/2}}{\sigma}S_n\xrightarrow{d}N(0,1). \tag{20}$$

To prove this theorem, we adapted the proof of the central limit theorem [3, pp. 316–118] in the case of an unweighted sum of a stationary α-mixing sequence. The proof is not given here.

The technical conditions of theorem 4 are satisfied in our case: theorem 2 guarantees that the sequence $(\log|\tilde{D}_{j,k}|)_{j\in\mathbb{Z}}$, where $k\in\mathbb{N}_0$, is α-mixing and that the α-mixing coefficients decay exponentially fast. Also the moments of $\log|X|$, where X is normally distributed, exist. From theorem (4), it follows immediately that the estimator $\hat{H}_n^k$ has asymptotically a normal distribution:

Theorem 5 *The estimator $\hat{H}_n^k$ has asymptotically a normal distribution:*

$$\hat{H}_n^k\sim N(H,\frac{\sigma_k^2}{n^3}),\ as\ n\to+\infty$$

where

$$\begin{aligned}\sigma_k^2 &= 12\mathrm{Var}[\log_2|D_{1,k}|]\\ &+\ 24\sum_{j=1}^{+\infty}\mathrm{Covar}[\log_2|D_{1,k}|,\log_2|D_{1+j,k}|].\end{aligned}$$

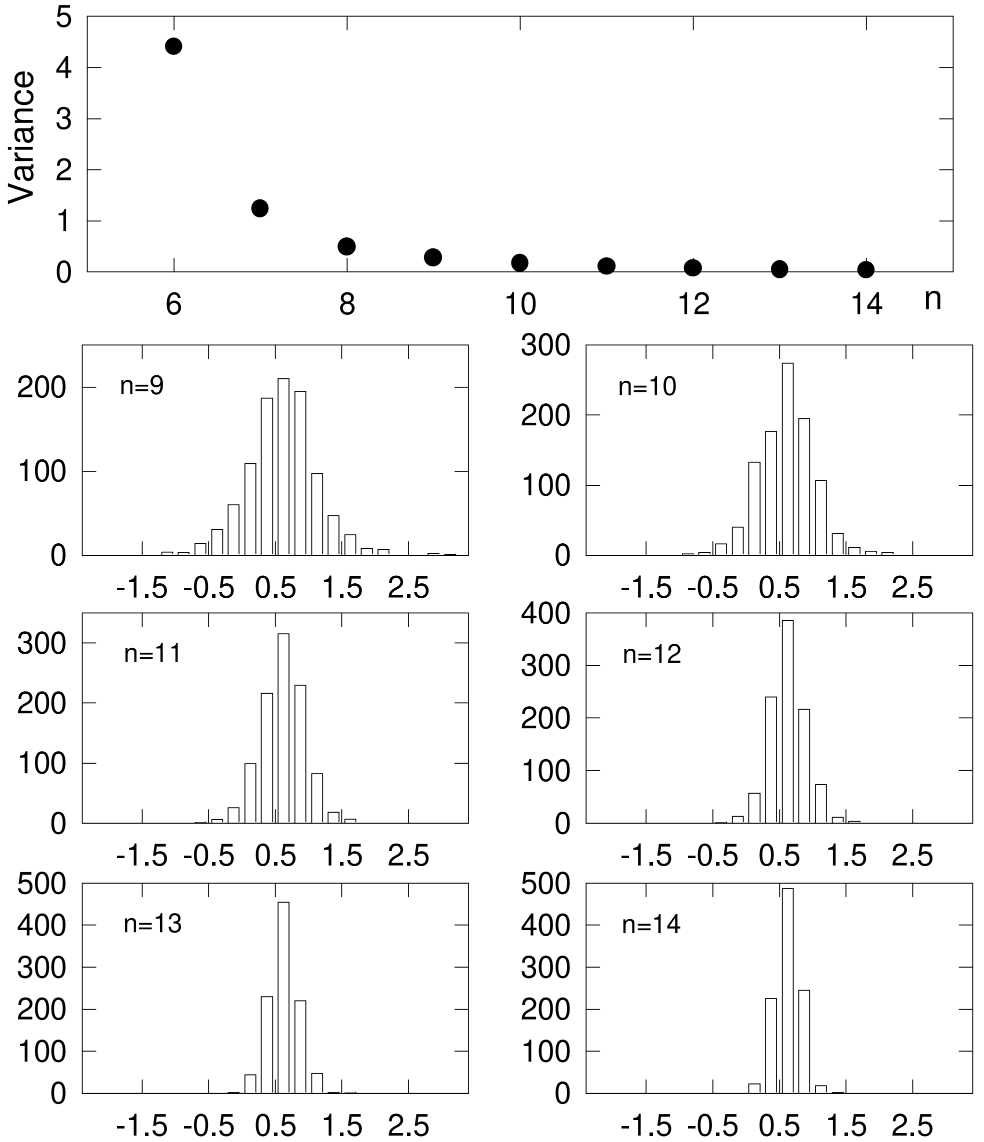

Figure 1: Sample variance and histograms (1000 realizations) of the estimator for H as a function of the number of scales n in the case of Brownian motion.

3 Simulation

We illustrate the theory by the following simulation of Brownian motion. We compute the estimator using 1000 realisations of Brownian motion. The Haar wavelet is used to estimate H, which is $1/2$ for Brownian motion. In figure 1, we first give a plot of the variance of the estimator $\hat{H}_n^k$ as a function of the number of scales used. If we use n scales, we simulate 2^n points of Brownian motion; n varies from 6 to 14. In figure 1, we also give histograms of the estimator for different scales. The histograms have a mean value around $\frac{1}{2}$ and they become narrower as the number of scales used increases.

4 Conclusion

The proposed estimator is useful to estimate the parameter of self-similarity of a self-similar process. In this note, we studied the basic properties of the estimator in the case that the process can be modeled as fractional Brownian motion. This class of self-similar processes can be extended to stable self-similar processes. The same estimator can be used and similar properties can be deduced.

References

[1] M. Abramowitz, I.A. Stegun, *Handbook of Mathematical Functions*, Dover Publications, New York, 1970.

[2] P.J. Bickel, K.A. Doksum, *Mathematical Statistics*, Holden-Day, San Francisco, 1977.

[3] P. Billingsley, *Probability and Measure*, John Wiley & Sons, New York, 1979.

[4] P.J. Brockwell, R.A. Davis, *Time series: Theory and Applications*, Springer-Verlag, New York, 1991.

[5] K.L. Chung, *A Course in Probability Theory*, Academic Press, New York, 1974.

[6] A. Cohen, *Wavelet Bases, Approximation Theory and Subdivision Schemes*, in 'Approximation Theory VII' (E.W. Cheney et al., eds.), Academic Press, Boston, 1992, pp. 63–85.

[7] A. Cohen, I. Daubechies, B. Jawerth, P. Vial, *Multiresolution Analysis, Wavelets and Fast Algorithms on an Interval*, C. R. Acad. Sci. Paris **316** (1993), 417–421.

[8] I.A. Ibragimov, Y.A. Rozanov, *Gaussian Random Processes* , Springer-Verlag, New York, 1978.

[9] W. Feller, *An Introduction to Probability Theory and its Applications, Vol. II*, John Wiley & Sons, New York, 1971.

[10] P. Flandrin, *On the Spectrum of Fractional Brownian Motion*, IEEE Trans. Inform. Theory **35** (1989), 197–199.

[11] P. Flandrin, *Fractional Brownian Motion and Wavelets*, in 'Wavelets, Fractals and Fourier Transforms' (M. Farge et al., eds.), Clarendon Press, Oxford, 1993, pp. 109–122.

[12] P. Flandrin, *Wavelet Analysis and Synthesis of Fractional Brownian Motion*, IEEE Trans. Inform. Theory **38** (1992), 910–917.

[13] M. Maejima, *Self-similar processes and limit theorems*, Sugaku Expositions, Amer. Math. Soc. **2**, (1) (1989), 103–123.

[14] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, Singapore, 1968.

[15] M.S. Taqqu, *A bibliographic guide to self-similar processes and long-range dependence*, in 'Dependence in Probability and Statistics' (E. Eberlein and M.S. Taqqu, eds.), Birkhäuser, Boston, 1986, pp. 137–162.

[16] A. Tewfik, M. Kim, *Correlation Structure of the Discrete Wavelet Coefficients of Fractional Brownian Motion*, IEEE Trans. Inform. Theory **38** (1992), 904–909.

[17] G. Samorodnitsky, M.S. Taqqu, *Stable Non-Gaussian Random Processes*, Chapman & Hall, New York, 1994.

[18] J.D. Scargle, T. Steiman-Cameron, K. Young, D.L. Donoho, J.P. Crutchfield, J. Imamura, *The Quasi-periodic Oscillations and Very Low Frequency Noise op Scorpius X-1 as Transient Chaos: a Dripping Handrail?*, Astrophys. J. **411** (1993), L91–L94.