Wavelet Tools for Scaling Processes — 1.

Patrick Flandrin

CNRS — École Normale Supérieure de Lyon, France
flandrin@ens-lyon.fr

Cergy, 29 mars 2006

Power laws and scaling

- **Power-law spectra.** Power-laws correspond to homogeneous functions:

$$\mathcal{S}(f) = C\,|f|^{-\alpha} \Rightarrow \mathcal{S}(kf) = C\,|kf|^{-\alpha} = k^{-\alpha}\mathcal{S}(f),$$

for any $k > 0$

- **Fourier transform.** Frequency scaling carries over to the time domain. If we let $s(t) := (\mathcal{F}^{-1}\mathcal{S})(f)$, we get:

$$\int_{-\infty}^{+\infty} \mathcal{S}(kf)\, e^{i2\pi ft}\, df = k^{-1} \int_{-\infty}^{+\infty} \mathcal{S}(f')\, e^{i2\pi f'(t/k)}\, df' = s(t/k)/k$$

It follows that $s(t/k) = s(t)/k^{\alpha-1} \Rightarrow$ self-similarity

Intuitive "self-similarity"

Beyond intuition

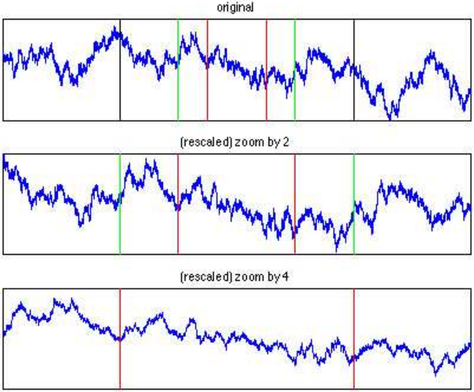
Definition

A process $\{X(t), t \in \mathbb{R}\}$ is said to be self-similar of index H (or "H-ss") if, for any $k > 0$,

$$\{X(kt), t \in \mathbb{R}\} \stackrel{d}{=} k^H \{X(t), t \in \mathbb{R}\}$$

- Invariance of statistical properties under dilations in time, up to a renormalization in amplitude ("self-affinity")
- Any zoomed (in or out) version of an H-ss process looks (statistically) the same $\Rightarrow$ no characteristic scale

Zooming



Self-similarity vs. stationarity 1.

Theorem

If a process X is self-similar, it is necessarily nonstationary

Proof.

Assuming that $\text{Var}\{X(t=1)\} \neq 0$, we have, for any $t > 0$,

$$\text{Var}\{X(t)\} = \text{Var}\{X(t \times 1)\} = t^{2H}\,\text{Var}\{X(1)\} \neq Const.$$

$\square$

Self-similarity vs. stationarity 2.

Theorem (Lamperti, 1962)

Stationary processes can be attached to self-similar processes, and vice-versa:

- *if* $\{X(t), t > 0\}$ *is* H*-ss, then* $\{Y(t) := e^{-Ht} X(e^t), t \in \mathbb{R}\}$ *is (strictly) stationary*
- *conversely, if* $\{Y(t), t \in \mathbb{R}\}$ *is (strictly) stationary, then* $\{X(t) := t^H Y(\log t), t > 0\}$ *is* H*-ss*

Stationary increments

Definition

A process $\{X(t), t \in \mathbb{R}\}$ is said to have stationary increments if and only if, for any $\theta \in \mathbb{R}$, the increment process:

$$\left\{X^{(\theta)}(t) := X(t+\theta) - X(t), t \in \mathbb{R}\right\}$$

has a distributional law which does not depend upon t

- **Extension.** The concept of stationary increments can be naturally extended to higher orders ("increments of increments")

Self-similarity and stationary increments

Definition

H-ss processes with stationary increments are referred to as "H-sssi" processes

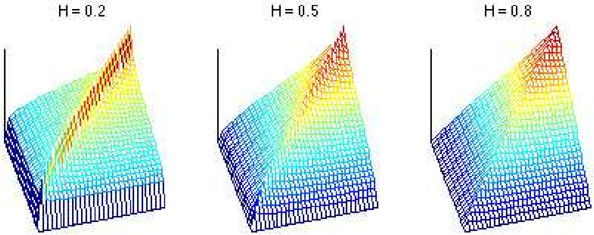
Theorem

The structure of the covariance function is the same for all H-sssi processes and reads

$$\mathbb{E}\{X(t)X(s)\} = \frac{Var\{X(1)\}}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right)$$

Proof

Covariance function of H-sssi processes



Fractional Brownian motion

Definition (Mandelbrot & van Ness, 1968)

A process $B_H(t)$ is referred to as a fractional Brownian motion (fBm) of index $0 < H < 1$, if and only if it is H-sssi and Gaussian

- fBm is an extension (anomalous diffusion) of the ordinary Brownian motion $B(t) \equiv B_H(t)|_{H=1/2}$
- The index H is referred to as the Hurst exponent, and its limited range guarantees the non-degeneracy ($H < 1$) and the mean-square continuity ($H > 0$) of fBm

Moving average

Definition (Mandelbrot & van Ness, 1968)

fBm admits the moving average representation:

$$B_H(t) - B_H(0) = \frac{1}{\Gamma(H+\frac{1}{2})} \left\{ \int_{-\infty}^{0} \left[(t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}} \right] B(ds) \right.$$

$$\left. + \int_0^t (t-s)^{H-\frac{1}{2}} B(ds) \right\}$$

- fBm results from a "fractional integration" of white noise
- no specific role attached to time $t = 0$

Harmonizability

Theorem

fBm admits the harmonizable spectral representation:

$$B_H(t) = C \,.\int_{-\infty}^{+\infty} |f|^{-(H+\frac{1}{2})}\,(e^{i2\pi tf} - 1)\,W(df),$$

with $W(df)$ the Wiener measure

- The "average spectrum" of fBm behaves as $|f|^{-(2H+1)}$
- fBm is a widespread model for (nonstationary) Gaussian processes with a power-law (empirical) spectrum

Fractional Gaussian noise 1.

Definition (Mandelbrot & van Ness, 1968)

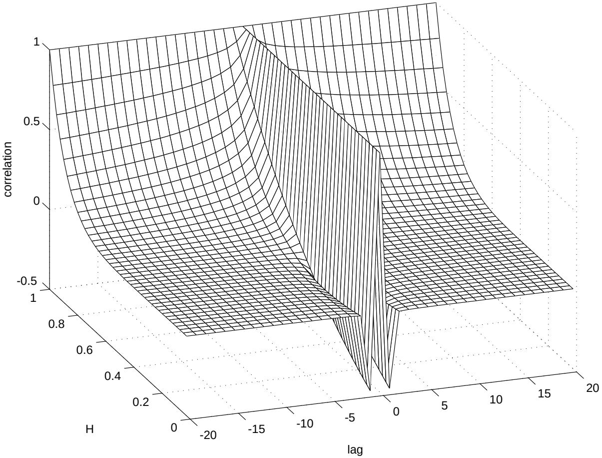
The (stationary) increment process $B_H^{(\theta)}(t)$ of fBm $B_H(t)$ is referred to as fractional Gaussian noise (fGn)

- **Autocorrelation.** The (stationary) autocorrelation function of fGn, $c_H(\tau) := \mathbb{E}\{B_H^{(\theta)}(t)B_H^{(\theta)}(t+\tau)\}$, reads:

$$c_H(\tau) = \frac{\sigma^2}{2}\left(|\tau+\theta|^{2H} - 2|\tau|^{2H} + |\tau-\theta|^{2H}\right).$$

- **White noise.** $\theta = 1$ and $H = \frac{1}{2} \Rightarrow c_H(k) = \sigma^2\,\delta(k), k \in \mathbb{Z}$
- **Asymptotics.** $\tau \to \infty \Rightarrow c_H(\tau) \sim \sigma^2\theta^2 H(2H-1)\tau^{2(H-1)}$ (subexponential, power-law decay)

Autocorrelation function of fGn



Spectrum of fGn 1.

- **Power Spectral Density.** If $\theta = 1$ (and, hence, $-\frac{1}{2} \leq f \leq +\frac{1}{2}$), the PSD of discrete-time fGn is given by:

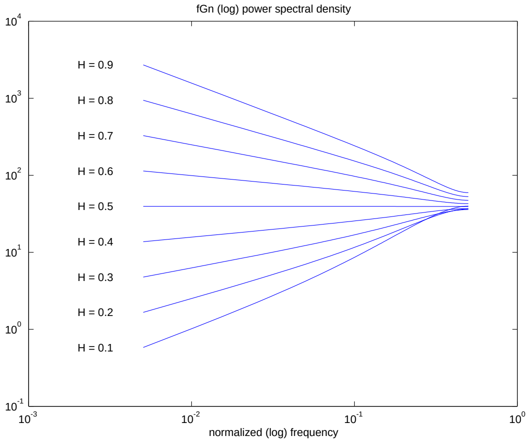
$$\mathcal{S}(f) = C\,\sigma^2\,|e^{i2\pi f} - 1|^2 \sum_{k=-\infty}^{\infty} \frac{1}{|f+k|^{2H+1}}$$

Fact

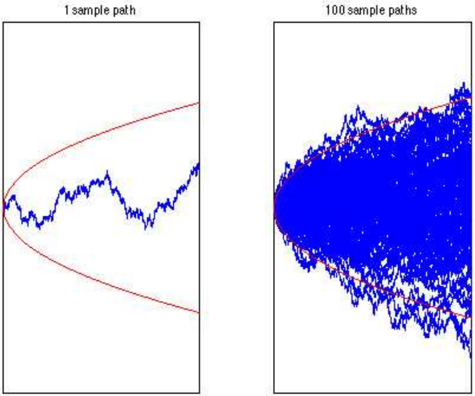
We observe that $\mathcal{S}(f) \sim C\,\sigma^2\,|f|^{1-2H}$ *when* $f \to 0$*:*

- $0 < H < \frac{1}{2} \Rightarrow \mathcal{S}(0) = 0$
- $\frac{1}{2} < H < 1 \Rightarrow \mathcal{S}(0) = \infty$ (*spectral divergence*)

Spectrum of fGn 2.



Sample paths of Bm



Sample paths of fBm

- **Local regularity.** For any (small enough) $\epsilon > 0$ and any $t \in \mathbb{R}$, we have $|B_H^{(\epsilon)}(t)| \leq C\, |\epsilon|^H$, with probability 1

Theorem

fBm is everywhere continuous but nowhere differentiable, and its sample paths have a uniform (Haussdorf and box) fractal dimension $\dim \mathrm{graph}\, B_H = 2 - H$

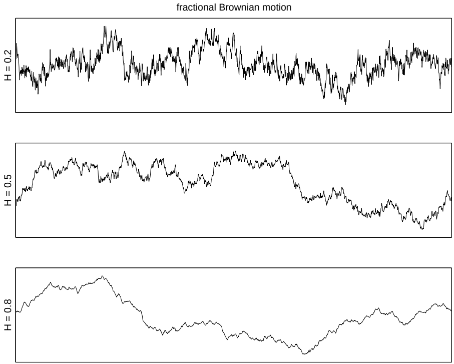
From Bm to fBm

- **Correlation between increments.** It follows from the covariance structure of fBm that, for any $t \in \mathbb{R}$,

$$C_H(\theta) := -\frac{\mathbb{E}\{B_H^{(-\theta)}(t)\, B_H^{(\theta)}(t)\}}{\mathrm{Var} B_H^{(\pm\theta)}(t)} = 2^{2H-1} - 1$$

  - $H = \frac{1}{2}$: no correlation (Brownian motion, $D = 1.5$)
  - $H < \frac{1}{2}$: negative correlation (more erratic, $\lim_{H \to 0} D = 2$)
  - $H > \frac{1}{2}$: positive correlation (less erratic, $\lim_{H \to 1} D = 1$)
- **Interpretation.** H is a roughness measure of sample paths.

H as a roughness measure



Long-range dependence

Definition

A stationary process $\{X(t), t \in \mathbb{R}\}$ is said to be asymptotically self-similar of index $\beta \in (0,1)$ if

$$(\mathrm{Var}\{X(t)\})^{-1}\,\mathbb{E}\{X(t)X(t+\tau)\} \sim \tau^{-\beta}$$

when $\tau \to \infty$

- H-sssi processes are asymptotically self-similar of index $\beta = 2(1-H)$ (example: fGn with $\frac{1}{2} < H < 1$)
- non-summability (and power-law decay) of the autocorrelation $\Rightarrow$ (power-law) divergence of the PSD at $f = 0$
- asymptotic self-similarity $\Rightarrow$ long-range dependence (LRD) (also referred to as long memory)

fGn as a limit

Definition

Given a stationary process $\{X(n), n \in \mathbb{Z}\}$, the recomposition rule

$$X(n) \mapsto X^T(n) := \frac{1}{T} \sum_{k=(n-1)T+1}^{nT} X(k)$$

is referred to as aggregation over T

Theorem

- *renormalized by T^{H-1}, fGn is invariant under aggregation*
- *as $T \to \infty$, aggregating any asymptotically H-ss process ends up with a process whose covariance structure is that of fGn*

"$1/f$" processes

Definition

A process is said to be of "$1/f$"-type if its empirical PSD behaves as $f^{-\alpha}$ ($\alpha > 0$) over some frequency range $[A, B]$

- **Special cases.** Depending on A and B, one can end up with:
  - LRD, if $A \to 0$ and $B < \infty$
  - scaling in some "inertial range", if $0 < A < B < \infty$
  - small-scale fractality, if $A < \infty$ and $B \to \infty$
- **Remark.** In the fBm/fGn case, the only Hurst exponent H controls all 3 situations

Evidencing scaling in data? 1.

Fact

Different and complementary signatures of scaling can be observed with respect to time (sample paths, correlation, increments . . .) or frequency/scale (spectrum, zooming . . .).

Idea

Use explicitly an approach which combines time and frequency/scale $\Rightarrow$ *wavelets!*

Evidencing scaling in data? 2.

Fact

Iterating aggregation reveals scale invariance

Idea

Use explicitly a multiresolution approach $\Rightarrow$ *wavelets!*

Multiresolution analysis 1.

Idea

"signal = (low-pass) approximation + (high-pass) detail"

$+$

iteration

- Successive approximations (at coarser and coarser resolutions) $\sim$ aggregated data
- Details (information differences between successive resolutions) $\sim$ increments

Multiresolution analysis 2.

Definition (Mallat & Meyer, 1986)

A MultiResolution Analysis (MRA) of $L^2(\mathbb{R})$ is given by:

1. A hierarchical sequence of embedded approximation spaces $\ldots V_1 \subset V_0 \subset V_{-1} \ldots$, whose intersection is empty and whose closure is dense in $L^2(\mathbb{R})$
2. A dyadic two-scale relation between successive approximations:

$$X(t) \in V_j \Leftrightarrow X(2t) \in V_{j-1}$$

3. A scaling function $\varphi(t)$ such that all of its integer translates $\{\varphi(t-n), n \in \mathbb{Z}\}$ form a basis of V_0

Wavelet decomposition 1.

- **Signal expansion.** For a given resolution depth J, any signal $X(t) \in V_0$ can be expanded as:

$$\underbrace{X(t)}_{signal} = \underbrace{\sum_k a_X(J,k)\,\varphi_{J,k}(t)}_{approximation} + \underbrace{\sum_{j=1}^{J}}_{J\ octaves} \underbrace{\sum_k \overbrace{d_X(j,k)}^{wav.\,coeffs.}\ \psi_{j,k}(t)}_{details},$$

with $\{\xi_{j,k}(t) := 2^{-j/2}\,\zeta(2^{-j}t - k), j \text{ and } k \in \mathbb{Z}\}$, for $\xi = \varphi$ and ψ

Idea

The wavelet $\psi(.)$ is constructed in such a way that all of its integer translates form a basis of W_0, defined as the complement of V_0 in V_{-1}

Wavelet decomposition 2.

Definition

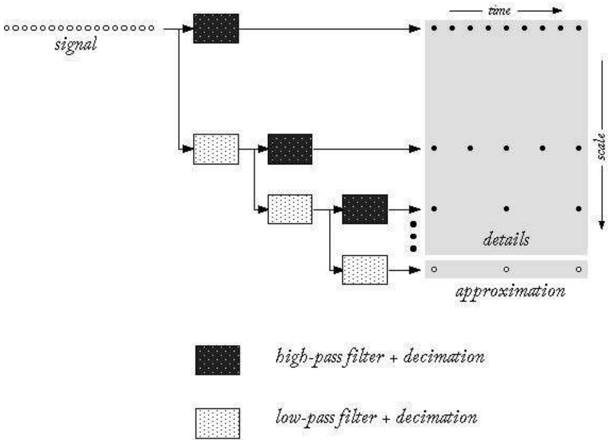
The wavelet coefficients $d_X(j,k)$ are given by the inner products:

$$d_X(j,k) := \langle X, \psi_{j,k} \rangle$$

- In practice, they can rather be computed in a recursive fashion, via efficient pyramidal algorithms (faster than FFT's)
- No need for knowing explicitly $\psi(t)$: enough to characterize a wavelet by its *filter coefficients* $\{g(n) := (-1)^n \, h(1-n), n \in \mathbb{Z}\}$, with

$$h(n) := \sqrt{2} \int_{-\infty}^{+\infty} \varphi(t) \, \varphi(2t-n) \, dt$$

Mallat's algorithm



Wavelet decomposition 3.

- **Example.** The simplest choice for a MRA is given by the Haar basis (Haar, 1911), attached to the scaling function $\varphi(t) = \chi_{[0,1[}(t)$ and wavelet $\psi(t) = \chi_{[0,1/2[}(t) - \chi_{[1/2,1[}(t)$
- **Remark.** When aggregated over dyadic intervals, data samples identify to Haar approximants
- **Interpretation.** Wavelet analysis offers a refined way of both aggregating data and computing increments

Wavelets as filters 1.

Result (Grossmann & Morlet, 1984)

By construction, a scaling function (resp., a wavelet) is a low-pass (resp., high-pass) function $\Rightarrow$ *an admissible wavelet* $\psi(t)$ *is necessarily zero-mean:*

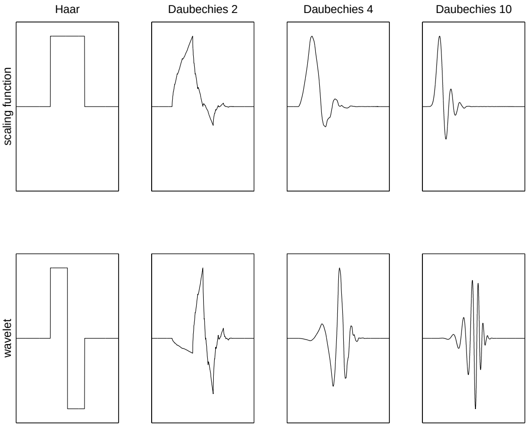
$$\Psi(0) := \int_{-\infty}^{+\infty} \psi(t)\, dt = 0$$

Definition

A further key property for a wavelet is the number of its vanishing moments, i.e., the integer $N \geq 1$ such that

$$\int_{-\infty}^{+\infty} t^k\, \psi(t)\, dt = 0, \;\; \text{for} \;\; k = 0, 1, \dots N-1$$

The example of Daubechies wavelets



Wavelets as filters 2.

- Given the statistics of the analyzed signal, statistics of its wavelet coefficients can be derived from input-ouput relationships of linear filters
- In the case of stationary processes with autocorrelation $\gamma_X(\tau) := \mathbb{E}\{X(t)X(t+\tau)\}$ and PSD $\Gamma_X(f)$, stationarity carries over to wavelet sequences and we end up with:

$$C_X(j,n) := \mathbb{E}\{d_X(j,k)d_X(j,k+n)\} = \int_{-\infty}^{+\infty} \gamma_X(\tau)\, \gamma_\psi(2^{-j}\tau+n)\, d\tau$$

$$\sum_{n=-\infty}^{\infty} C_X(j,n)\, e^{-i2\pi fn} = \Gamma_X(2^{-j}f) \times \underbrace{\sum_{n=-\infty}^{\infty} \gamma_\psi(n)\, e^{-i2\pi fn}}_{\textit{wavelet spectrum}}$$

Wavelets as stationarizers 1.

Theorem (F., 1989 & 1992)

Wavelet admissibility ($N \geq 1$) guarantees that, if $X(t)$ has stationary increments, then $d_X(j,k)$ is stationary in k, for any given scale 2^j

Proof

Wavelets as stationarizers 2.

- **Extension.** Stationarization can be extended to processes with stationary increments of order $p > 1$, under the vanishing moments condition $N \geq p$
- **Application.** Stationarization applies to H-sssi processes (e.g., fBm), with $\rho(t) = |t|^{2H}$;
- **Remark.** Nonstationarity is contained in the approximation sequence.

Self-similarity in wavelet space

Theorem

The multiresolution nature of wavelet analysis guarantees that, if $X(t)$ is H-ss, then

$$\{d_X(j,k), k \in \mathbb{Z}\} \stackrel{d}{=} 2^{j(H+1/2)} \{d_X(0,k), k \in \mathbb{Z}\}$$

for any $j \in \mathbb{Z}$

- **Spectral interpretation.** Given a "$1/f$" process, the wavelet tuning condition $N > (\alpha - 1)/2$ guarantees that

$$\mathcal{S}_X(f) \propto |f|^{-\alpha} \Rightarrow \mathbb{E}\{d_X^2(j,k)\} \propto 2^{j\alpha}$$

Wavelets as decorrelators 1.

Theorem (F., 1992; Tewfik & Kim, 1992)

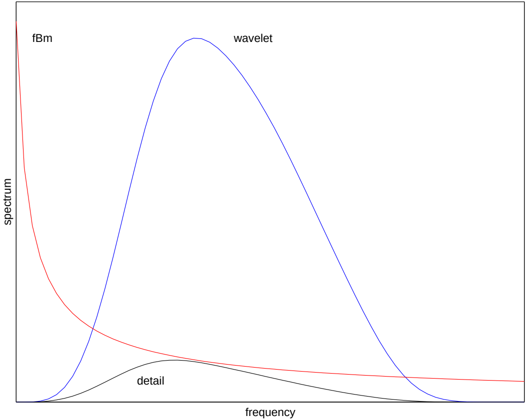
In the case where $X(t)$ is H-sssi, the condition $N > H + 1/2$ guarantees that wavelets coefficients are almost uncorrelated:

$$\mathbb{E}\{d_X(j,k)d_X(j,k+n)\} \sim n^{2(H-N)}, \ \ n \to \infty$$

- **Interpretation.** Competition, at $f = 0$, between the (divergent) spectrum of the process and the (vanishing) transfer function of the wavelet:

$$\mathbb{E}\{d_X(j,k)d_X(j,k+n)\} \propto \int_{-\infty}^{+\infty} \frac{|\Psi(2^j f)|^2}{|f|^{2H+1}} \, e^{i2\pi nf} \, df$$

Wavelets as decorrelators 2.



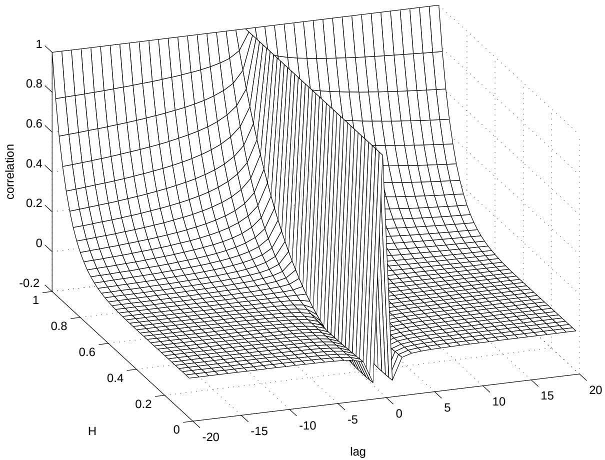
Wavelets as decorrelators 3.

Corollary

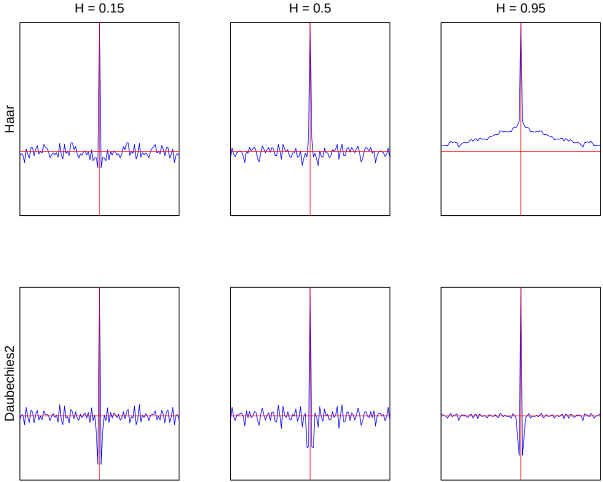
Long-range dependence (LRD) of a process $X(t)$ can be transformed into short-range dependence (SRD) in the space of its wavelet coefficients $d_X(j,.)$, provided that the number N of the vanishing moments is high enough

- **Remark.** Residual LRD in the approximation sequence
- **The case of H-sssi processes.** LRD when $H > 1/2 \Rightarrow$ SRD when $N > 1 \Rightarrow$ Haar not suitable

Wavelet correlation of fBm in the Haar case (theory)



Wavelet correlation and vanishing moments (experiment)



Rationale

Result

Given the variance $v_X(j) := \mathbb{E}\{d_X^2(j,k)\}$, scale invariance is revealed by the linear relation $\log_2 v_X(j) = \alpha\, j + Const.$

Idea (Abry, F. & Gonçalvès, 1995)

The further properties of 1) stationarization and 2) quasi-decorrelation suggest to use as estimator of $v_X(j)$ the empirical variance

$$\hat{v}_X(j) := \frac{1}{N_j} \sum_{k=1}^{N_j} d_X^2(j,k),$$

where N_0 stands for the data size and $N_j := 2^{-j} N_0$

LogScale Diagram

Definition (Abry & Veitch, 1998)

Given that $\log \mathbb{E}\{.\} \neq \mathbb{E}\{\log .\}$, the effective estimator ("LogScale Diagram") is $y_X(j) := \log_2 \hat{v}_X(j) - g(j)$, with

$$g(j) = \psi(N_j/2)/\log 2 - \log_2(N_j/2)$$

and $\psi(.)$ the derivative of the Gamma function

- **Bias.** $\mathbb{E}\{y_X(j)\} = \alpha j + \mathit{Const.}$: no bias in the *uncorrelated* case
- **Variance.** Assuming stationarization and quasi-decorrelation guarantees furthermore that

$$\sigma_j^2 := \mathrm{Var}\{y_X(j)\} = \zeta(2, N_j/2)/\log^2 2,$$

where $\zeta(z, \nu)$ is the generalized Riemann function

Scaling exponent estimation

- **From $y_X(j)$ to $\widehat{\alpha}$.** The slope α is estimated via a weighted linear regression in a log-log diagram:

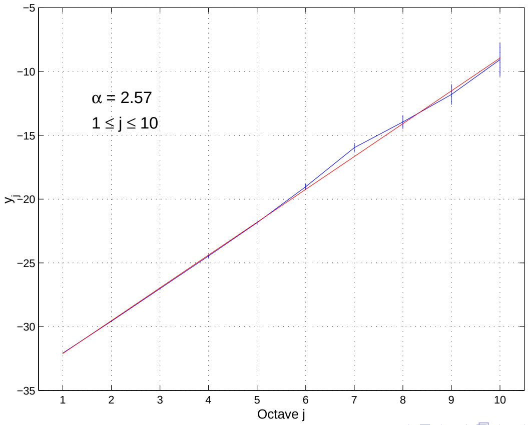
$$\widehat{\alpha} = \sum_{j=j_{min}}^{j_{max}} \frac{S_0\, j - S_1}{S_0\, S_2 - S_1^2} \, \frac{1}{\sigma_j^2} \, y_X(j),$$

  with $S_k := \sum_j k/\sigma_j^2$, $k = 0, 1, 2$
- **Bias and variance.** We have $\mathbb{E}\{\widehat{\alpha}\} \equiv \alpha$, by construction. Assuming Gaussianity, the estimator is moreover asymptotically efficient in the limit $N_j \to \infty$ (for any j), with

$$\mathrm{Var}\{\widehat{\alpha}\} \sim 1/N_0$$

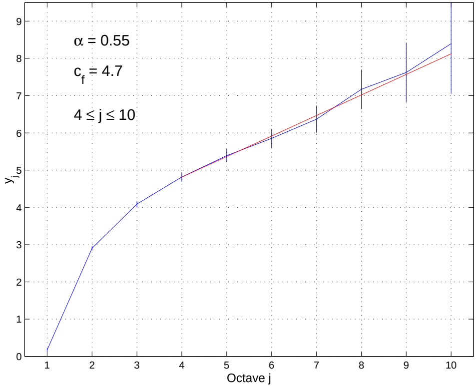
Example 1 (H-ss)

fBm with $H = 0.8 \Rightarrow \alpha = 2H + 1 = 2.6$



Example 2 (LRD)

FARIMA $(1, d, 0)$ with $d = 0.3 \Rightarrow \alpha = 2d = 0.6$



Robustness

- **Cancellation.** The vanishing moments condition

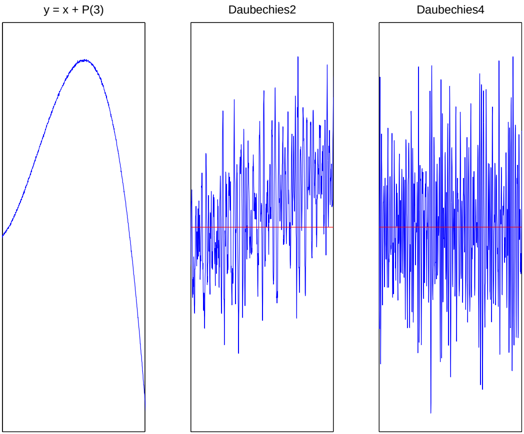
$$\int_{-\infty}^{+\infty} t^k \, \psi(t) \, dt = 0, \quad \textit{for} \quad k = 0, 1, \ldots N-1,$$

guarantees that $d_T(j,n) \equiv 0$ for any $T(t)$ of the form

$$T(t) = \sum_{k=0}^{N-1} a_k \, t^k$$

- **Interpretation.** A wavelet with enough vanishing moments makes the transform of $Z(t) := X(t) + T(t)$ blind to a superimposed polynomial trend

Robustness to polynomial trends



Wavelets and . . . 1.

- **Aggregation.** Wavelets offer a natural generalization to aggregation: Haar approximants $\mapsto$ Haar details $\mapsto$ wavelet details with higher N
- **Variogram.** — Wavelets generalize as well variogram techniques (Matheron, 1967), which are based on the increment property $\mathbb{E}\{(X(t+\tau) - X(t))^2\} = \sigma^2|\tau|^{2H}$, since increments can be viewed as constructed on the "poorman's wavelet":

$$\psi(t) := \delta(t+\tau) - \delta(t)$$

Wavelets and ... 2.

Definition (Allan, 1966)

A refined notion of variance — introduced in the study of atomic clocks stability — is the so-called Allan variance, defined by

$$\mathrm{Var}_X^{(Allan)}(T) := \frac{1}{2T^2}\,\mathbb{E}\left\{\int_{t-T}^{t} X(s)\,ds - \int_{t}^{t+T} X(s)\,ds\right\}^2$$

- In the case of H-ss processes, Allan variance is such that $\mathrm{Var}_X^{(Allan)}(T) \sim T^{2H}$ when $T \to \infty$
- When evaluated over dyadic intervals, Allan variance identifies to the variance of Haar details:

$$\mathrm{Var}_X^{(Allan)}(2^j) = \mathrm{Var}\{d_X^{(Haar)}(j,k)\}$$

Wavelets and ... 3.

Definition

In the case of a Poisson process $P(t)$ of counting process $N(.)$, one can define the Fano factor as:

$$F(T) := \mathrm{Var}\{N(T)\}/\mathbb{E}\{N(T)\}$$

- For a uniform density λ, we have $F(T) = 1$ for any T whereas, for a "fractal" density $\lambda(t) = \lambda + B_H^{(\theta)}(t)$, we have $F(T) \sim T^{2H-1}$ when $T \rightarrow \infty$
- Interpretation as fluctuations/average suggests the wavelet generalization given by:

$$F(T) \mapsto F_W(j) := 2^{j/2}\,\mathrm{Var}\{d_P(j,k)\}/\mathbb{E}\{a_P(j,k)\} \sim 2^{j(2H-1)}$$

when $j \rightarrow \infty$, and $F_W^{(Haar)}(j) \equiv F^{(Allan)}(2^j)$

Higher-order moments

- **Exact model**. LogScale Diagram 2nd order but H-ss $\Rightarrow$

$$\mathbb{E}\{|d_X(j,k)|^q\} \propto (2^j)^{Hq}$$

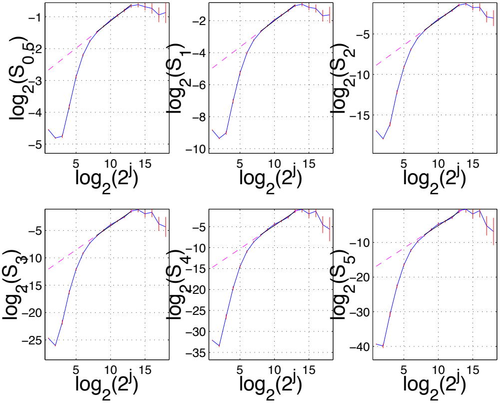
for any q (and all j's).

- **Variations**. Restrict scaling to intervals and/or make the scaling exponent a nonlinear function of q:

$$Hq \longrightarrow \zeta(q).$$

- **Issues**. Assessment? Models? Estimation?

Example (turbulence)



à suivre. . . (P. Abry)

Covariance of H-sssi processes

Proof.

Assuming that $X(t)$ is H-sssi, with $X(0) = 0$ and $X(1) \neq 0$, we have necessarily:

$$\begin{aligned}
\mathbb{E}X(t)X(s) &= \frac{1}{2}\left(\mathbb{E}X^2(t) + \mathbb{E}X^2(s) - \mathbb{E}\left(X(t) - X(s)\right)^2\right) \\
&= \frac{1}{2}\left(\mathbb{E}X^2(t) + \mathbb{E}X^2(s) - \mathbb{E}\left(X(t-s) - X(0)\right)^2\right) \\
&= \frac{\mathrm{Var}X(1)}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right).
\end{aligned}$$

$\square$

Back

Wavelets as stationarizers (1/3)

1. Assuming that $X(t)$ is a s.i. process with $X(0) = 0$ and $\mathrm{Var}\{X(t)\} := \rho(t)$, we have:

$$\begin{aligned}
\mathbb{E}\{X(t)X(s)\} &= \frac{1}{2}\left(\mathbb{E}\{X^2(t)\} + \mathbb{E}\{X^2(s)\} - \mathbb{E}\{(X(t) - X(s))^2\}\right) \\
&= \frac{1}{2}\left(\mathbb{E}\{X^2(t)\} + \mathbb{E}\{X^2(s)\} - \mathbb{E}\{(X(t-s) - X(0))^2\}\right) \\
&= \frac{1}{2}\left(\rho(t) + \rho(s) - \rho(t-s)\right)
\end{aligned}$$

Wavelets as stationarizers (2/3)

2. It follows that:

$$
\begin{aligned}
\mathbb{E}\{d_X(j,n)d_X(j,m)\} &= \int\int_{-\infty}^{+\infty} \mathbb{E}\{X(t)X(s)\}\,\psi_{jn}(t)\,\psi_{jm}(s)\,dt\,ds \\
&= \frac{1}{2}\int_{-\infty}^{+\infty} \rho(t)\,\psi_{jn}(t)\underbrace{\left(\int_{-\infty}^{+\infty} \psi_{jm}(s)\,ds\right)}_{=0}\,dt \\
&\quad +\frac{1}{2}\int_{-\infty}^{+\infty} \rho(s)\,\psi_{jm}(s)\underbrace{\left(\int_{-\infty}^{+\infty} \psi_{jn}(t)\,dt\right)}_{=0}\,ds \\
&\quad -\frac{1}{2}\int\int_{-\infty}^{+\infty} \rho(t-s)\,\psi_{jn}(t)\,\psi_{jm}(s)\,dt\,ds
\end{aligned}
$$

Wavelets as stationarizers (3/3)

3. And then:

$$\begin{aligned}\mathbb{E}\{d_X(j,n)d_X(j,m)\} &= -\frac{1}{2}\int\int_{-\infty}^{+\infty} \rho(t-s)\,\psi_{jn}(t)\,\psi_{jm}(s)\,dt\,ds \\ &= -\frac{1}{2}\int_{-\infty}^{+\infty} \rho(\tau)\,\gamma_\psi(2^{-j}\tau-(n-m))\,d\tau\end{aligned}$$

Back