WAVELET BASED ESTIMATOR FOR THE SELF-SIMILARITY PARAMETER OF α-STABLE PROCESSES

Patrice ABRY[1] *, Lieve DELBEKE*[2] *and Patrick FLANDRIN*[1]

(1) - *CNRS URA 1325 - ENS Lyon - 46, allée d'Italie 69 364 LYON Cedex 07 - France*
tel: (+33) 4 72 72 84 93 - Fax: (+33) 4 72 72 80 80
pabry,flandrin@physique.ens-lyon.fr - http://www.physique.ens-lyon.fr/ts

(2) *KU Leuven, Department of Mathematics, Celestijnenlaan 200 B, 3001 Heverlee, Belgium*
tel: (+32) 16 32 70 49 - Fax: (+32) 16 32 79 89 - Lieve.Delbeke@wis.kuleuven.ac.be

ABSTRACT

We, here, study self-similar processes with possibly infinite second-order statistics and long-range dependence. To do so, we detail the statistical properties of the wavelet coefficients of α-stable self similar processes, used as a paradigm for those situations. We, then, propose a wavelet-based estimator for the self-similarity parameter and analyse its statistical performance both theoretically and numerically. We show that it is unbiased, that its variance decreases as the inverse of the length of the data and that it can be easily implemented.

1. MOTIVATION

Evidencing self-similarity and estimating the corresponding parameters in experimental data is a major, yet difficult, issue. Self-similarity indeed implies non stationarity [14] and can, quite often, be related to long-range statistical dependences (LRD), which result in severe complications when data analysis and parameter estimation tasks are undertaken [3]. Fractional Brownian motion of index H is as example of such a self similar process which exhibits long-range statistical dependence whenever $d = H - 1/2 > 0$. For the analysis of such kind of processes, whose second order statistics are finite, the wavelet transform already proved to be a relevant tool [10, 15, 1, 2]. We are here interested in the study of self similar processes with infinite second order statistics (ISOS) which may also exhibit long-range dependences. For those situations, α-stable processes can be regarded as a paradigm. The linear fractional stable motion (LFSM) model indeed provides us with a broad class of α-stable self similar processes with stationary increments and possibly long-range dependence [14] .

The aim of this work is therefore to propose a generalization of previous wavelet-based estimators of the self-similarity parameter [1, 10], dedicated to processes that may present both type of complications (LRD and ISOS) and show theoretically and numerically that it exhibits excellent statistical properties (unbiasedness and $1/n$ variance decrease).

2. SYMMETRIC α-STABLE SELF SIMILAR PROCESSES

Self similar process. A process x is said to be self-similar with self similarity parameter H, and denoted H-ss, if and only if:

$$(x(t_1), \dots x(t_n)) \stackrel{d}{=} c^{-H}\,(x(ct_1), \dots x(ct_n))\,, \forall n,\ \forall c > 0,$$

where $\stackrel{d}{=}$ means equality in law. Though self-similarity implies non stationarity [14], there is an interesting subclass of H-ss processes, whose increments are stationary. A process x is said to possess stationary increments y_h if the finite-dimensional distributions of $y_h(t) = x(t+h) - x(t)$ do not depend on t. A self-similar process with stationary increments is denoted by H-sssi.

Symmetric α-stable process. Let $x(t)$ be a symmetric α-stable (SαS) process defined using the integral representation [14, 6]:

$$x(t) = \int f(t,u) M(du)\,,$$

where $M(du)$ is an SαS measure and $f(t,u)$ an integration kernel that controls the time dependence of the statistical properties of x. Two examples of particular interest here are given in [14].

- When $f(t,u) = 1$ for $0 < u \leq t$ and $f(t,u) = 0$ elsewhere, x reduces to the so-called Lévy process. It has stationary and independent increments and is a self similar process with self-similarity parameter $H = 1/\alpha$.
- A simple version of the LFSM is obtained through

$$f(t,u) = (t-u)_+^d - (-u)_+^d,$$

where $(u)_+ = u$ if $u \geq 0$ and 0 elsewhere and the fractional parameter d satisfy $-\infty < d < 1/2$ [14]. It basically amounts to fractionally integrating a stationary sequence of i.i.d. α-stable variables. Note that $d = 0$ requires a specific definition not developed here [14]. The LFSM is self similar with $H = d + 1/\alpha$ and has stationary increments, whose statistical dependence is controlled by d. Long-range dependence occurs for $d > 0$, i.e., $H > 1/\alpha$ [14]. It can more precisely be shown [14] that if x is LFSM, the codifference of its increments y_h slowly decreases as a power-law of the lag τ:

$$\text{Cod}\; y_h(t) y_h(t+\tau) \sim |\tau|^{(\alpha-1)H}\,, |\tau| \to +\infty,$$

for all the pairs (α, H) considered here. For further technical details and definition of the codifference, see [14]. We simply here recall that it plays for α-stable process the role of the covariance for processes with finite second order statistics.

3. WAVELET COEFFICIENTS OF α-STABLE SYMMETRIC SELF SIMILAR PROCESSES

Wavelet transform. Let $d_x(j,k)$ be the coefficients of the discrete wavelet transform (DWT) defined as $d_x(j,k) = \langle x, \psi_{j,k}\rangle$. The analysing functions $\psi_{j,k}(t) = 2^{-j/2}\psi_0(2^{-j}t - k)$ are dilated and translated templates of the mother wavelet $\psi_0(t)$ [5]. We will not go further into the introduction of the DWT and of the underlying multiresolution analysis, however, we insist here on the two following items that play a key role in the analysis below:

- **I1:** The analysis basis is designed using the dilation operator, $\psi_{j,0}(t) = \psi_0(t/2^j)/2^{j/2}$, and therefore is, by nature, self-similar.
- **I2:** The mother-wavelet is characterized by its number N of vanishing moments:

$$\int t^k \psi_0(t)dt \equiv 0,\ k = 0, \ldots, N-1. \qquad (1)$$

By definition of the wavelet transform, $N \geq 1$ [5].

Moreover, most of the results stated below hold for both the continuous (CWT) and discrete wavelet transform, however for the estimation purposes developped here the computation of a CWT as compared to a DWT proves useless [8]. We therefore hereafter restrict to the DWT which can be easily implemented using a fast pyramidal filter-bank based algorithm [5].

H-ss process. Let x be $H-$ss process, its wavelet coefficients reproduce the self-similarity through [6, 7, 13]:

$$\textbf{P0:}\ (d_x(j,0), d_x(j,1), \ldots, d_x(j,N_j-1)) \stackrel{d}{=} 2^{j(H+\frac{1}{2})}(d_x(0,0), d_x(0,1), \ldots, d_x(0,N_j-1)). \qquad (2)$$

which indicates self-similarity with parameter $H + 1/2$, the $+1/2$ term being simply due to the chosen normalization of the wavelet. This is a direct consequence of **I1**.

Process with stationary increments. Let x be a process with stationary increments, which is not necessary stationary. Its wavelet coefficients satisfy:

P1: $\{d_x(j,1), \ldots d_x(j,k)\}$ form, at each octave j, stationary sequences.

This is a consequence of **I2** [10, 15, 11, 4].

SαS processes with stationary increments. Let x be a SαS process, with stationary increments, then the sequences of wavelet coefficients satisfy [7, 6, 13]:

P2: $\{d_x(j,1), \ldots d_x(j,k)\}$ form, at each octave j, identically distributed SαS processes.

LFSM. Let x be a LFSM. For each octave j, the statistical dependence structure of the wavelet coefficients can be analyzed through their codifferences. It has been shown [6, 9] that, when $|2^j k - 2^{j'}k'| \to +\infty$,

$$\textbf{P3:}\ |\text{Cod}\, d_x(j,k)d_x(j',k')| \leq C|2^j k - 2^{j'}k'|^{-(\alpha/2)(N-H)}. \qquad (3)$$

This results from both **I1** and **I2** and shows that the range of dependence of any two wavelet coefficients can be significantly shortened by increasing the number of vanishing moments of the mother wavelets. More specifically the long range dependence structure that can be introduced in the LFSM if choosing a fractional integration of parameter $d > 0$ can be significantly shortened by increasing N. This effect receives more attention in the following section.

4. ESTIMATION OF THE SELF-SIMILARITY PARAMETER

After proposing an intuitive understanding of the wavelet based estimator for $\hat{H}$, we give its definition and derive its statistical performance both from theoretical and numerical arguments.

Data generation and numerical simulations. Using the algorithm described in [14], page 371, we synthetized $L = 60$ realizations of LFSMs (with $n = 8192$ samples each), α and H are chosen such that x is LRD: $\alpha = 1.5, H = 0.85$. The wavelet transforms are performed using Daubechies wavelets (because of compact support and easy control of N, but not because of orthonormality - note that the theoretical results proven here do not depend on mother wavelet except on its number of vanishing moments N. More details can be found in [8]). The covariance functions shown in Fig. 1 result from an average of the L covariance estimates obtained using a standard biased covariance estimator. Independently on each realization, an estimation $\hat{H}_l$ of H is performed. Mean $\mu_{\hat{H}}$ and variance $\sigma^2_{\hat{H}}$ for $\hat{H}$ are computed using the standard sample estimators. Bias and variances are represented in Fig. 2 and compared to the theoretical values. Confidence intervals (cf. Fig. 2 (a)) for the mean are given using a Gaussian approximation for $\hat{H}$ (see below).

Principle of the estimation. Let x be a H-sssi SαS process, from **P0**, we obtain:

$$\mathbb{E}\log_2|d_x(j,k)| = j(H+1/2) + \mathbb{E}\log_2|d_x(0,k)|, \qquad (4)$$

which allows to estimate H by measuring the slope in a linear fit in the $\mathbb{E}\log_2|d_x(j,k)|$ versus j plot. The difficulty lies in the fact that $\mathbb{E}\log_2|d_x(j,k)|$ needs to be estimated. Using the stationary property **P1**, one can perform an average along the time index k, at each octave j:

$$Y_j = 1/n_j \sum_{k=1}^{n_j} \log_2|d_x(j,k)|,$$

where n_j is the number of wavelet coefficients at scale j. Obviously, one has:

$$\begin{aligned}\mathbb{E}Y_j &= j(H+1/2) + 1/n_j \textstyle\sum_{k=1}^{n_j} \mathbb{E}\log_2|d_x(0,k)| \\ &= j(H+1/2) + \mathbb{E}\log_2|d_x(0,0)|.\end{aligned} \qquad (5)$$

Moreover, from **P2**, we have that $\log_2|d_x(j,k)|$ has finite second-order statistics [6, 12]. In the specific case where x is a LFSM, the covariance function of $\log_2|d_x(j,k)|$ behaves as the codifference of $d_x(j,k)$, i.e., when $|k-k'| \to +\infty$,

$$|\text{Cov}\ \log_2|d_x(j,k)|, \log_2|d_x(j,k')|| \leq C|k-k'|^{-(\alpha/4)(N-H)}. \qquad (6)$$

This result, that comes from **P3**, was recently proven in [6, 9] and is a crucial point. It is known indeed [3] that the performance of the sample mean estimator are poor if LRD exists among data; for instance its variance decreases much slower than the usual $1/n$ behaviour. Eq. (6) shows that LRD, that can be introduced in the LFSM by choosing $d > 0$, can be significantly reduced among the $\log_2|d_x(j,k)|$ by increasing the number of vanishing moments N and turned to short range dependencies; allowing, for instance, an efficient use of the sample mean estimator. Note, however, that, for a given H, this decorrelation effect requires larger N's for smaller α's: basically, N has to behave as $1/\alpha$. This theoretical result is here further illustrated using the numerical simulations described above. Fig. 1 compares the covariance functions of $\log|y|$ (where $y(t) = x(t+1) - x(t)$ and x is LFSM), and of $\{\log|d_x(j,k)|, k \in \mathbb{Z}\}$ obtained using Daubechies1 and Daubechies3 wavelets. Fig. 1 clearly shows that covariance functions behave as power-laws of the lag for large lags and that their rates of decrease significantly increase when N is increased, in agreement with (6). It therefore confirms that an increase of N insures a better decorrelation among the $\log_2|d_x(j,k)|$.

An exact decorrelation hypothesis, **HYP1**, among the $\log|d_x(j,k)|$ can be used to derive closed-form relations for the variance of Y_j [6]:

$$\begin{aligned} \text{Var}\, Y_j &= \text{Var}\ \left(1/n_j \textstyle\sum_k \log_2|d_x(0,k)|\right) \\ &= \text{Var}\ (\log_2|d_x(0,0)|)/n_j \\ &= (\log_2(e))^2\pi^2(1+2/\alpha^2)/(12n_j) \end{aligned} \quad (7)$$

Both stationarity (**P1**) and short range dependence (**P3**) indicate that measuring the slope in a Y_j versus j plot will provide us with a relevant estimate of H. This can be made more precise.

Definition of $\hat{H}$. Let $\hat{H}$ denote the estimate of H obtained as the slope of a linear fit in a Y_j versus j plot:

$$\hat{H} = \sum_j w_j Y_j - 1/2,$$

where, $\sum_j$ is to be understood as $\sum_{j_1}^{j_2}$ if the linear fit is performed on the range of scales $j_1 \le j \le j_2$, and the weigths w_j satisfy $\sum_j jw_j = 1$ and $\sum_j w_j = 0$. They follow the general form $w_j = 1/a_j(S_0 j - S_1)/(S_0S_2 - S_1^2)$ where $S_m = \sum_{j=j_1}^{j_2} a_j^{-1} j^m$ ($m = 0, 1, 2$) and the a_j are arbitrary numbers.

Bias of $\hat{H}$. Using Eq. (5) above, one obtains:

$$\begin{aligned} \mathbb{E}\hat{H} &= \textstyle\sum_j w_j \mathbb{E} Y_j - 1/2 \\ &= \underbrace{\sum_j j w_j}_{=1}(H+1/2) + \underbrace{\sum_j w_j}_{=0} \mathbb{E}\log_2|d_x(0,0)| - 1/2 \\ &= H, \end{aligned}$$

which shows that the estimate is strictly unbiased even for finite length duration of analyzed data. Note that this result is valid regardless of the value of α ($0 < \alpha \le 2$). Fig. 2 (a) consists of the bias of $\hat{H}$, with 95% confidence intervals, as a function of n. It shows that the estimator is unbiased even for small length of analysed process. Let us insist again on the fact that this directly results from the wavelet basis being built from a dilation operator (**I1**).

Variance of $\hat{H}$. From inequality (6), it can be shown, taking into account the residual correlations between wavelet coefficients, that [6, 9]:

$$\text{Var}\ \hat{H} \le Cn^{-1/(1+1/(\alpha(N-H)))}. \quad (8)$$

Assuming exact decorrelation (**HYP1**) among the $\log_2|d_x(j,k)|$, one also has decorrelation between the Y_j, and one obtains:

$$\begin{aligned} \text{Var}\ \hat{H} &\simeq \textstyle\sum_j w_j^2 \text{Var}\, Y_j \\ &\simeq (\log_2(e))^2\pi^2(1+2/\alpha^2)(\textstyle\sum_j w_j^2/n_j)/12. \end{aligned} \quad (9)$$

These results require the following comments.

1. For a given α, the variance is minimum when $\sum_j w_j^2/n_j$ is minimum, which can be obtained using $a_j = \text{Var}\, Y_j \sim n_j^{-1}$. Such a choice has always been made here.

2. Var $\hat{H}$ all the less depend on H as N is increased. Under **HYP1**, it does not depend on H.

3. Neglecting border effects, n_j basically behaves as $n_j = n2^{-j}$. Therefore, $\sum_j w_j^2/n_j$ reduces to $(\sum_j w_j^2 2^j)/n$ which shows that, under **HYP1**, the variance of the estimate exhibits the standard $1/n$ decrease. The prefactor $\sum_j w_j^2/n_j$ is actually also depending on n through the number of available scales, but this dependence is weak. Such a $1/n$ decrease of the variance is obtained regardless of the possibly LRD nature of the process, which is not a trivial result, see e.g., [3]. A careful examination of the upper bound (8) however shows that the decrease of Var $\hat{H}$ depends both on H and α and that an increase of N reduces this dependence not only with respect to H, but also with respect to α.

4. Under **HYP1**, the variance of H explicitly depend, through the prefactor, on α which can be an unknown parameter. It could however be estimated using Eq. (7) above. The performances of such a joint (α, H) estimator are under current analysis.

Fig. 2 (b) compares the theoretical (Eq. (9)) and numerically obtained variances of $\hat{H}$ as functions of n. It shows that **HYP1** used in the analytical derivation of Var $\hat{H}$ is relevant.

Confidence intervals for $\hat{H}$. Fig. 2 (c), obtained from numerical simulations as described above, shows that the random variable $\hat{H}$ exhibits a probability density function that is very close to that of a Gaussian. In [6], theoretical arguments allow us to believe, that under mild conditions on the mother wavelet, $\hat{H}$ should asymptotically follow a Gaussian law, but no proof is available yet. From this assumed asymptotic Gaussianity, confidence intervals for $\hat{H}$ can be derived using Eq. (9): $\hat{H} \pm z_{\beta,2}\sqrt{\text{Var}\ \hat{H}}$. This is another interesting feature of our estimator because it allows to qualify the credit that can be given to the estimated value. Again, deriving these confidence intervals requires the knowledge of α, indicating that achieving a joint (α, H) estimate is crucial.

5. CONCLUSION

We proposed here a wavelet based estimator for the self similar parameter of SαS self-similar processes and we showed (theoretically and numerically) that it exhibits excellent statistical performances (unbiasedness, $1/n$ variance decrease, asymptotic Gaussianity). We explained how this directly results from an intimate adequacy between self-similarity and the wavelet transform.

This estimator applies in fact, without modification, to any other α-stable self similar processes (not necessarily SαS or LFSM) as well as to other possibly non-Gaussian self-similar processes with finite second-order statistics: it will remain unbiased since, as shown, unbiasedness only depend on self similarity and asymptotically Gaussian and efficient (the value of the variance and therefore of the confidence intervals will change with probability density functions of the analyzed processes) [8].

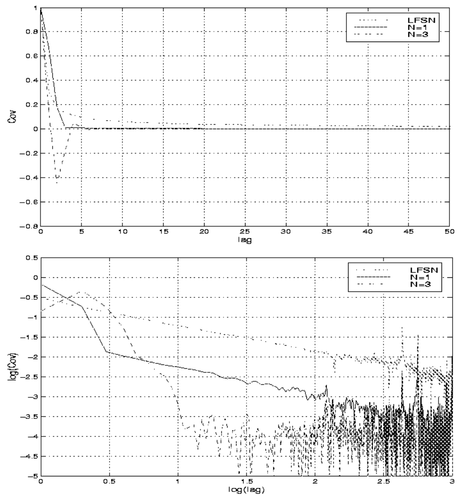

Figure 1: **Covariances of the** log **of the increments of a LFSM and of the log of its wavelet coefficients.** Autocovariance functions of $\log|y|$ where y is the increment process (LFSN) of the LFSM x, of $\log|d_x(j,k)|$ at fixed octave j using Daubechies1 and a Daubechies3 wavelets. Top, lin-lin plot; bottom, log-log plot of the absolute value of the covariance. It shows that the range of correlation is significantly decreased when the number N of vanishing moments of the mother wavelet is increased.

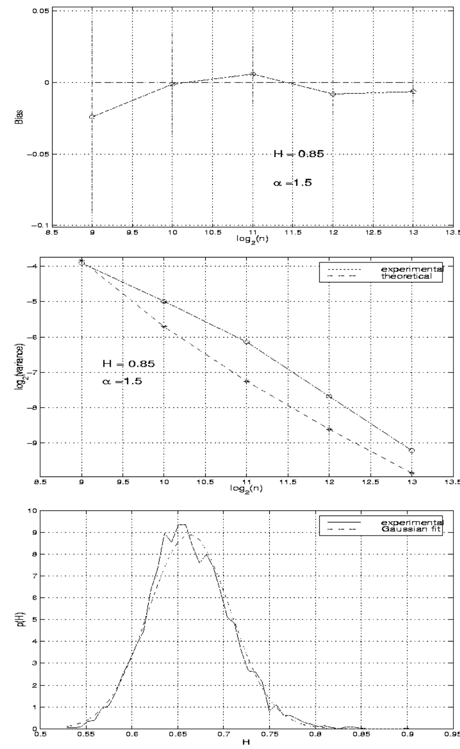

Figure 2: **Statistical performance of $\hat{H}$.** Top: bias (as a function of n), with 95% confidence intervals, shows that $\hat{H}$ is unbiased even for small length of data; middle: theoretical (derived under exact decorrelation hypothesis **HYP1**) and experimental variances (as functions of n) are very close one of the other and decrease with the standard $1/n$ behaviour despite LRD in x; bottom: the estimated probability density function of $\hat{H}$ is very close to a Gaussian distribution.

6. REFERENCES

[1] P. Abry, P. Gonçalvès and P. Flandrin, Wavelets, spectrum estimation and $1/f$ processes, in A. Antoniadis and G. Oppenheim, eds, *Wavelets and Statistics*, *Lectures Note in Statistics* **103**, pp. 15–30. Springer-Verlag, New York, 1995.

[2] P. Abry, D. Veitch, Wavelet analysis of long-range dependent traffic, *IEEE Trans. on Info. Theory*, 44(1):2–15, 1998.

[3] J. Beran, *Statistics for Long-Memory Processes*. Chapman and Hall, New York, 1994.

[4] S. Cambanis and C. Houdré, On the continuous wavelet transform of second-order random processes. In *IEEE Trans. on Info. Theory*, Vol. 41, No. 3, May 1995, pp. 628–642.

[5] I. Daubechies, *Ten Lectures on Wavelets*. SIAM, Philadelphia (PA), 1992.

[6] L. Delbeke, *Wavelet based estimators for the scaling index of a self-similar with stationary increments*, PhD Thesis, KU Leuven, Belgium, 1998.

[7] L. Delbeke, P. Abry, Stochastic integral representation and properties of the wavelet coefficients of linear fractional stable motion, submitted to *Stochastic Processes and their Applications*, preprint, 1997.

[8] L. Delbeke, P. Abry, Wavelet based estimators for the self-similarity parameter of the fractional Brownian motion, submitted to *Applied and Computational Harmonic Analysis*, preprint, 1998.

[9] L. Delbeke, J. Segers, The covariance of the logarithm of jointly symmetric stable random variables, preprint, 1998.

[10] P. Flandrin, Wavelet analysis and synthesis of fractional Brownian motion, *IEEE Trans. on Info. Theory*, 38:910-917, 1992.

[11] E. Masry, The wavelet transform of stochastic processes with stationary increments and its application to fractional Brownian motion. *IEEE Trans. on Info. Theory*, Vol. 39, No. 1, January 1993, pp. 260–264.

[12] C.L. Nikias, M. Shao, *Signal processing with Alpha-Stable distributions and applications*. John Wiley and Sons, Inc., New York, 1995.

[13] B. Pesquet-Popescu, Statistical properties of the wavelet decomposition of some non Gaussian self-similar processes, submitted to *Signal Processing*, preprint, 1997.

[14] G. Samorodnitsky, M. S. Taqqu, *Stable Non-Gaussian Processes: Stochastic Models with Infinite Variance*. Chapman and Hall, New York, London, 1994.

[15] A.H. Tewfik, M. Kim, Correlation structure of the discrete wavelet coefficients of fractional Brownian motion, *IEEE Trans. Info. Theory*, 38:904–909, 1992.