Wavelet Tools for Scaling Processes*

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

†thanks to
Patrice Abry (CNRS–ENS Lyon),
Pierre Borgnat (ENS Lyon), Pierre Chainais (ENS Lyon),
Lieve Delbeke (KU, Leuven), Paulo Gonçalvès (INRIA Rhône-Alpes),
Stéphane Roux (ENS Lyon), Murad Taqqu (Boston Univ.),
and Darryl Veitch (Univ. of Melbourne)

*Lecture given at the Summer School "Turbulence: measurements and signals," Cargèse, May 2002.

The idea of "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*.

Scaling processes

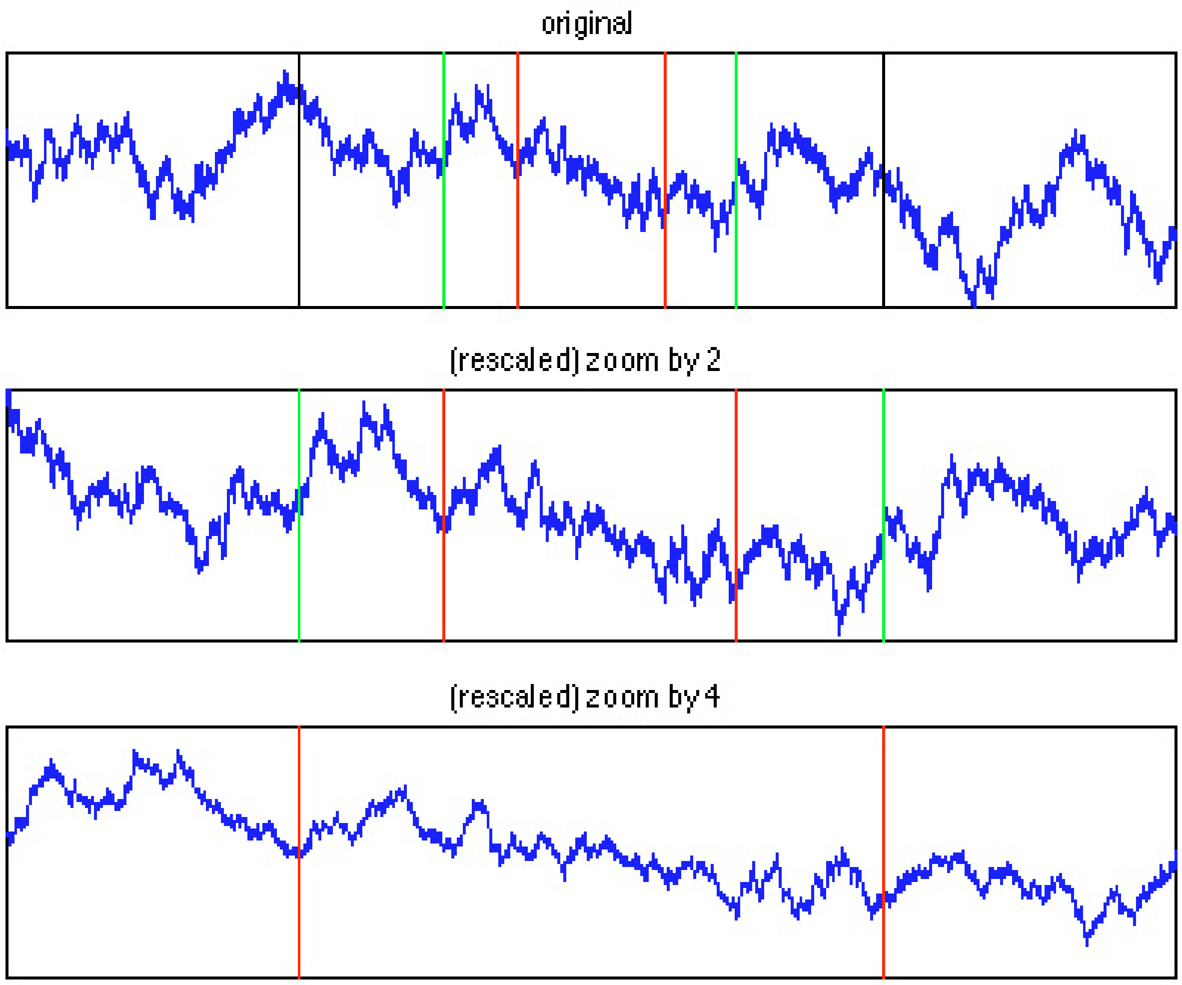
Self-similarity

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

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

- *Interpretation* — *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*.


original
(rescaled) zoom by 2
(rescaled) zoom by 4

Self-similarity vs. stationarity

- *Exclusion* — If a process X is self-similar, it is necessarily *nonstationary*. *Proof* — Assuming that Var$X(t=1) \neq 0$, we have, for any $t > 0$,

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

- *Transformations* — *Stationary* processes can be attached to *self-similar* processes, and vice-versa (Lamperti, 1962):

  - 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").

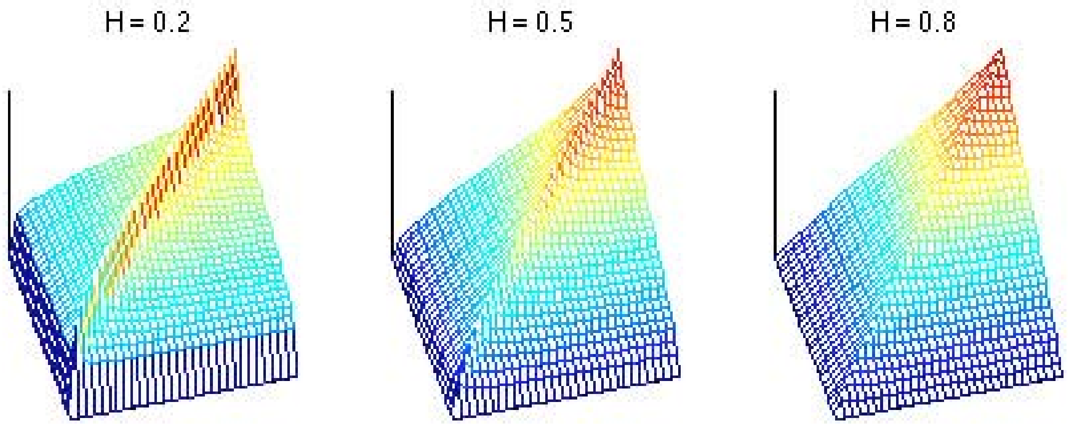
Self-similarity and stationary increments

- *Definition* — H-ss processes with stationary increments are referred to as "H-sssi" processes.

- *Covariance* — The structure of the covariance function is the same for all H-sssi processes. Indeed, assuming that $X(t)$ is H-sssi, with $X(0) = 0$ and $X(1) \neq 0$, we have necessarily:

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

Covariance function of H-sssi processes



Fractional Brownian motion — 1.

- *Definition 1* — 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 has been introduced in (Mandelbrot & van Ness, 1968), as an extension of the ordinary Brownian motion $B(t) \equiv B_H(t)|_{H=1/2}$ (*anomalous diffusion*).

  - 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.

Fractional Brownian motion — 2.

- *Definition 2* — fBm admits the *moving average representation*:

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

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

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

Fractional Brownian motion — 3.

- *Definition 3* — 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.

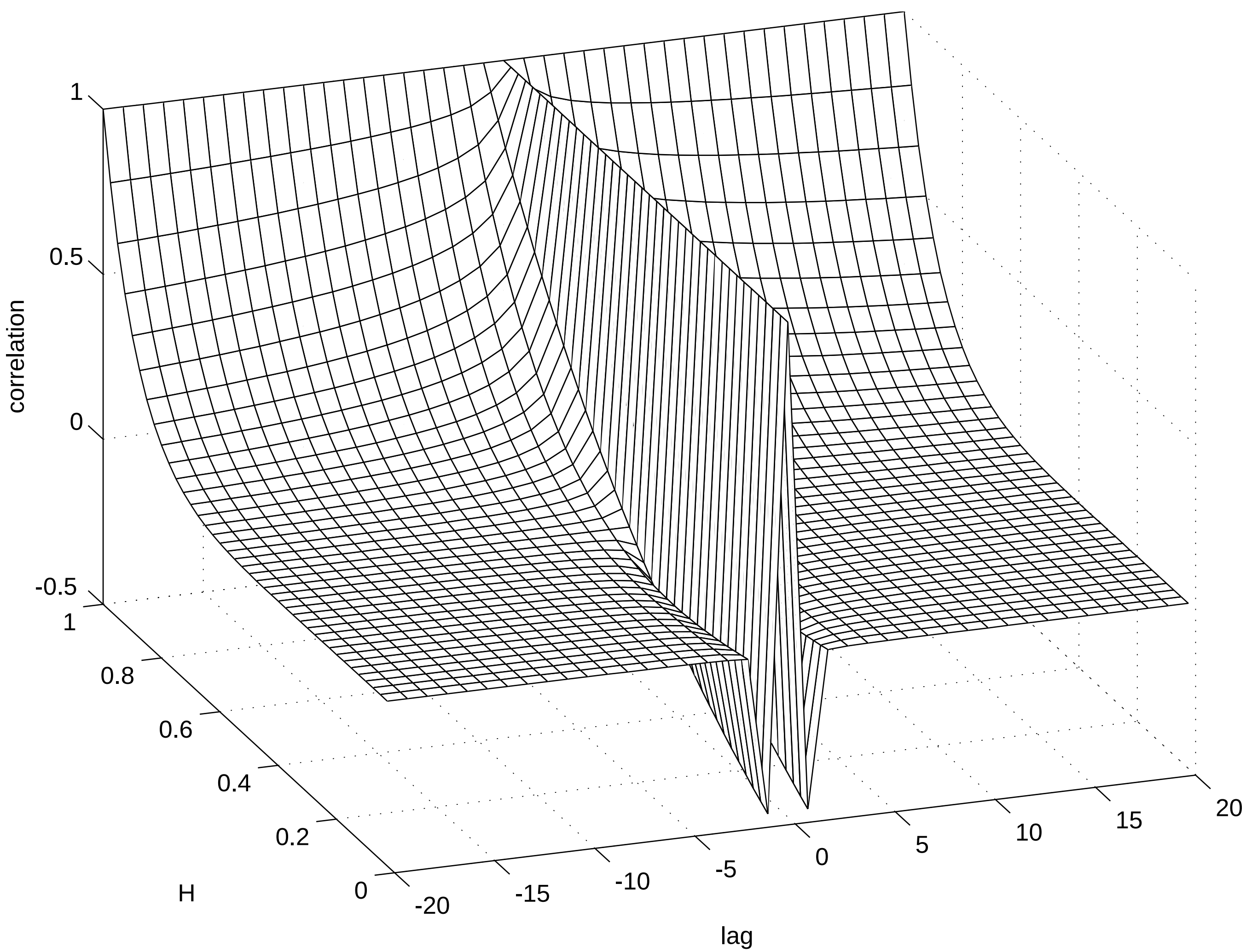
Fractional Gaussian noise — 1.

- *Definition* — 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:

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

  - if $\theta = 1$ and $H = \frac{1}{2}$, we have $c_H(k) = \sigma^2 \delta(k), k \in \mathbb{Z}$ (*discrete-time white noise*);
  - for large lags τ, one has $c_H(\tau) \sim \sigma^2 \theta^2 H(2H-1)\tau^{2(H-1)}$ (subexponential, *power-law* decay).


correlation
1
0.5
0
-0.5
H
1
0.8
0.6
0.4
0.2
0
lag
-20
-15
-10
-5
0
5
10
15
20

Fractional Gaussian noise — 2.

- *Spectrum* — If $\theta = 1$, the *power spectral density* of discrete-time fGn is given by:

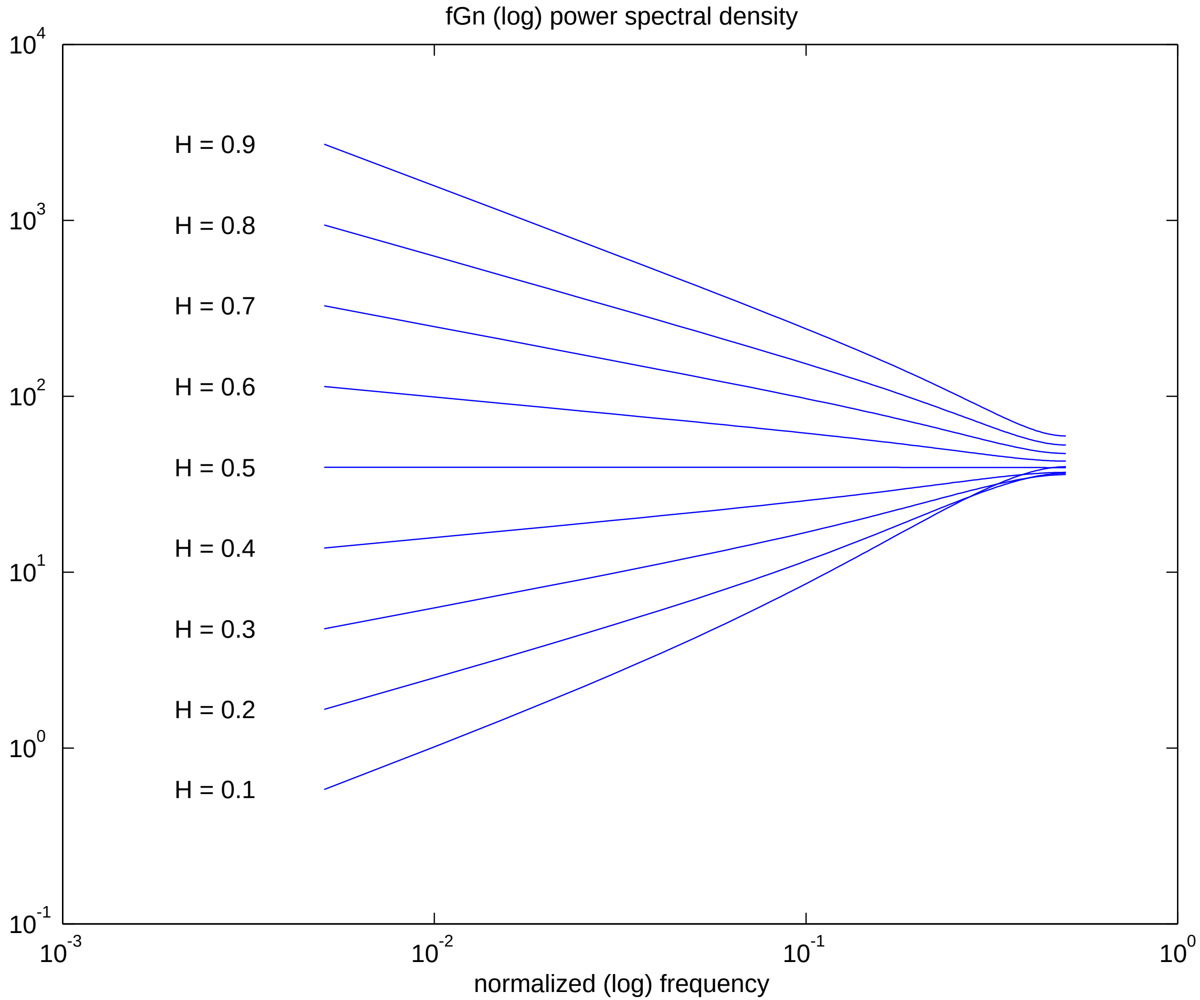
$$\mathcal{S}(f) = C\,\sigma^2\,|e^{i2\pi f} - 1|^2 \sum_{k=-\infty}^{\infty} \frac{1}{|f+k|^{2H+1}},$$

with $-\frac{1}{2} \leq f \leq +\frac{1}{2}$.

– if $H \neq \frac{1}{2}$, we have $\mathcal{S}(f) \sim C\,\sigma^2\,|f|^{1-2H}$ 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*).



fGn (log) power spectral density
H = 0.9
H = 0.8
H = 0.7
H = 0.6
H = 0.5
H = 0.4
H = 0.3
H = 0.2
H = 0.1
10^4
10^3
10^2
10^1
10^0
10^-1
10^-3
10^-2
10^-1
10^0
normalized (log) frequency

Fractional Gaussian noise — 3.

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

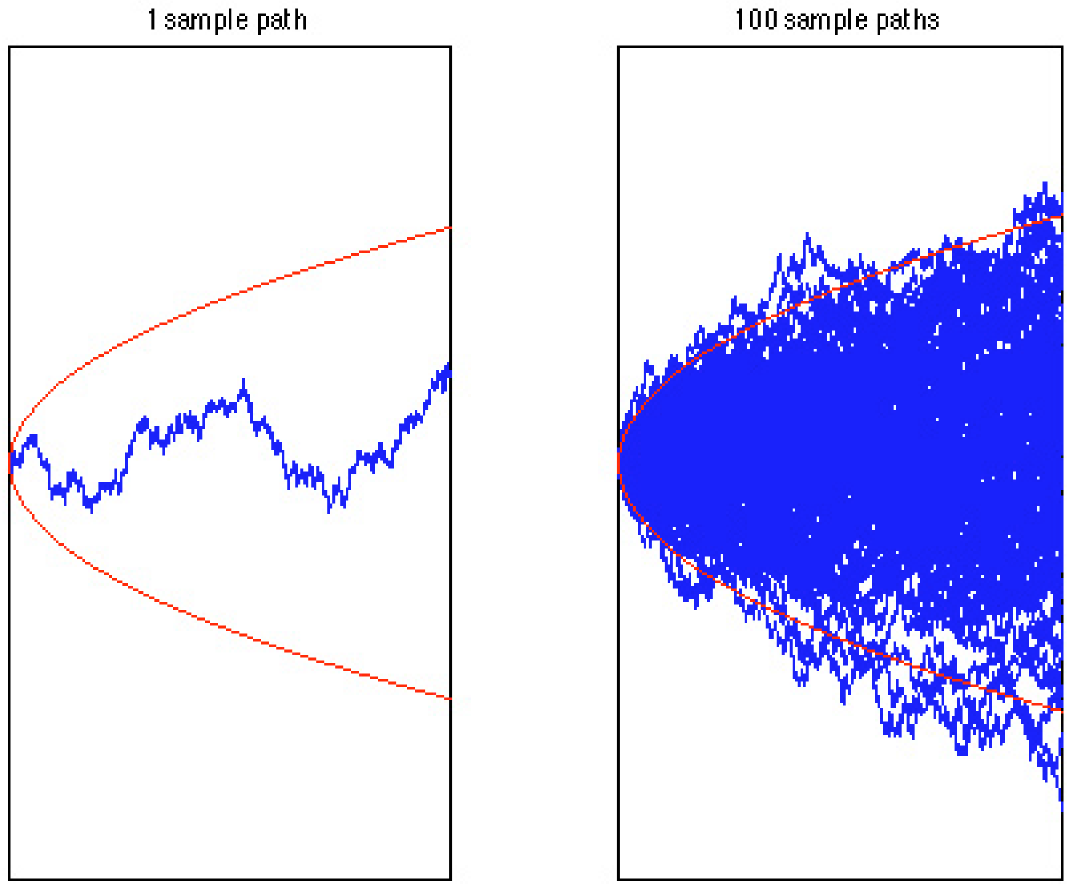
$$X(n) \mapsto X^T(n) := \frac{1}{T} \sum_{k=(n-1)T+1}^{nT} X(k)$$

is referred to as *aggregation* over T.

  – 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.

Sample paths — The example of Bm



Sample paths of fBm — 1.

- *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.
  - fBm is *everywhere continuous*, but *nowhere differentiable*;
  - sample paths have a uniform *Hölder regularity* $h < H$;
  - sample paths have a uniform (Haussdorf and box) *fractal dimension*: $\dim_B \text{graph}\, B_H = 2 - H$.

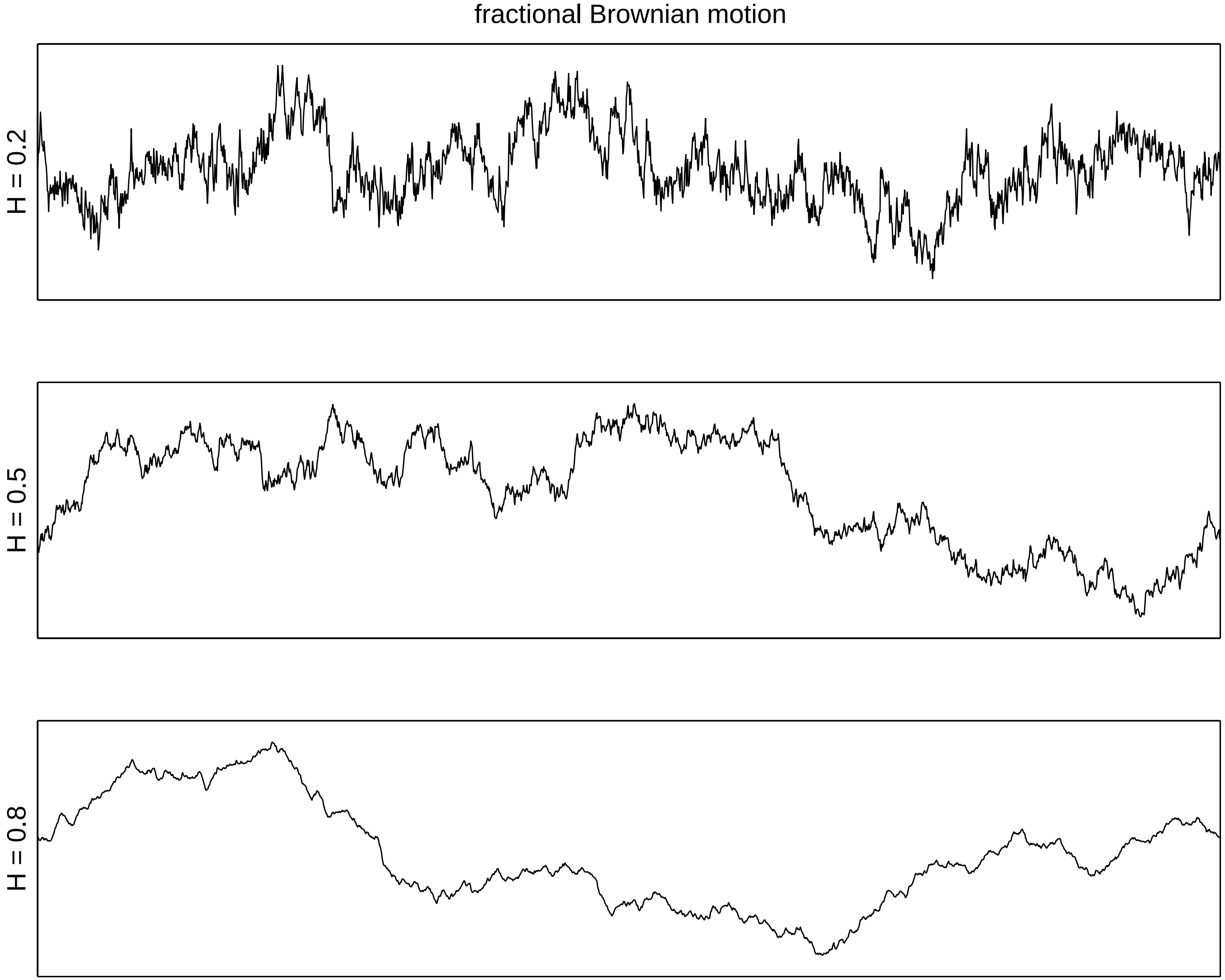
Sample paths of fBm — 2.

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

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

  - $H = \frac{1}{2}$: 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.


fractional Brownian motion
H = 0.2
H = 0.5
H = 0.8

Asymptotic self-similarity

- *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*).

“$1/f$” 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 $[A, B]$. 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 case, the only Hurst exponent H controls all 3 situations.

Evidencing scaling in data ? — 1.

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

- *Suggestion* — Use explicitly an approach which *combines* time and frequency/scale.

- *Formalization* — *Wavelets* !

Evidencing scaling in data ? — 2.

- *Fact* — *Iterating* aggregation reveals scale invariance.

- *Suggestion* — Use explicitly a *multiresolution* approach.

- *Formalization* — *Wavelets* !

Wavelets

Multiresolution analysis — 1.

"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 is a natural language for scaling processes.

Multiresolution analysis — 2.

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.

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

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

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.

- *Definition*. — 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.

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

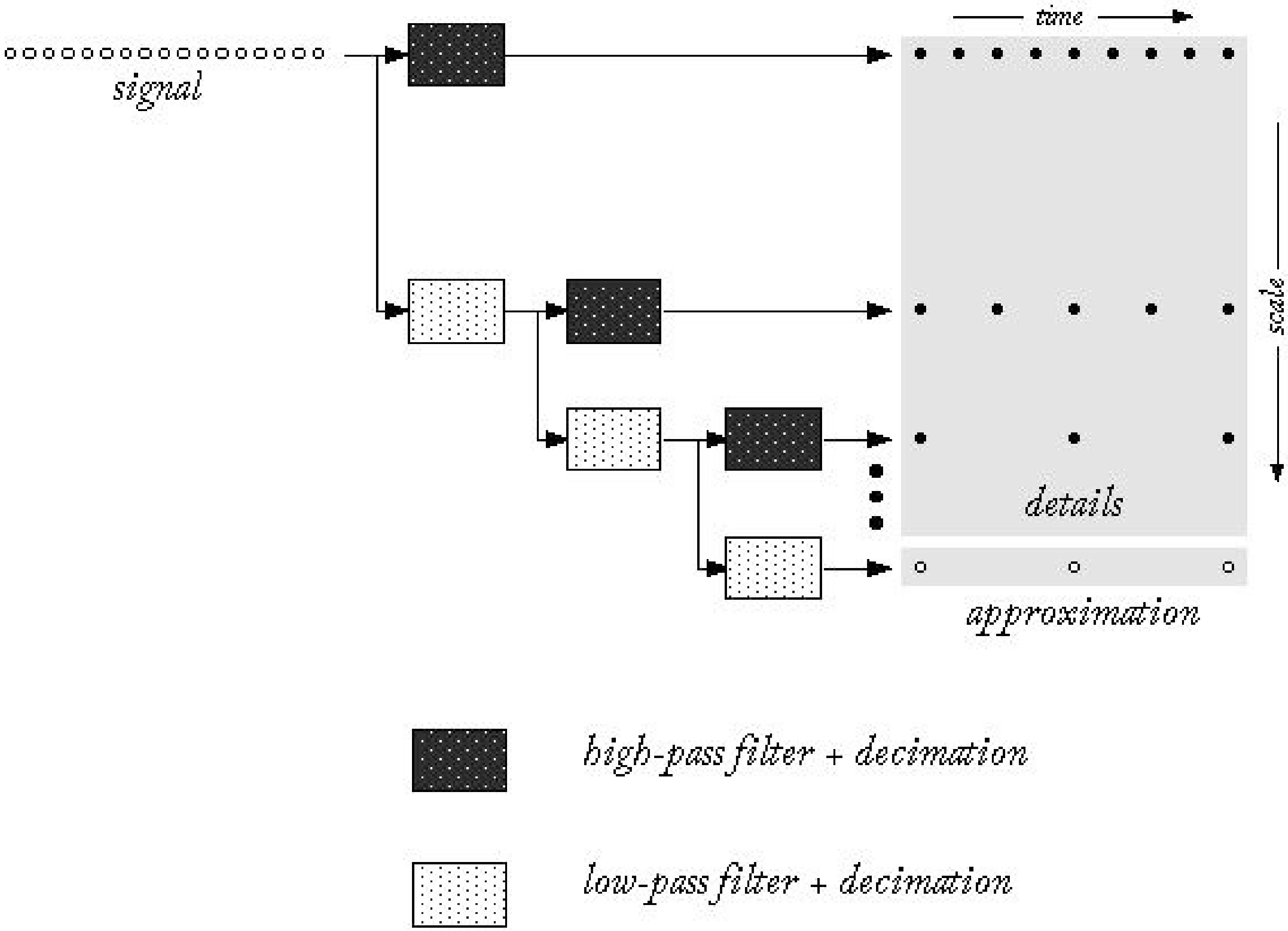
$$d_X(j,k) := \langle X, \psi_{j,k} \rangle \,.$$

- *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

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


signal
time
scale
details
approximation
high-pass filter + decimation
low-pass filter + decimation

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 the 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.

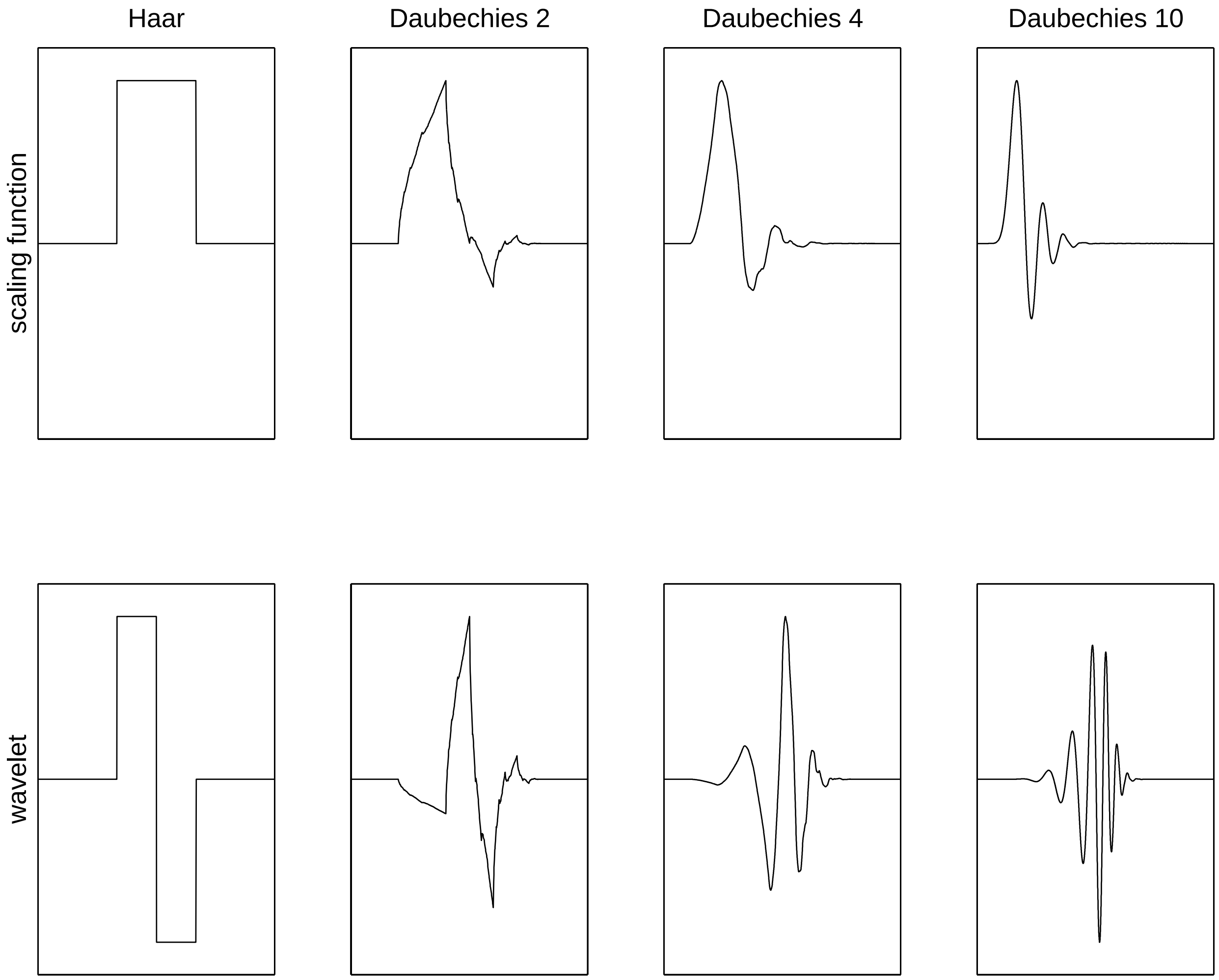
- *Admissibility* — By construction, a scaling function (resp., a wavelet) is a low-pass (resp., high-pass) function $\Rightarrow$ an *admissible* wavelet $\psi(t)$ is necessarily *zero-mean*:

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

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

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

The example of Daubechies wavelets



Wavelets as filters — 2.

- *Input-output* — Given the statistics of the analyzed signal, statistics of its wavelet coefficients can be derived from imput-ouput relationships of *linear filters*.

- *Stationary processes* — In the case of *stationary processes* with autocorrelation $\gamma_X(\tau) := \mathbb{E}X(t)X(t+\tau)$, stationarity carries over to wavelet sequences and we end up with:

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

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

Framework

Wavelets as stationarizers — 1.

- *Stationarization* — 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 — Assuming that $X(t)$ is a s.i. process with $X(0) = 0$ and Var $X(t) := \rho(t)$, we have

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

Wavelets as stationarizers — 2.

It follows that

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

Wavelets as stationarizers — 3.

- *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.

Wavelets and scale invariance

- *Self-similarity* — 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

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

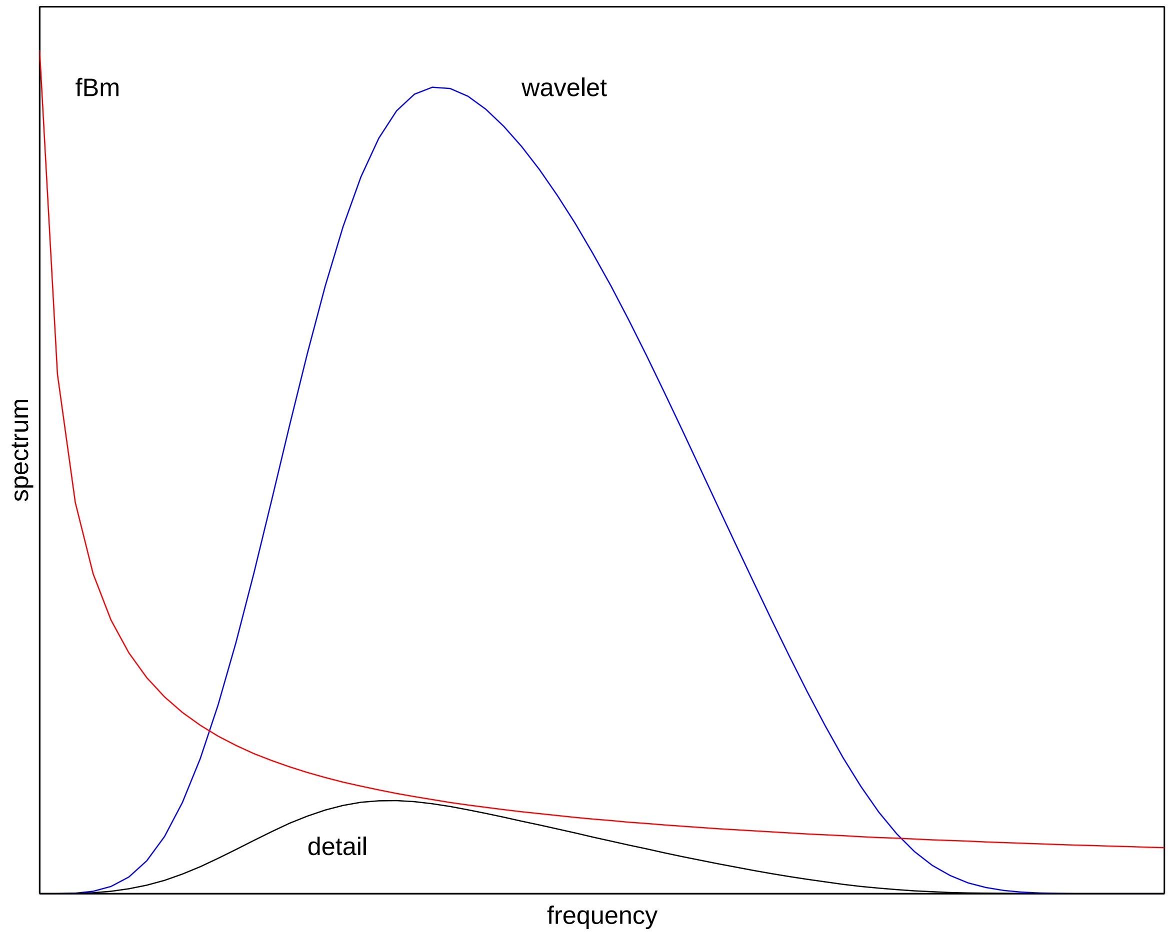
Wavelets as decorrelators — 1.

- *Quasi-decorrelation* — In the case where $X(t)$ is H-sssi, the condition $N > H + 1/2$ guarantees that

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

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

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


fBm
wavelet
detail
spectrum
frequency

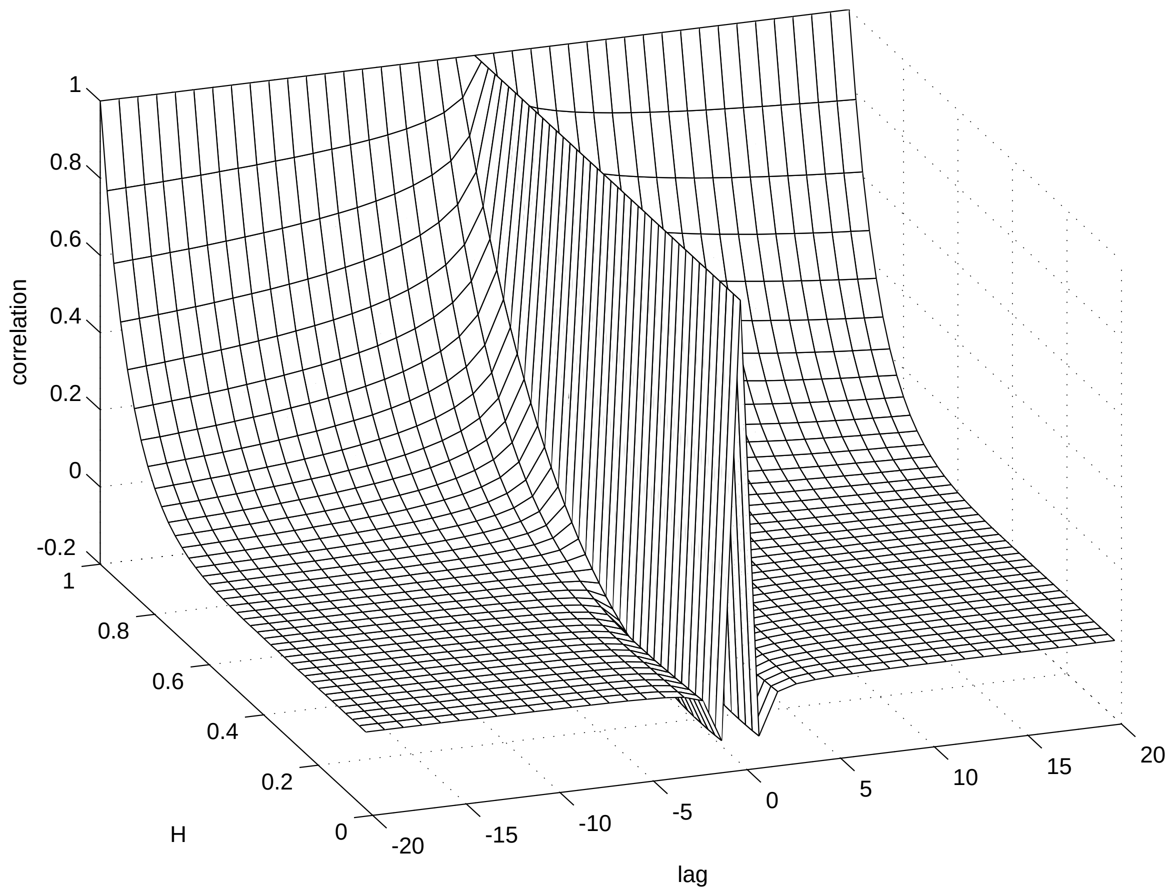
Wavelets as decorrelators — 2.

- *Consequence* — *Long-range dependence* (LRD) of a process X 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.

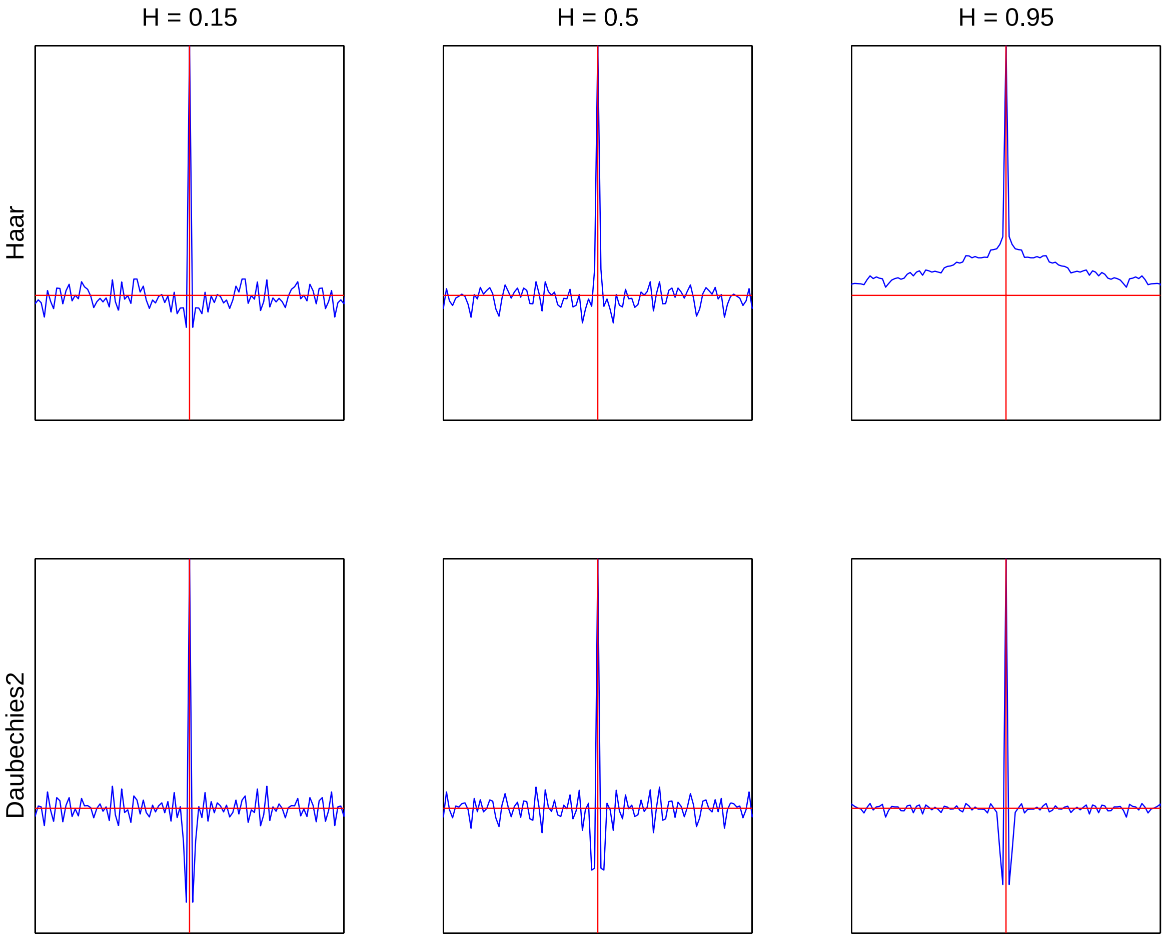
- *Remark* — *Residual* LRD in the approximation sequence.

- *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



Wavelet correlation and vanishing moments



Wavelets and scaling estimation — 1.

- *Theory* — Given the variance $v_X(j) := \mathbb{E} d_X^2(j,k)$, scale invariance is revealed by the *linear relation* :

$$\log_2 v_X(j) = \alpha\, j + \text{Const.}$$

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

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

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

Wavelets and scaling estimation — 2.

- *Bias correction* — Given that $\log \mathbb{E} \neq \mathbb{E} \log$, the effective estimator 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, so that $\mathbb{E} y_X(j) = \alpha j + \text{Const.}$ in the *uncorrelated* case.

- *Variance* — Assuming *stationarization* and *quasi-decorrelation* guarantees furthermore that

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

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

Wavelets and scaling estimation — 3.

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

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

with $S_k := \sum_j k/\sigma_j^2$, $k = 0, 1, 2$.

- *Bias and variance* — We have $\mathbb{E}\hat{\alpha} \equiv \alpha$, by construction. Assuming *Gaussianity*, the estimator is moreover *asymptotically efficient* in the limit $N_j \to \infty$ (for any j), with

$$\text{Var } \hat{\alpha} \sim 1/N_0.$$

Wavelets and scaling estimation — 4.

- *Robustness* — The *vanishing moments* condition

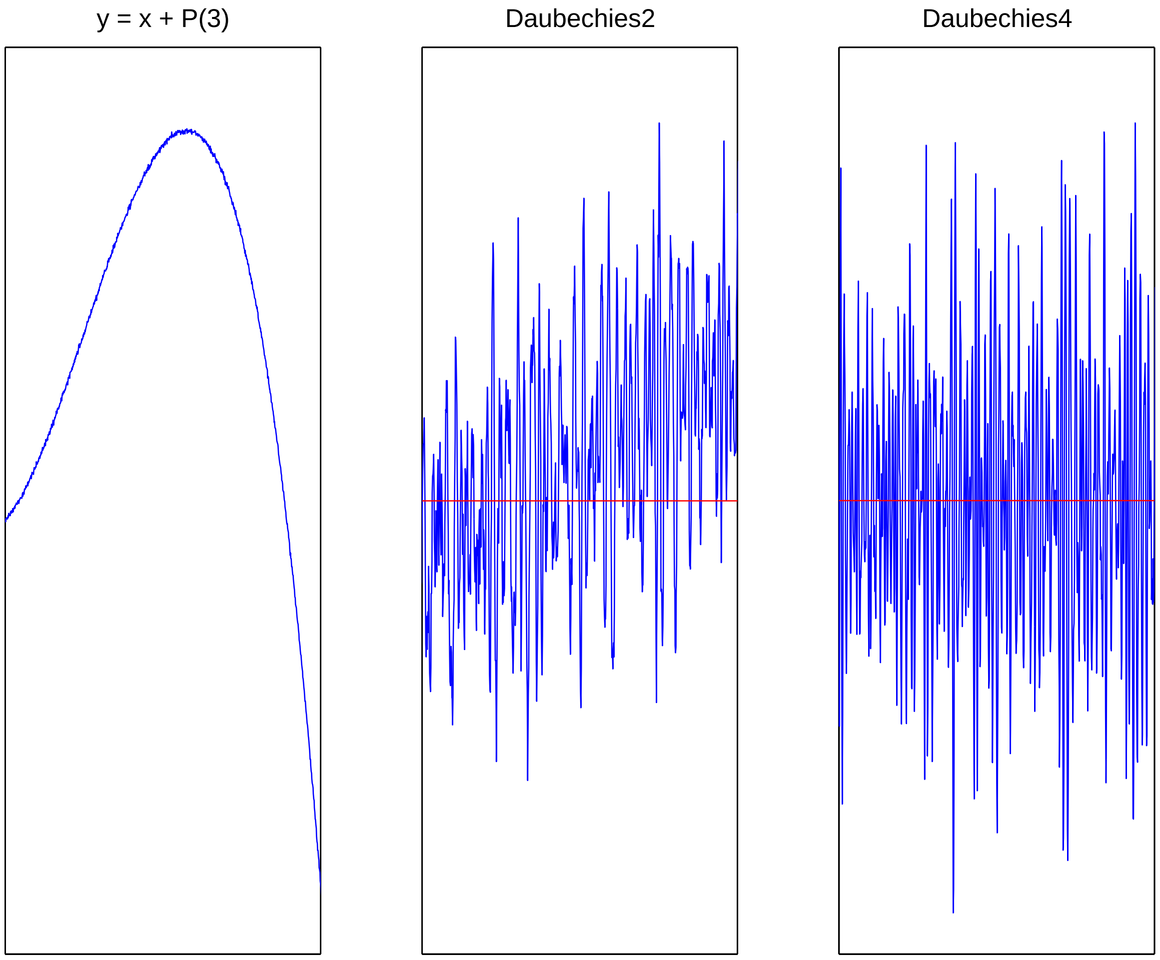
$$\int_{-\infty}^{+\infty} t^k \, \psi(t) \, dt = 0, \ \text{ for } \ k = 0, 1, \ldots N-1,$$

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

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

In other words, 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.

- *Allan variance* — A refined notion of variance — introduced in the study of atomic clocks stability (Allan, 1966) — is the so-called *Allan variance*, defined by

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

  - in the case of H-ss processes, Allan variance is such that $\text{Var}_X^{(\text{Allan})}(T) \sim T^{2H}$ when $T \to \infty$;

  - when evaluated over dyadic intervals, Allan variance identifies to the variance of Haar details:

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

Wavelets and . . . — 3.

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

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

  - 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 $T \to \infty$.

  - interpretation as *fluctuations/average* suggests the *wavelet generalization* given by:

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

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

Variations

Beyond 2nd order — 1.

- *Stable motions* — Let $X(t)$ be a (zero-mean) "bursty" process, with possibly *infinite variance*. A possible model is given by *stable motions*, whose representation reads

$$X(t) = \int_{-\infty}^{+\infty} f(t,u) M(du)\,,$$

  with:

  - $M(du)$ some *symmetric α-stable* ("SαS") measure, with scale parameter σ;

  - $f(t,u)$ an integration kernel that controls the time dependence of the statistics of the process.

Symmetric α-stable variables

- *Definition* — A random variable X is said to be *symmetric* α*-stable* (SαS) if its characteristic function is of the form:

$$\mathbb{E}\exp\{i\theta X\} = \exp\{-\sigma^{\alpha}\,|\theta|^{\alpha}\}.$$

(*Remark*: $\alpha = 1 \Rightarrow$ *Cauchy* and $\alpha = 2 \Rightarrow$ *Gauss*.)

- *Heavy tails* — Let $X \sim S_{\alpha}(\sigma)$ with $0 < \alpha < 2$. We then have:

$$\beta \geq \alpha \Rightarrow \mathbb{E}|X|^{\beta} = \infty.$$

- *Stability* — Let $\{X_i \sim S_{\alpha_i}(\sigma_i); i = 1, 2\}$ be independent SαS variables, and $X := X_1 + X_2$. We then have $X \sim S_{\alpha}(\sigma)$, with $\sigma = (\sigma_1^{\alpha} + \sigma_2^{\alpha})^{1/\alpha}$.

Some stable motions

If $f(t,u) \equiv f(0, u-t)$, then $X(t)$ is a *stationary* process and, if $f(ct, cu) = c^{H-1/\alpha} f(t,u)$ for any $c > 0$, $X(t)$ is an H-ss process.

- *Lévy flight* — $f(t,u) := 1$ if $t \geq \max(u, 0)$, and 0 otherwise. Lévy flight (LF) is an H-ss process with $H = 1/\alpha$, and its increments are *stationary* and *independent*.

- *Linear fractional stable motion* — $f(t,u) := (t-u)_+^d - (-u)_+^d$, where $(t)_+ = t$ if $t \geq 0$, and 0 otherwise. Linear fractional stable motion (LFSM) depends on a parameter $d \leq 1/2$ and is an H-ss process with $H = d + 1/\alpha$. Its increments are *stationary* but *dependent*, dependence being controlled by d.

H-sssi stable processes and wavelets

- *Representation* — Under mild conditions on the wavelet ψ and the kernel $f(t,u)$, the wavelet coefficients of a stable motion are SαS random variables with integral representation:

$$d_X(j,k) = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(t,u)\psi_{jk}(t)dt \right) M(du).$$

- *Scaling* — If X is H-sssi stable, the scale parameters of its wavelet coefficients satisfy $\sigma_{jk}^{\alpha} = 2^{j(H+1/2)\alpha}\sigma_{00}^{\alpha}$.

  While the covariance structure of a stable process X is not defined when $\alpha < 2$, the logarithmically transformed process $Y := \log|X|$ has finite second order statistics $\Rightarrow$ *considering wavelet log-coefficients*.

H-sssi stable processes and wavelet log-coeffs.

- *Scaling* — Wavelet log-coefficients of H-sssi stable processes are such that:

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

- *From LRD to SRD* — In the LFSM case, the asymptotic dependence structure is bounded as:

$$|\text{Cov}\;(\log_2|d_x(j,k)|, \log_2|d_x(j,k+n)|)| \leq C\,|n|^{-(\alpha/4)(N-H)},$$

when $|n| \to \infty$: the decay can be made *as fast as desired* by increasing the number of vanishing moments N.

Estimation in the stable case — 1.

- *Variance substitute* — The quantity of interest is in this case $w_X(j) := \mathbb{E} \log_2 |d_X(j,k)|$, that can be estimated by:

$$\widehat{w}_X(j) := \frac{1}{N_j} \sum_{k=1}^{N_j} \log_2 |d_X(j,k)| \, .$$

- *Bias and variance* — Assuming an exact *decorrelation*, one has $\mathbb{E} \widehat{w}_X(j) = (H + 1/2)j + \text{Const.}$, and

$$\text{Var } \widehat{w}_X(j) = \left(1 + \frac{2}{\alpha^2}\right) \frac{(\pi \log_2 e)^2}{12} \frac{1}{N_j}.$$

Estimation in the stable case — 2.

- *Hurst exponent* — The corresponding estimator for H is *unbiased* and of variance decreasing as $1/N_0$.

- *Stability parameter* — Since wavelet details $d_X(j^*, .)$ form, at *any* scale j^*, sequences of variables that are 1) α-*stable* and 2) *almost decorrelated*, the stability parameter α can be estimated as $\widehat{\alpha} := 1/H^*$, where H^* is the Hurst exponent obtained from any wavelet analysis $d_{S^*}(j,k)$ of the cumulative sum $S^*(k) := \sum_{m=-\infty}^{k} d_X(j^*, m)$ (Abry *et al.*, 2000).

  (*Remark*: in practice, minimizing variance $\Rightarrow$ maximizing the number of data points $\Rightarrow j^* = 1$.)

Beyond 2nd order — 2.

Given the *renormalized* definition $T_X(a) := 2^{-j/2}\, d_X(j,n)\Big|_{j=\log_2 a}$, one can consider scaling laws which *generalize* second order behaviors:

$$\mathbb{E}|T_X(a)|^q \propto a^{Hq} \;=\; \begin{array}{ll} \exp\{Hq \log a\} & (\text{“monoscaling”}) \\ \downarrow & \\ \exp\{H(q) \log a\} & (\text{“multiscaling”}) \\ \downarrow & \\ \exp\{H(q)\, n(a)\} & (\text{“cascade”}) \end{array}$$

From self-similarity to cascades

- *Self-similarity* — If a process is H-ss, the *probability density functions* of its wavelet coefficients ($a < a'$) are such that:

$$p_a(d_X) = p_{a'}(d_X/\alpha)/\alpha,$$

with $\alpha := (a'/a)^H$.

- *Generalization* — Based on Castaing's approach (Castaing, 1993), one introduces a *propagator* $G_{a,a'}$ such that:

$$p_a(d_X) = \int_0^{+\infty} G_{a,a'}(\log\alpha)\, p_{a'}(d_X/\alpha)\, d\log\alpha/\alpha.$$

($G_{a,a'}(u) = \delta\left(u - H\log(a'/a)\right) \Rightarrow$ *exact* self-similarity.)

Infinitely divisible cascades — 1.

- *Convolution* — It follows from the cascade relation that the pdf's of the *log-details* are given by:

$$\begin{aligned} p_a(\log|d_X|) &= \int G_{a,a'}(u)\, p_{a'}(\log|d_X| - u)\, du \\ &= (G_{a,a'} \star p_{a'})(\log|d_X|). \end{aligned}$$

- *Propagation* — If $p_a = G_{a,a''} \star p_{a''}$ and $p_{a''} = G_{a'',a'} \star p_{a'}$, one has directly $p_a = G_{a,a'} \star p_{a'}$, with $G_{a,a'} = G_{a,a''} \star G_{a'',a'}$.

Infinitely divisible cascades — 2.

- *Infinite divisibility* — If there is *no characteristic scale* between a and a', the intermediate scale a'' is *arbitrary*. Iterating the argument thus leads to:

$$G_{a,a'} = \underbrace{G_0 \star G_0 \star \ldots \star G_0}_{n(a)-n(a')}.$$

- *Moments* — Letting $H(q) := \log \tilde{G}_0(q)$, with $\tilde{G}_0$ the *Laplace transform* of G_0, one gets:

$$\mathbb{E}\,|T_X(a)|^q \propto \exp\{H(q)\,n(a)\}$$

$\Rightarrow$ *Separability* between order q and scale a.

Cascades and scale invariance

A *scale invariant* cascade is characterized by $n(a) \equiv \log a$.

- *"Multiscaling"* — In the scale invariant case, one gets directly $\tilde{G}_{a,a'}(q) = (a/a')^{\log H(q)}$ and, therefore,

$$\mathbb{E}\,|T_X(a)|^q \propto a^{H(q)}.$$

- *Multifractality* — In the *small scale* limit ($a \to 0$), this is equivalent to the *multifractal model* (Riedi, 2000), with the identification $H(q) \equiv \zeta_q$.

Cascades and model testing

- *Extended self-similarity (ESS)* — From the general cascade relationship, one can infer, for any p and any q, the ESS property (Benzi *et al.*, 1993):

$$\log \mathbb{E}\,|T_X(a)|^q = (H(q)/H(p))\,\log \mathbb{E}\,|T_X(a)|^p + \mathrm{Const}(p,q).$$

- *Test* — Estimating $\mathbb{E}\,|T_X(a)|^q$ by:

$$S_q(j) := \frac{1}{N_j}\sum_{k=1}^{N_j} |d_X(j,k)|^q\,,\;\; j = \log_2 a,$$

testing for ESS amounts to testing for the *linearity* of $\log S_q$ versus $\log S_p$ (while taking into account the *estimation variances* $\mathrm{Var}\ \log S_q(j)$).

Cascades and estimation

- *Estimation of $H(q)$* — Given some (arbitrary) reference order p, the quantity $\widehat{H}(q)/H(p)$ is estimated as the *slope* in the *weighted linear regression* of $\log S_q(j)$ versus $\log S_p(j)$.

- *Estimation of $n(a)$* — For dyadic scales $a \equiv 2^j$, the estimation of $n(a)$ follows from the ESS property and reads (Chainais *et al.*, 1999):

$$\widehat{n}(a) := \left\langle \frac{1}{\widehat{H}(q)} \left(\log S_q(j) - \left\langle \log S_q(j) - \frac{\widehat{H}(q)}{H(p)} \log S_p(j) \right\rangle_j \right) \right\rangle_q .$$

Some references

Reading guide

- *Scaling processes*

  - Schrœder, M.R. (1991). *Fractals, Chaos, Power Laws*, Freeman.

  - Beran, J. (1994). *Statistics for Long-Memory Processes*, Chapman & Hall.

  - Samorodnitsky, G. and Taqqu, M.S. (1994). *Stable Non-Gaussian Processes: Stochastic Models with Infinite Variance*, Chapman and Hall.

  - Tricot, C. (1995). *Curves and Fractal Dimension*, Springer Verlag.

- *Wavelets and algorithms*

  - Mallat, S. (1998). *A Wavelet Tour of Signal Processing*, Academic Press.
  - Carmona, R., Hwang, H.L., Torrésani, B. (1998). *Practical Time-Frequency Analysis*, Academic Press.
  - Flandrin, P. (1999). *Time-Frequency/Time-Scale Analysis*, Academic Press.
  - Vetterli, M. and Kovacevic, J (1995). *Wavelets and Subband Coding*, Prentice-Hall.

- *Wavelets and scaling processes*

  – Abry, P., Flandrin, P., Taqqu, M.S. and Veitch, D. (2000). "Wavelets for the analysis, estimation and synthesis of scaling data," in *Self-Similar Network Traffic and Performance Evaluation* (K. Park and W. Willinger, *eds.*), 39–88, Wiley.

  – Abry, P., Gonçalvès P. et Lévy-Véhel, J., éds. (2002). *Lois d'Échelle, Fractales et Ondelettes* (2 vol.), Hermes.

  – Vidakovic, B. (1999). *Statistical Modeling by Wavelets*, Wiley.

  – Wornell, G.W. (1995). *Signal Processing with Fractals: A Wavelet-Based Approach*, Prentice-Hall.

Web links

- *Publications, preprints, software*

  `www.ens-lyon.fr/~flandrin/`

  `www.ens-lyon.fr/~pabry/`

  `www.ens-lyon.fr/PHYSIQUE/Signal/index.html`

  `www.emulab.ee.mu.oz.au/~darryl/`

  `www.cmap.polytechnique.fr/~bacry/LastWave/`

- *Wavelet Digest*

  `www.wavelet.org`