

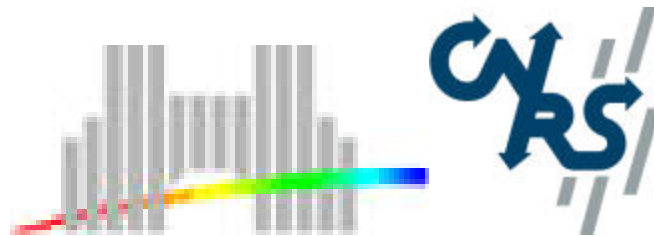
NAVIGATING SCALING: MODELLING AND ANALYSING

P. Abry⁽¹⁾,

P. Gonçalves⁽²⁾

(1) SISYPH, CNRS,
Ecole Normale Supérieure
Lyon, France

(2) INRIA,
Grenoble, France

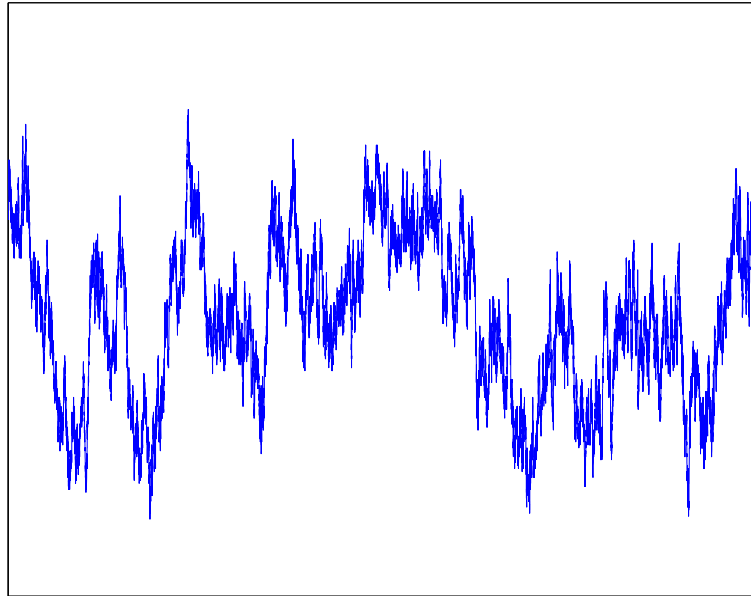


IN COLLABORATIONS WITH :

P. Flandrin, D. Veitch,
P. Chainais, B. Lashermes, N. Hohn,
S. Roux, P. Borgnat,
M. Taqqu, V. Pipiras, R. Riedi

Wavelet And Multifractal Analysis, Cargèse, France, July 2004.

SCALING PHENOMENA ?



- **DETECTION:** SCALING ? WHAT DOES IT MEAN ? NON STATIONARITY ?
- **IDENTIFICATION:** RELEVANT STOCHASTIC MODELS ?
- **ESTIMATION:** RELEVANT PARAMETER ESTIMATION ?
- **SIDE ISSUES:**
ROBUSTNESS ? COMPUTATIONAL COST ? REAL TIME ? ON LINE ?

OUTLINE

I. INTUITIONS, MODELS, TOOLS

- I.1 INTUITIONS, DEFINITION, APPLICATIONS
- I.2 STOCHASTIC MODELS: SELF-SIMILARITY VS MULTIFRACTAL
- I.3 MULTIREOLUTION TOOLS, AGGREGATION, INCREMENTS
- I.4 WAVELETS, CONTINUOUS, DISCRETE

II. SECOND ORDER ANALYSIS, SELF SIMILARITY AND LONG MEMORY

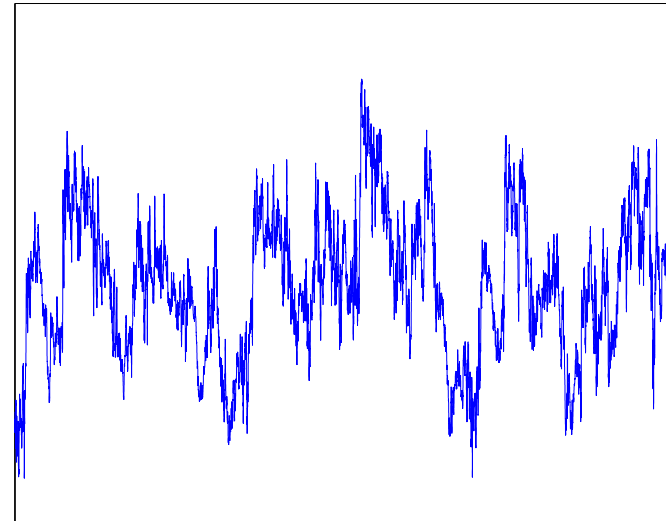
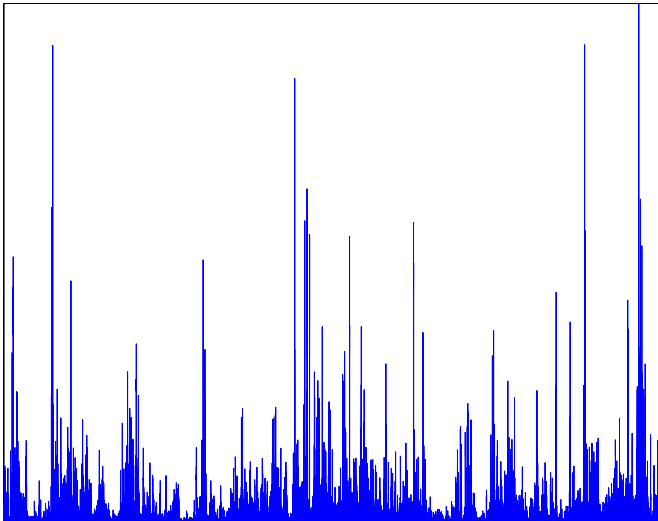
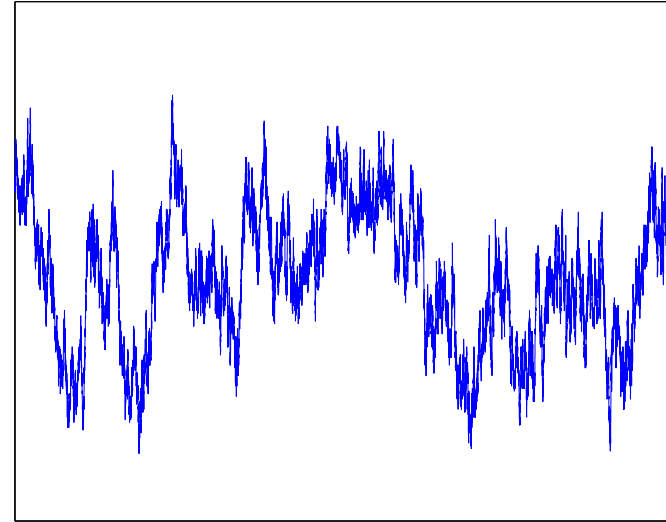
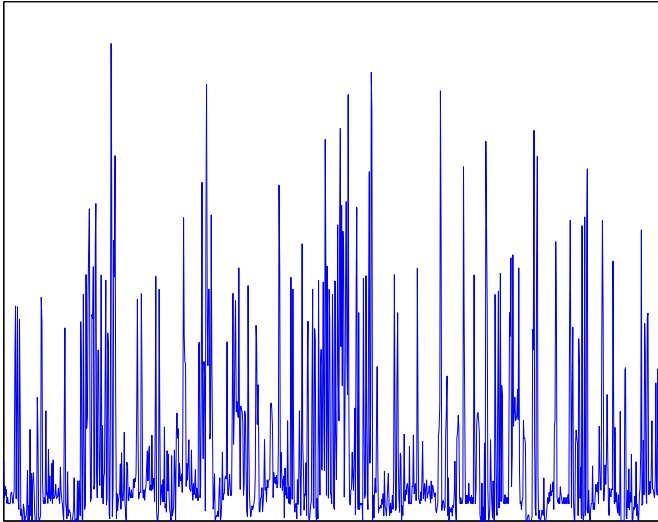
- II.1 RANDOM WALKS, SELF SIMILARITY, LONG MEMORY,
- II.2 2ND ORDER WAVELET STATISTICAL ANALYSIS,
- II.3 ESTIMATION, ESTIMATION PERFORMANCE,
- II.4 ROBUSTNESS AGAINST NON STATIONARITIES,

III. HIGHER ORDER ANALYSIS, MULTIFRACTAL PROCESSES

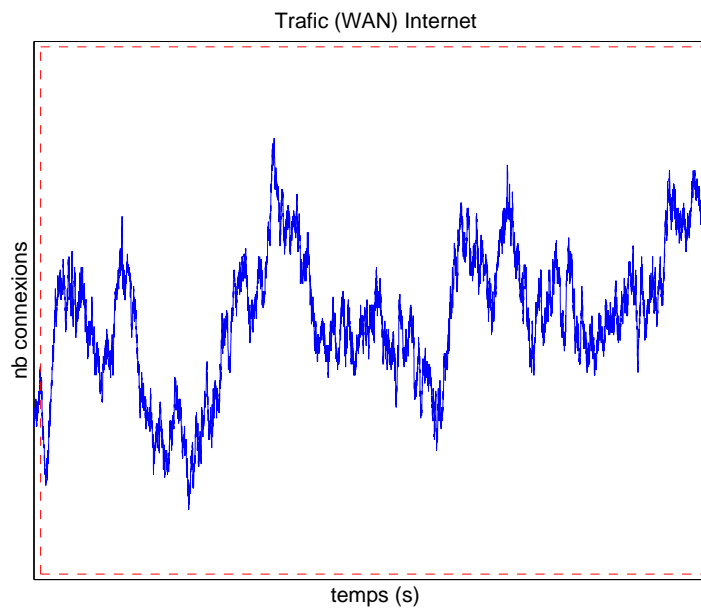
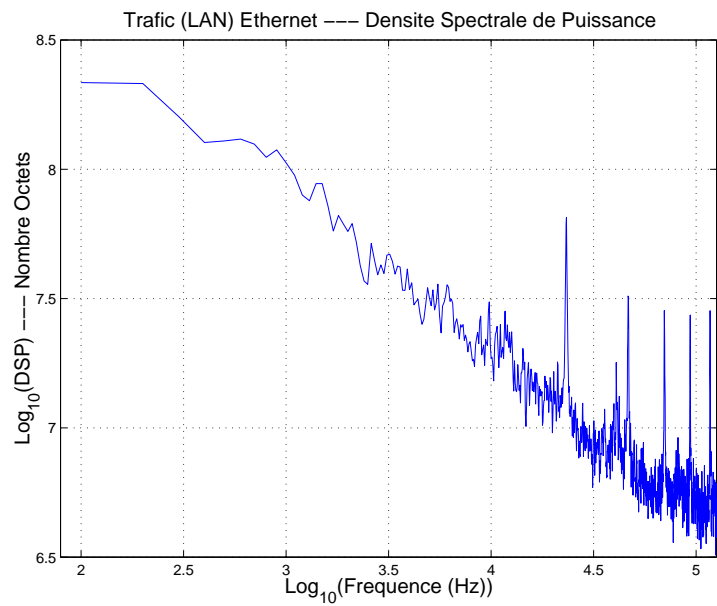
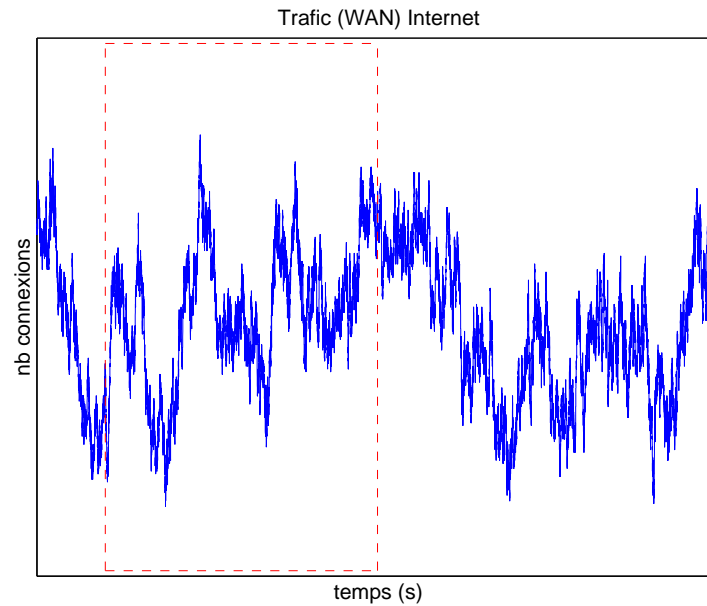
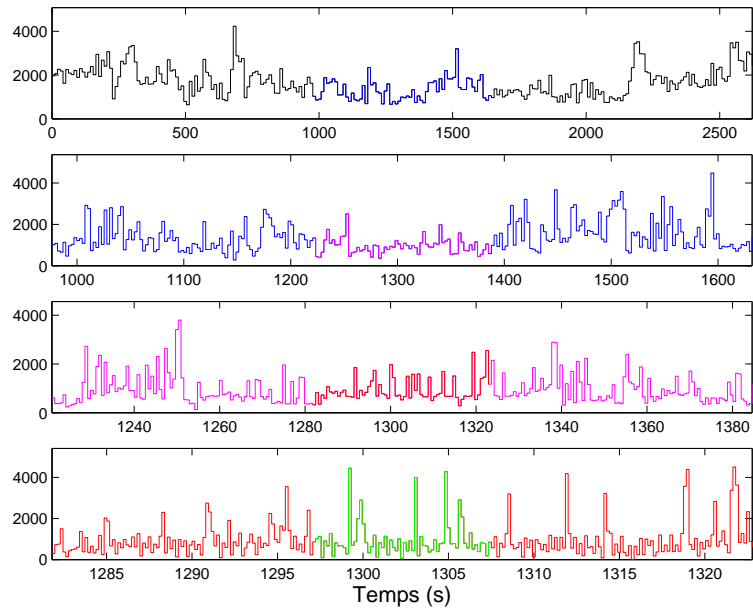
- III.1 MULTIPLICATIVE CASCADES, MULTIFRACTAL PROCESSES,
- III.2 HIGHER ORDER WAVELET STATISTICAL ANALYSIS,
- III.3 FINITENESS OF MOMENTS,
- III.4 ESTIMATION, ESTIMATION PERFORMANCE,
- III.5 NEGATIVE ORDERS,
- III.6 BEYOND POWER LAWS.

IRREGULARITIES, VARIABILITIES

SCALING OR NON STATIONARITIES?



SCALING ?



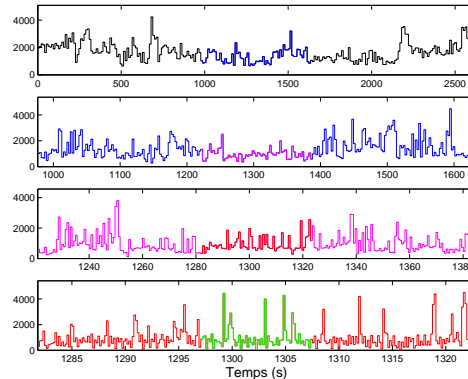
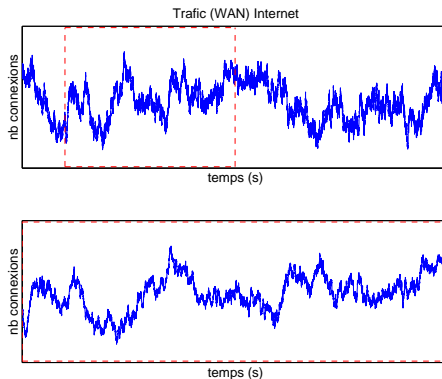
SCALING !

- **DEFINITION :**

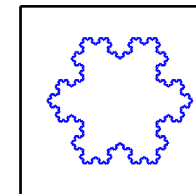
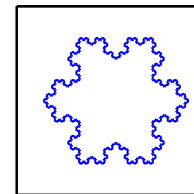
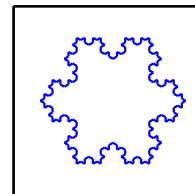
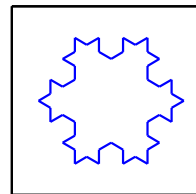
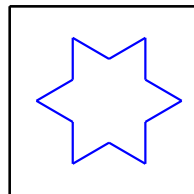
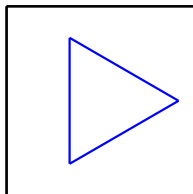
NON PROPERTY: NO CHARACTERISTIC SCALE.
NON GAUSSIAN, NON STATIONARY, NON LINEAR

- **EVIDENCE:**

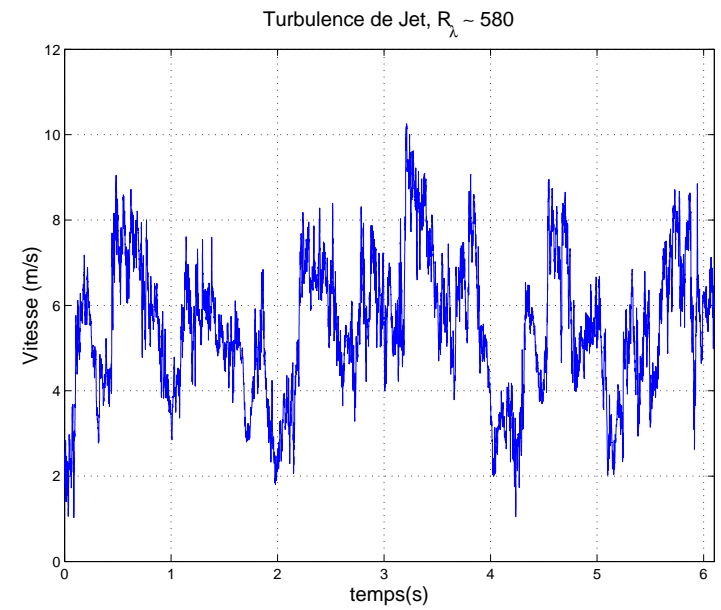
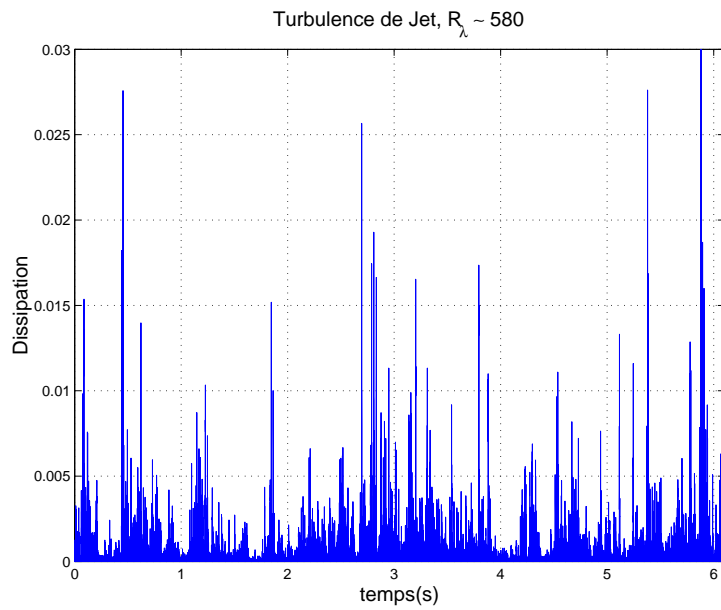
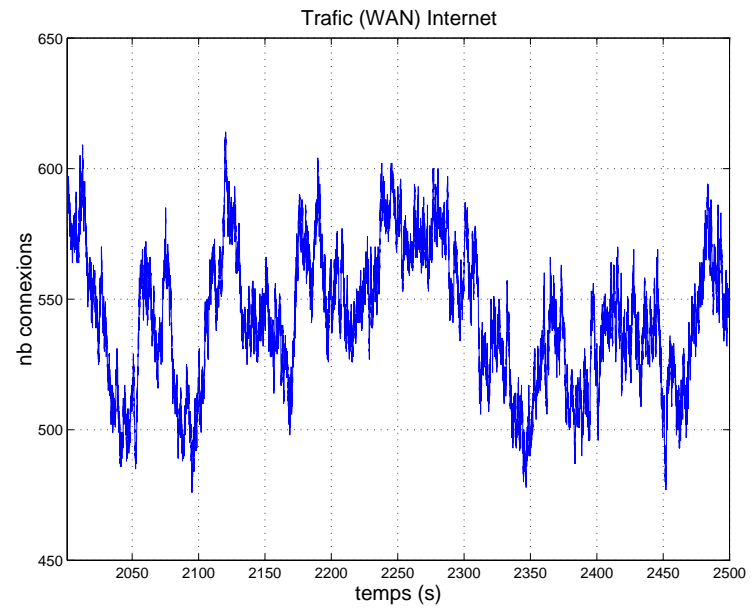
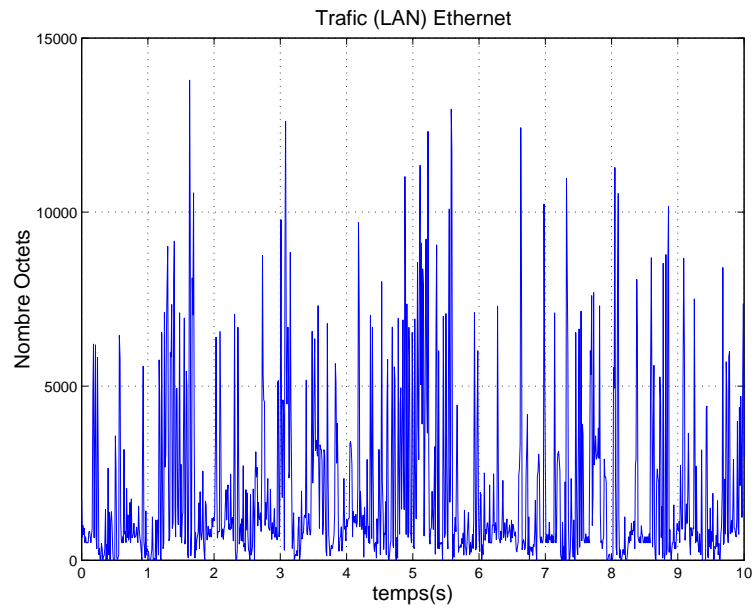
The whole resembles to its part, the part resembles to the whole.



- **ANALYSIS:** Rather than for a characteristic scale, look for a relation, a mechanism, a cascade between scales.



UBIQUITY ?



UBIQUITY !

- Hydrodynamic Turbulence,
- Physiology, Biological Rhythms (Heart beat, walk),
- Geophysics (Faults Repartition, Earthquakes),
- Hydrology (Water Levels),
- Statistical Physics (Long Range Interactions),
- Thermal Noises (semi-conductors),
- Information Flux on Networks, Computer Network Traffic,
- Population Repartition (local: cities, global: continent),
- Financial Markets (Daily returns, Volatility, Currencies Exchange Rates),
- ...

MODELLING TOOLS

- **SELF-SIMILARITY:** $\mathbf{E}|X(t + a\tau_0) - X(t)|^q = C_q a^{qH}$
 - Power Laws,
 - $\forall a$ (for all scales),
 - $\forall q / \mathbf{E}|d_X(j, k)|^q < \infty$,
 - A single parameter H
 - Additive Structure.
- **MULTIFRACTAL:** $\mathbf{E}|X(t + a\tau_0) - X(t)|^q = C_q a^{\zeta(q)}$
 - Power Laws,
 - $\forall a < L$, (for fine scales only, in the limit $a \rightarrow 0$),
 - $\forall q$?
 - A whole collection of scaling parameter $\zeta(q)$
 - Multiplicative Structure.

- **BEYOND POWER LAWS**

$$\mathbf{E}|X(t + a\tau_0) - X(t)|^q = C_q a^{qH} = C_q \exp(qH \ln a)$$

$$\mathbf{E}|X(t + a\tau_0) - X(t)|^q = C_q a^{\zeta(q)} = C_q \exp(\zeta(q) \ln a)$$

$$\mathbf{E}|X(t + a\tau_0) - X(t)|^q = \quad \quad \quad = C_q \exp(\zeta(q)n(a))$$

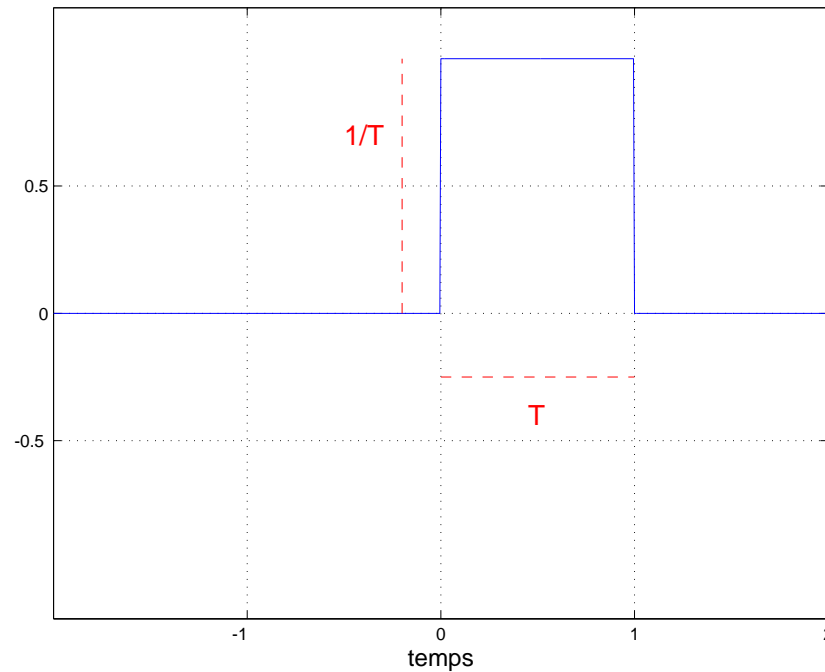
→ VISIT PIERRE CHAINAIS'S POSTER

ANALYSING TOOL 1 : AGGREGATION

COMPARE DATA AGAINST A BOX, THEN VARY a

$$T_X(a, t) = \frac{1}{aT_0} \int_t^{t+aT_0} X(u) du$$

AVERAGE



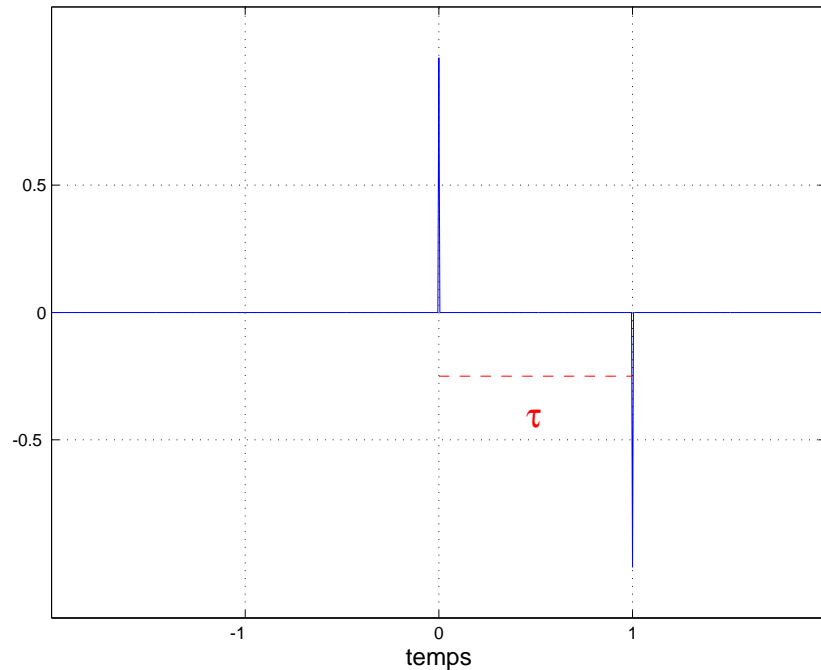
WORKS ONLY FOR POSITIVE TIME SERIES, DENSITY

ANALYSING TOOL 2 : INCREMENTS

COMPARE DATA AGAINST A DIFFERENCE OF DELTA FUNCTIONS, THEN VARY a

$$T_X(a, t) = X(t + a\tau_0) - X(t)$$

DIFFERENCE



INCREMENTS OF HIGHER ORDERS OR GENERALISED N -VARIATIONS

– Order 2 : $T_X(a, t) = -X(t + 2a\tau_0) + 2X(t + a\tau_0) - X(t)$,

– Order N : $T_X(a, t) = \sum_{p=0}^N (-1)^p a_p X(t + pa\tau_0)$,

where $\sum_{p=0}^N (-1)^p a_p p^k \equiv 0$, $k = 0, \dots, N - 1$.

ANALYSING TOOL: MULTIREOLUTION ANALYSIS

- MULTIREOLUTION QUANTITIES:

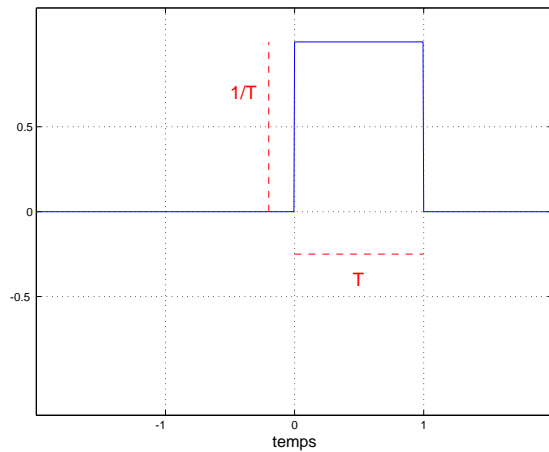
$$X(t) \rightarrow T_X(a, t) = \langle f_{a,t} | X \rangle, \quad f_{a,t}(u) = \frac{1}{a} f_0\left(\frac{u-t}{a}\right)$$

AGGREGATION

$$f_0(u) = (\beta_0)$$

$$= \frac{1}{aT_0} \int_t^{t+aT_0} X(u) du$$

BOX, AVERAGE

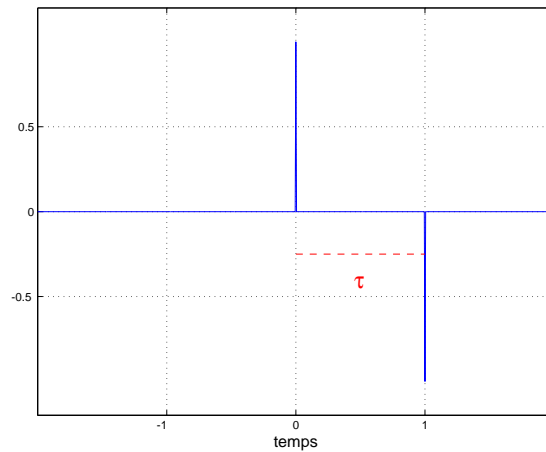


INCREMENTS

$$f_0(u) = (I_0)$$

$$= X(t + a\tau_0) - X(t)$$

DIFFERENCE



?

?

?

?

?

ANALYSING TOOL: MULTIREOLUTION ANALYSIS

- MULTIREOLUTION QUANTITIES:

$$X(t) \rightarrow T_X(a, t) = \langle f_{a,t} | X \rangle, \quad f_{a,t}(u) = \frac{1}{a} f_0\left(\frac{u-t}{a}\right)$$

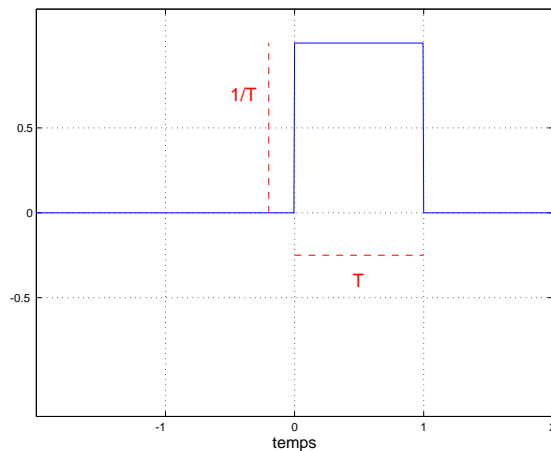
- CHOICES FOR MOTHER FUNCTIONS: f_0 ,

AGGREGATION

$$f_0(u) = (\beta_0)$$

$$= \frac{1}{aT_0} \int_t^{t+aT_0} X(u) du$$

BOX, AVERAGE

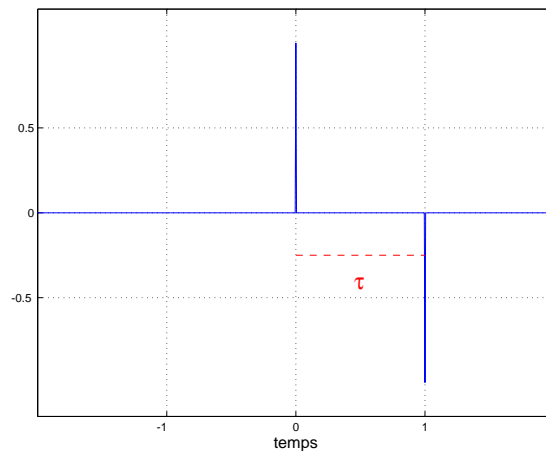


INCREMENTS

$$f_0(u) = (I_0)$$

$$= X(t + a\tau_0) - X(t)$$

DIFFERENCE

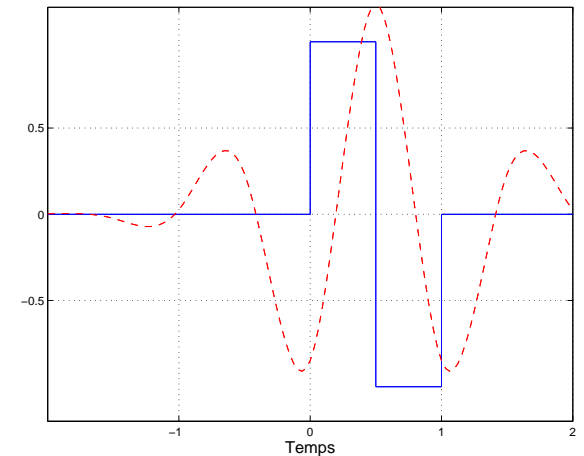


WAVELETS

$$f_0(u) = \psi_{0,N}$$

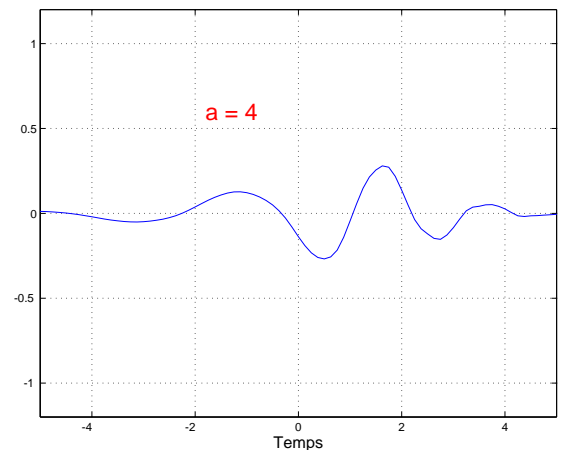
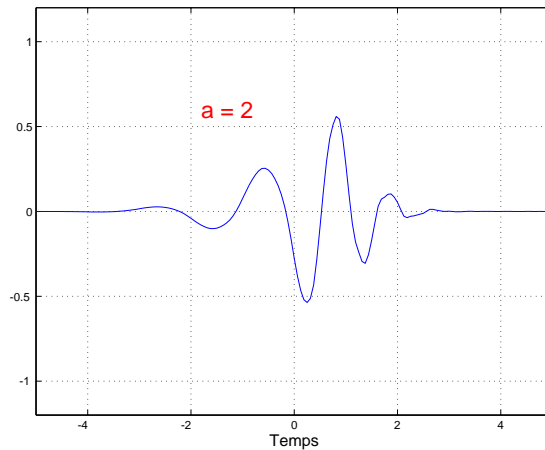
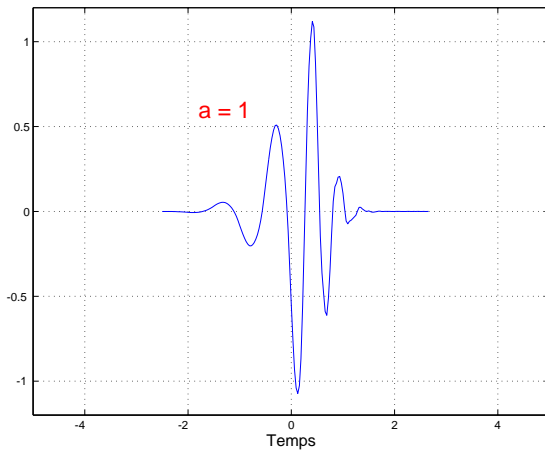
$$= \int X(u) \frac{1}{|a|} \psi_0\left(\frac{u-t}{a}\right),$$

AVERAGE, DIFFERENCE



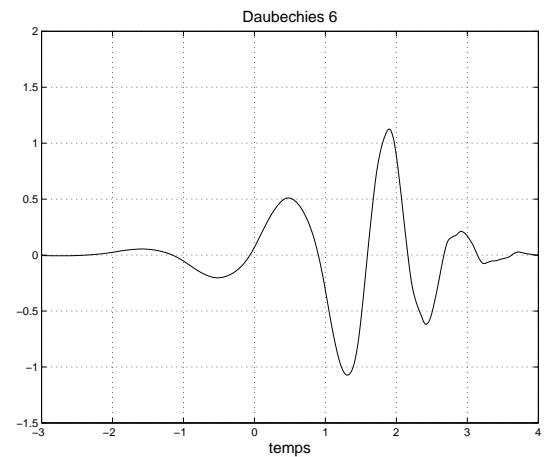
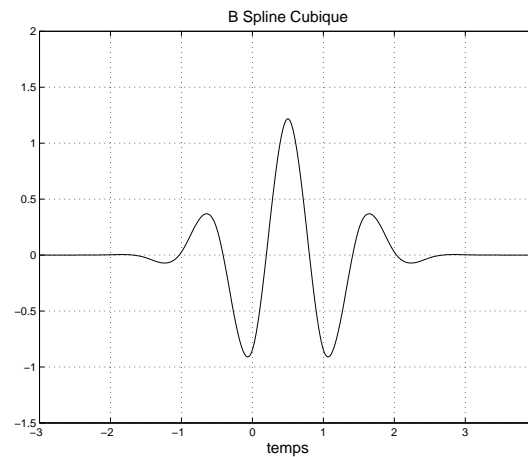
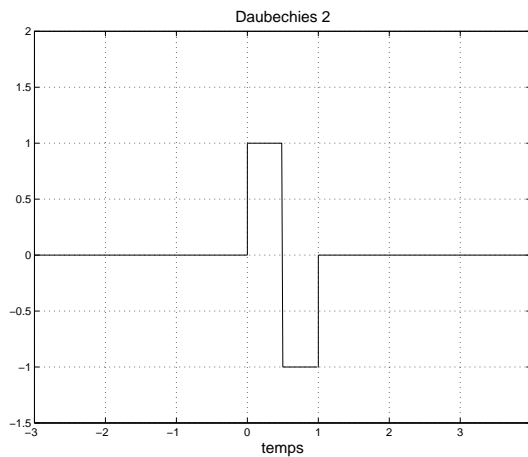
WAVELETS AND SCALING: KEY INGREDIENTS

- DILATION OPERATOR, $\frac{1}{|a|}\psi_0\left(\frac{t}{|a|}\right)$



- NUMBER OF VANISHING MOMENTS,

$$N \geq 1, \int t^k \psi_0(t) dt \equiv 0, \quad k = 0, 1, \dots, N - 1.$$



ANALYSING TOOL: MULTIREOLUTION ANALYSIS

- MULTIREOLUTION QUANTITIES:

$$X(t) \rightarrow T_X(a, t) = \langle f_{a,t} | X \rangle, \quad f_{a,t}(u) = \frac{1}{a} f_0\left(\frac{u-t}{a}\right)$$

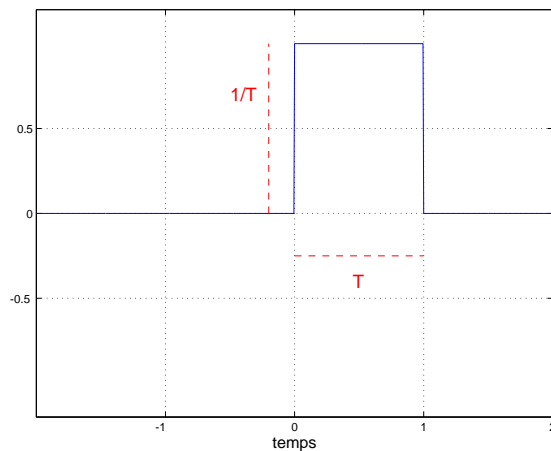
- CHOICES FOR MOTHER FUNCTIONS: f_0 ,

AGGREGATION

$$f_0(u) = (\beta_0)^{*N}$$

$$= \frac{1}{aT_0} \int_t^{t+aT_0} X(u) du$$

BOX, AVERAGE



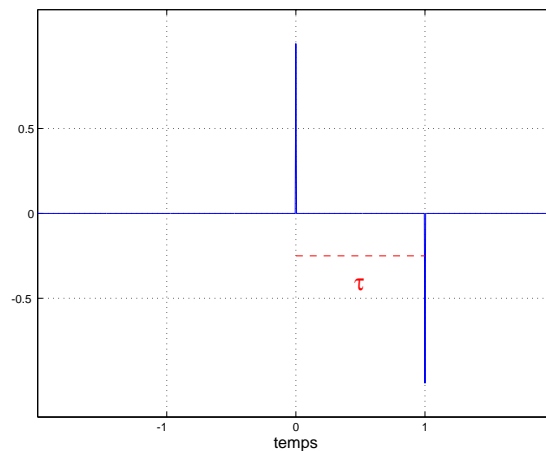
N

INCREMENTS

$$f_0(u) = (I_0)^{*N}$$

$$= X(t + a\tau_0) - X(t)$$

DIFFERENCE



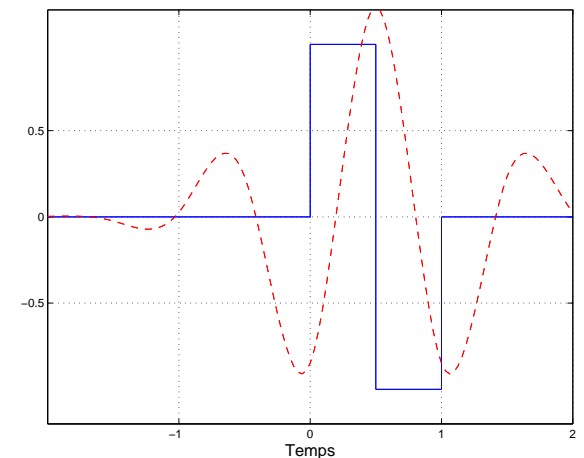
N

WAVELETS

$$f_0(u) = \psi_{0,N}$$

$$= \int X(u) \frac{1}{|a|} \psi_0\left(\frac{u-t}{a}\right),$$

AVERAGE, DIFFERENCE



N

WAVELET TRANSFORMS

- **MOTHER-WAVELET AND "BASIS":** $\int \psi_0(u)du = 0, \quad \psi_{a,t}(u) = \frac{1}{|a|}\psi_0\left(\frac{u-t}{a}\right)$
- **WAVELET COEFFICIENTS:**

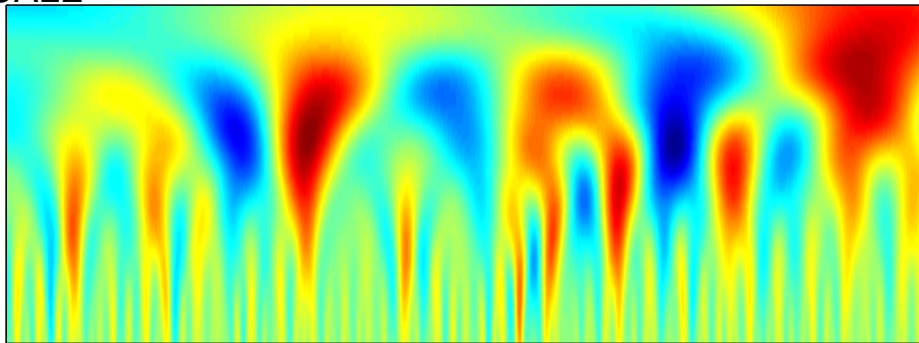
CONTINUOUS WT

$$T_X(a, t) = \langle X, \psi_{a,t} \rangle$$

AND DISCRETE WT

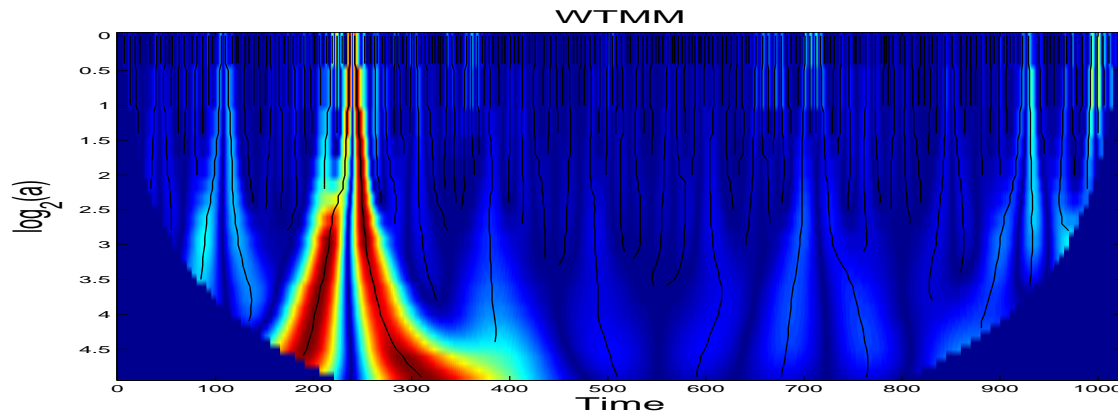
$$d_X(j, k) = T_X(a = 2^j, t = 2^j k)$$

SCALE



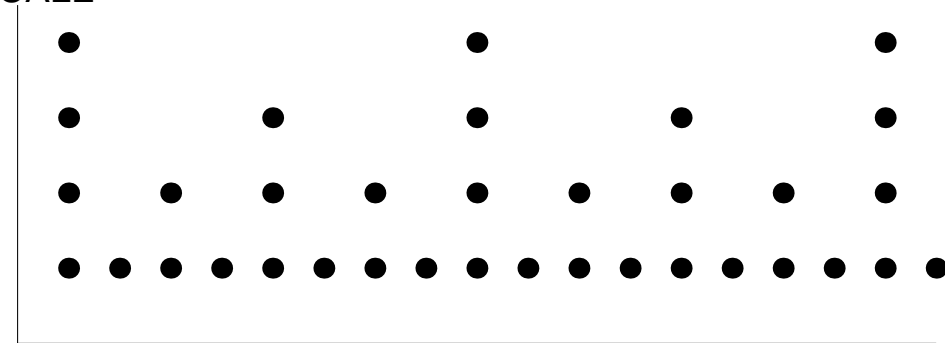
TIME

MODULUS MAXIMA WT



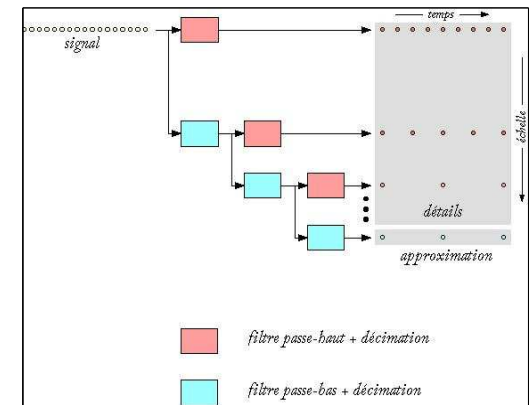
WTMM

SCALE



TIME

AND FAST PYRAMIDAL ALGORITHM



OUTLINE

I. INTUITIONS, MODELS, TOOLS

- I.1 INTUITIONS, DEFINITION, APPLICATIONS
- I.2 STOCHASTIC MODELS: SELF-SIMILARITY VS MULTIFRACTAL
- I.3 MULTIREOLUTION TOOLS, AGGREGATION, INCREMENTS
- I.4 WAVELETS, CONTINUOUS, DISCRETE

II. SECOND ORDER ANALYSIS, SELF SIMILARITY AND LONG MEMORY

- II.1 RANDOM WALKS, SELF SIMILARITY, LONG MEMORY,
- II.2 2ND ORDER WAVELET STATISTICAL ANALYSIS,
- II.3 ESTIMATION, ESTIMATION PERFORMANCE,
- II.4 ROBUSTNESS AGAINST NON STATIONARITIES,

III. HIGHER ORDER ANALYSIS, MULTIFRACTAL PROCESSES

- III.1 MULTIPLICATIVE CASCADES, MULTIFRACTAL PROCESSES,
- III.2 HIGHER ORDER WAVELET STATISTICAL ANALYSIS,
- III.3 FINITENESS OF MOMENTS,
- III.4 ESTIMATION, ESTIMATION PERFORMANCE,
- III.5 NEGATIVE ORDERS,
- III.6 BEYOND POWER LAWS.

MOD. TOOL 1: RAND. WALKS AND SELF SIMILARITY

RANDOM WALK: $X(t + \tau) = X(t) + \underbrace{\delta_\tau X(t)}_{\text{Steps or Increments}}$

STATISTICAL PROPERTIES OF THE STEPS:

- **A1:** Stationary,
- **A2:** Independent,
- **A3:** Gaussian,
 - \Rightarrow Ordinary Random Walk, Ordinary Brownian Motion,
 - $\Rightarrow \mathbf{E}X(t)^2 = 2D|t|$, Einstein relation,
 - $\Rightarrow \mathbf{E}X(t)^q = 2D|t|^{q/2}$, $q > -1$.

ANOMALIES:

- $\Rightarrow \mathbf{E}X(t)^2 = 2D|t|^\gamma$,
- $\Rightarrow \mathbf{E}X(t)^2 = \infty$.

SELF SIMILAR RANDOM WALKS:

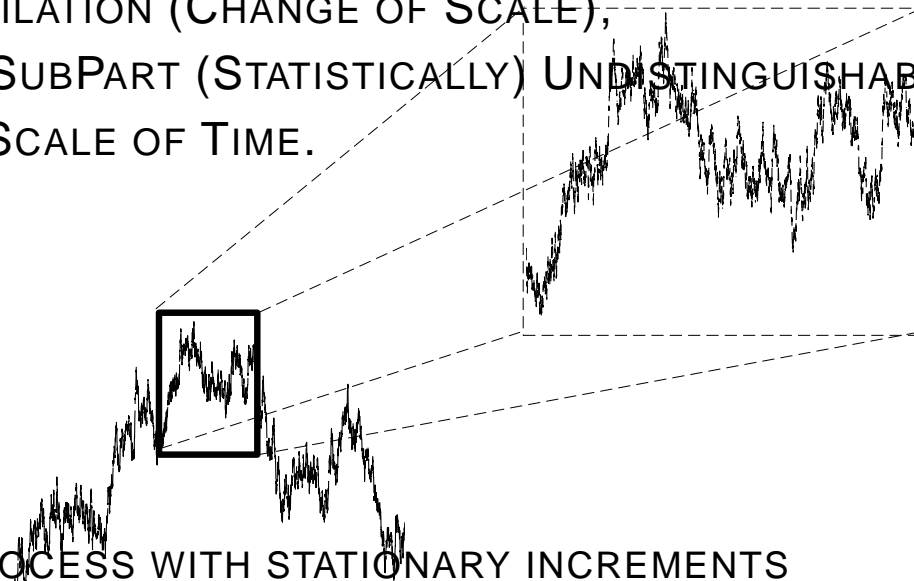
- **B1:** Stationary,
- **B2:** Self Similarity

MODELLING TOOL 1: SELF-SIMILARITY

- **DEFINITION:** $\delta_\tau X(t) \stackrel{fdd}{=} c^H \delta_{\tau/c} X(t/c), \forall c > 0$, DILATION FACTOR,
 $H > 0$: SELF-SIMILARITY EXPONENT

- **INTERPRETATIONS:**

- COVARIANCE UNDER DILATION (CHANGE OF SCALE),
- THE WHOLE AND THE SUBPART (STATISTICALLY) UNDISTINGUISHABLE,
- NO CHARACTERISTIC SCALE OF TIME.



- **IMPLICATIONS:**

- NON STATIONARITY PROCESS WITH STATIONARY INCREMENTS
- $\mathbf{E}|X(t + a\tau_0) - X(t)|^q = C_q |a|^{qH}$,
- $\forall a > 0, \forall c > 0, \forall q \quad \mathbf{E}|X(t)|^q < \infty$,
- A **SINGLE** SCALING EXPONENT H .
- ADDITIVE STRUCTURE,
- (CORRELATED) RANDOM WALK, LONG MEMORY, LONG RANGE CORRELATIONS.

MOD. TOOL 1 (BIS): LONG RANGE DEPENDENCE

● **DEFINITIONS** : LET X BE A 2ND STATIONARY PROCESS WITH,

- COVARIANCE : $c_X(\tau) = \mathbb{E}X(t)X(t + \tau)$

- SPECTRUM : $\Gamma_X(\nu)$

$$c_X(\tau) = c_\tau |\tau|^{-\beta}, \quad 0 < \beta < 1, \quad |\tau| \rightarrow +\infty$$

$$\Gamma_X(\nu) = c_f |\nu|^{-\alpha}, \quad 0 < \alpha < 1, \quad |\nu| \rightarrow 0$$

WITH $\alpha = 1 - \beta$ AND $c_f = 2(2\pi) \sin((1 - \gamma)\pi/2)c_\tau$.

● **CONSEQUENCES** :

- $\sum_A^{+\infty} c_X(\tau) d\tau = +\infty, A > 0,$

- NO CHARACTERISTIC SCALE,

- AGGREGATION : $X_{aT_0}(t) = \int_t^{t+aT_0} X(u) du, \Rightarrow \text{VAR } X_{aT_0}(t) \sim a^{\alpha-1}, Ca \rightarrow +\infty,$

- INCREMENTS OF SELF.-SIM. PROC. (WITH $H > 1/2$) ARE LONG RANGE DEP. (WITH $\alpha = 2H - 1$).

WAVELETS AND SELF-SIMILAR PROCESSES WITH STATIONARY INCREMENTS - SUMMARY

(FLANDRIN ET AL., TEWFIK AND KIM)

- **P1:** $\{d_X(j, k), k \in \mathbb{Z}\}$ **STATIONARY** Sequences for each Scale 2^j .

$N \geq 1$

- **P2:** **SELF-SIMILARITY : Dilation**

$$\{X(t)\} \stackrel{d}{=} \{c^H X(t/c)\} \Rightarrow \{d_X(0, k)\} \stackrel{d}{=} \{2^{-jH} d_X(j, k)\}$$

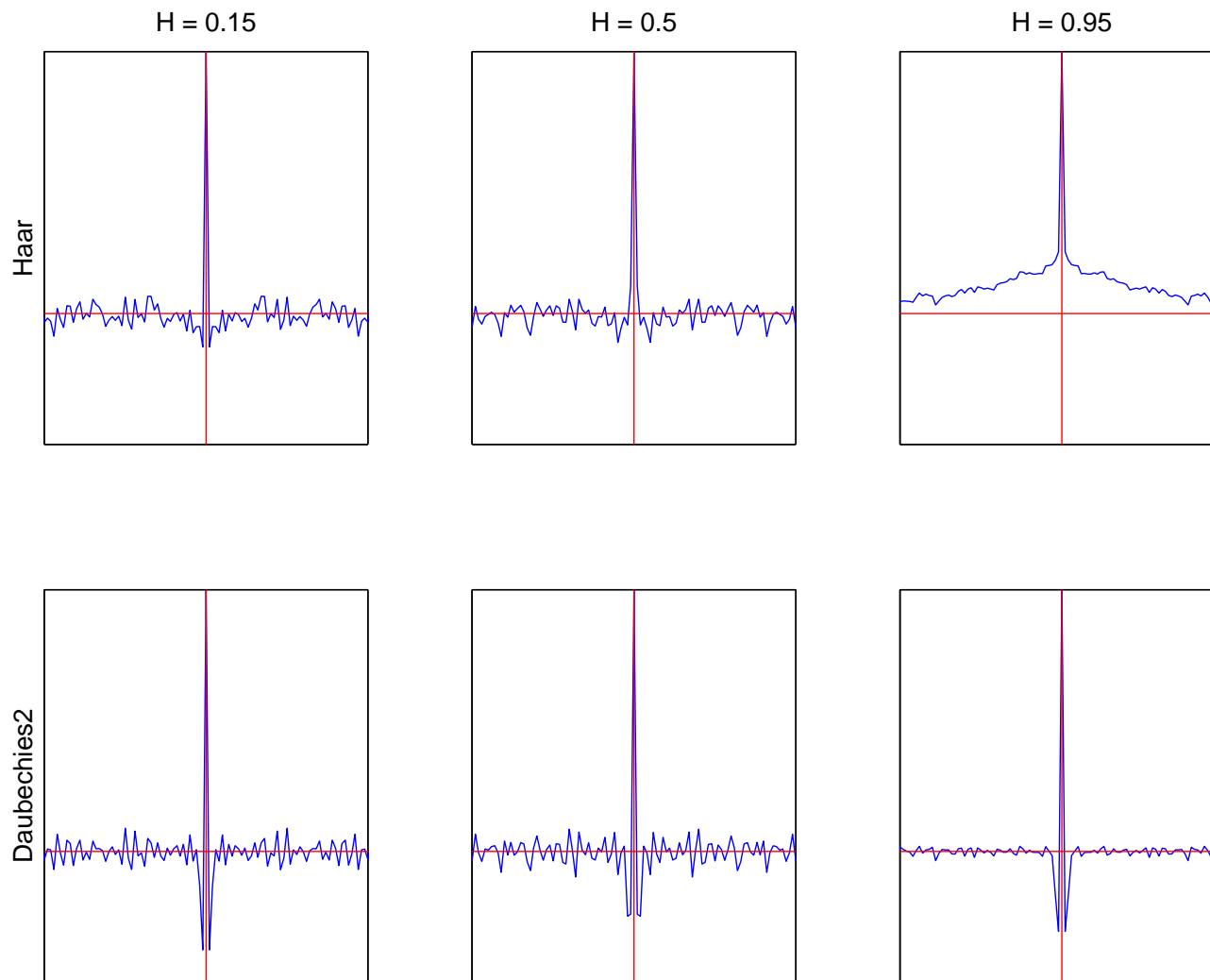
- **P3 :** **MARGINAL DIST.** $P_j(d) = \frac{1}{\beta_0} P_{j'}(\frac{d}{\beta_0}), \quad \beta_0 = \left(\frac{2^{j'}}{2^j}\right)^H$.

- **P4 :** $\{d_X(j, k)\}$ **SHORT RANGE DEPENDENT** IF $N > H + 1/2$.

$$|2^j k - 2^{j'} k'| \rightarrow +\infty, \quad |\text{Cov } d_X(j, k) d_X(j', l)| \leq D |2^j k - 2^{j'} k'|^{2(H-N)},$$

$N \geq 1$ and **Dilation**

WAVELETS AND LONG RANGE DEPENDENCE



WAVELETS AND SELF-SIMILAR PROCESSES WITH STATIONARY INCREMENTS - SUMMARY

• **P1:** $\{d_X(j, k), k \in \mathcal{Z}\}$ **STATIONARY** Sequences for each Scale 2^j .
 $N \geq 1$

• **P2: SELF-SIMILARITY : Dilation**

$$\{X(t)\} \stackrel{d}{=} \{c^H X(t/c)\} \Rightarrow \{d_X(0, k)\} \stackrel{d}{=} \{2^{-jH} d_X(j, k)\}$$

- **P3 : MARGINAL DIST.** $P_j(d) = \frac{1}{\beta_0} P_{j'}(\frac{d}{\beta_0}), \quad \beta_0 = \left(\frac{2^{j'}}{2^j}\right)^H$.

• **P4 :** $\{d_X(j, k)\}$ **SHORT RANGE DEPENDENT** IF $N > H + 1/2$.
 $|2^j k - 2^{j'} k'| \rightarrow +\infty, \quad |\text{Cov } d_X(j, k) d_X(j', k')| \leq D |2^j k - 2^{j'} k'|^{2(H-N)},$
 $N \geq 1$ and **Dilation**

\Rightarrow **IDEALISATION :** $d_X(j, k)$ **INDEPENDENT VARIABLES** .

\Rightarrow **INTERPRETATIONS:** $X(t) = \sum_k a_X(J, k) \varphi_{J,k}(t) + \sum_{j=1, \dots, J, k} d_X(j, k) \psi_{j,k}(t)$.

\Rightarrow **IMPLICATIONS:** $\mathbf{E}|d_X(j, k)|^q = \mathbf{E}|d_X(0, k)|^q 2^{j\zeta_q} \quad \forall q / \mathbf{E}|d_X(0, k)|^q < \infty$.

WAVELETS AND LONG RANGE DEPENDENCE

- **SPECTRAL ANALYSIS :**

Let X be a 2nd Order stationary process,

Let Ψ be the FT of ψ with central frequency ν_0 and bandwidth $\Delta\nu_0$.

$$\begin{aligned}\mathbf{E}|d_X(j, k)|^2 &= \int \Gamma_X(\nu) |\Psi(2^j \nu)| d\nu \\ &\simeq 2^{-j} \Gamma_X(2^{-j} \nu_0) \text{ within bandwidth } 2^{-j} \Delta\nu_0.\end{aligned}$$

- **LET X BE LONG RANGE DEPENDENT :**

- **POWER LAW:** $\Gamma_X(\nu) = c_f |\nu|^{-\alpha}, 0 < \alpha < 1, |\nu| \rightarrow 0$

- **POWER LAW:** $\mathbf{E}|d_X(j, k)|^2 \sim C 2^{j(\alpha-1)}, j \rightarrow +\infty,$

- $\{d_X(j, k)\}$ **SHORT RANGE DEPENDENT** IF $N > \alpha - 1$.

$|2^j k - 2^{j'} k'| \rightarrow +\infty, |\text{Cov } d_X(j, k) d_X(j', k')| \leq D |2^j k - 2^{j'} k'|^{\alpha-1-2N},$

$N \geq 1$ and Dilation

2ND ORDER WAVELET STATISTICAL ANALYSIS

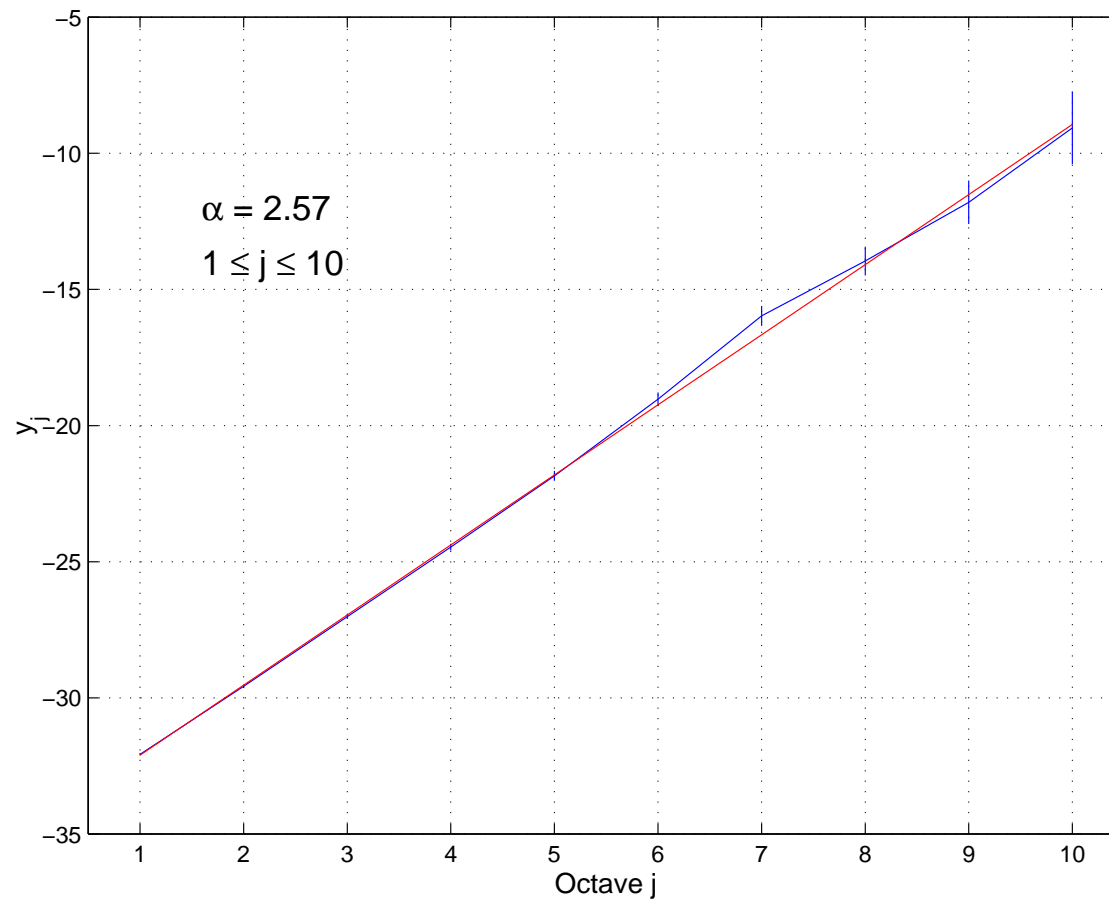
Abry, Gonçalves, Flandrin

PRINCIPLES:

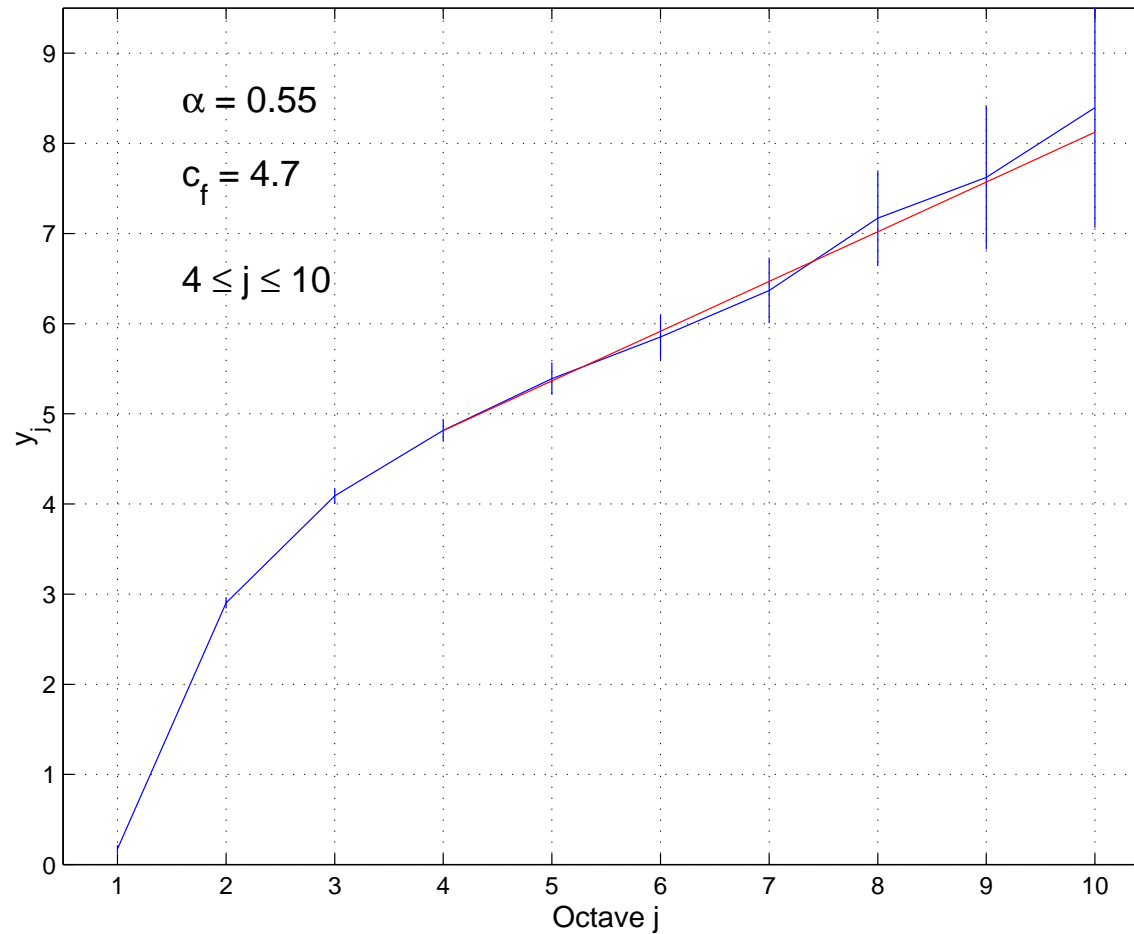
- IDEAS : **P1** $\Rightarrow \mathbf{E}|d_X(j, k)|^2 = C_2 2^{j2H}$
 $\Rightarrow \log_2 \mathbf{E}|d_X(j, k)|^2 = j2H + \beta_q,$
- PROBLEMS: ESTIMATE $\mathbf{E}|d_X(j, k)|^2$ FROM A SINGLE FINITE LENGTH OBSERVATION ?
- SOLUTION : **P2 et P3** \Rightarrow STATISTICAL AVERAGES \Rightarrow TIME AVERAGES,
 $S_2(j) = (1/n_j) \sum_{k=1}^{n_j} |d_X(j, k)|^2$

LOG-SCALE DIAGRAMS: $\log_2 S_2(j)$ vs $\log_2 2^j = j$

2ND ORDER WAVELET-BASED STATISTICAL ANALYSIS FOR SELF-SIMILARITY

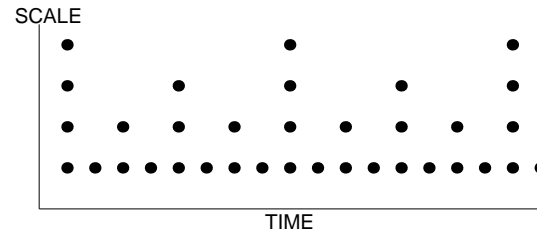


2ND ORDER WAVELET-BASED STATISTICAL ANALYSIS FOR LONG RANGE DEPENDENCE



WAVELETS AND 2ND-ORDER SCALING: ESTIMATION

- **DYADIC GRID** (DISCRETE WAVELET TRANSFORM): $a_j = 2^j, t_{j,k} = k2^j,$



- **STRUCTURE FUNCTION (TIME AVERAGE):**

$$Y_j = \left(\frac{1}{2} \log_2 S_2(2^j)\right) = \frac{1}{2} \log_2(1/n_j) \sum_{k=1}^{n_j} |d_X(j, k)|^2$$

- **DEFINITION :**

$$Y_j \text{ versus } \log_2 2^j = j,$$

$$\hat{H} = \sum_{j=j_1}^{j_2} w_j Y_j .$$

WHERE $\sum_j j w_j \equiv 1, \quad \sum_j w_j \equiv 0,$ WITH $w_j \equiv \frac{B_0 j - B_1}{B_0 B_2 - B_1^2},$
 AND $p = 0, 1, 2, \quad B_p = \sum_j j^p / a_j, \quad a_j$ ARBITRARY NUMBERS.

- **WHAT ARE THE PERFORMANCE OF SUCH AN ESTIMATOR ?**
 WHEN APPLIED TO A SELF-SIMILAR. OR LRD PROCESS

WAVELETS AND 2ND-ORDER SCALING: ESTIMATION

Abry, Gonçalves, Flandrin,

Abry, Veitch

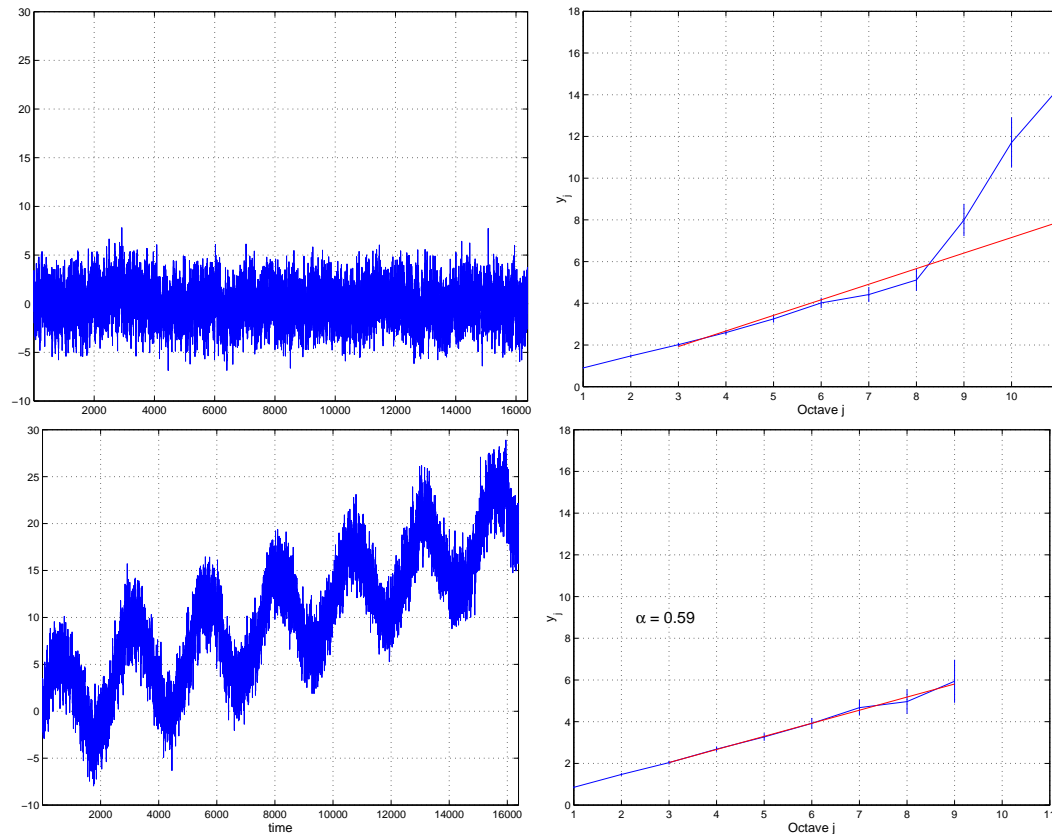
- **ASSUME:** - *i*) X GAUSSIAN,
- *ii*) IDEALISATION: EXACT INDEPENDENCE.
- **BIAS :** $\mathbf{E} \log_2 S_2(j) = \log_2 \mathbf{E} S_2(j) + \underbrace{\Gamma'(n_j/2) - \log_2(n_j/2)}_{g_j}.$
 $\Rightarrow \mathbf{E} \hat{H} = H + \frac{1}{2} \sum_j w_j g_j.$
- **VARIANCE:**
 - $\text{Var} \hat{H} = \frac{1}{4} \sum_j w_j^2 \sigma_j^2,$
 - $\min \text{Var} \hat{H} \implies a_j \propto \text{Var} \log_2 S_2(j)$
 - $\text{Var} \log_2 S_2(j) \simeq C/n_j \simeq 2^j C/n,$
 $\Rightarrow \text{VAR} \hat{H} \simeq \left((\log_2 e)^2 (\sum_j w_j^2 2^j) \right) / n,$
 \Rightarrow ANALYTICAL (APPROXIMATE) CONFIDENCE INTERVAL
(DOES NOT DEPEND ON UNKNOWN H).
- **ACTUAL PERFORMANCES :** NEGLIGIBLE BIAIS, EXTREMELY CLOSE TO MLE.
- **CONCEPTUAL AND PRACTICAL SIMPLICITY :** MATLAB CODE AVAILABLE.

WAV. AND 2ND-ORDER SCALING: ROBUSTNESS

Superimposed Trends

$$Y(t) = X(t) + T(t) \Rightarrow d_Y(j, k) = d_X(j, k) + d_T(j, k)$$

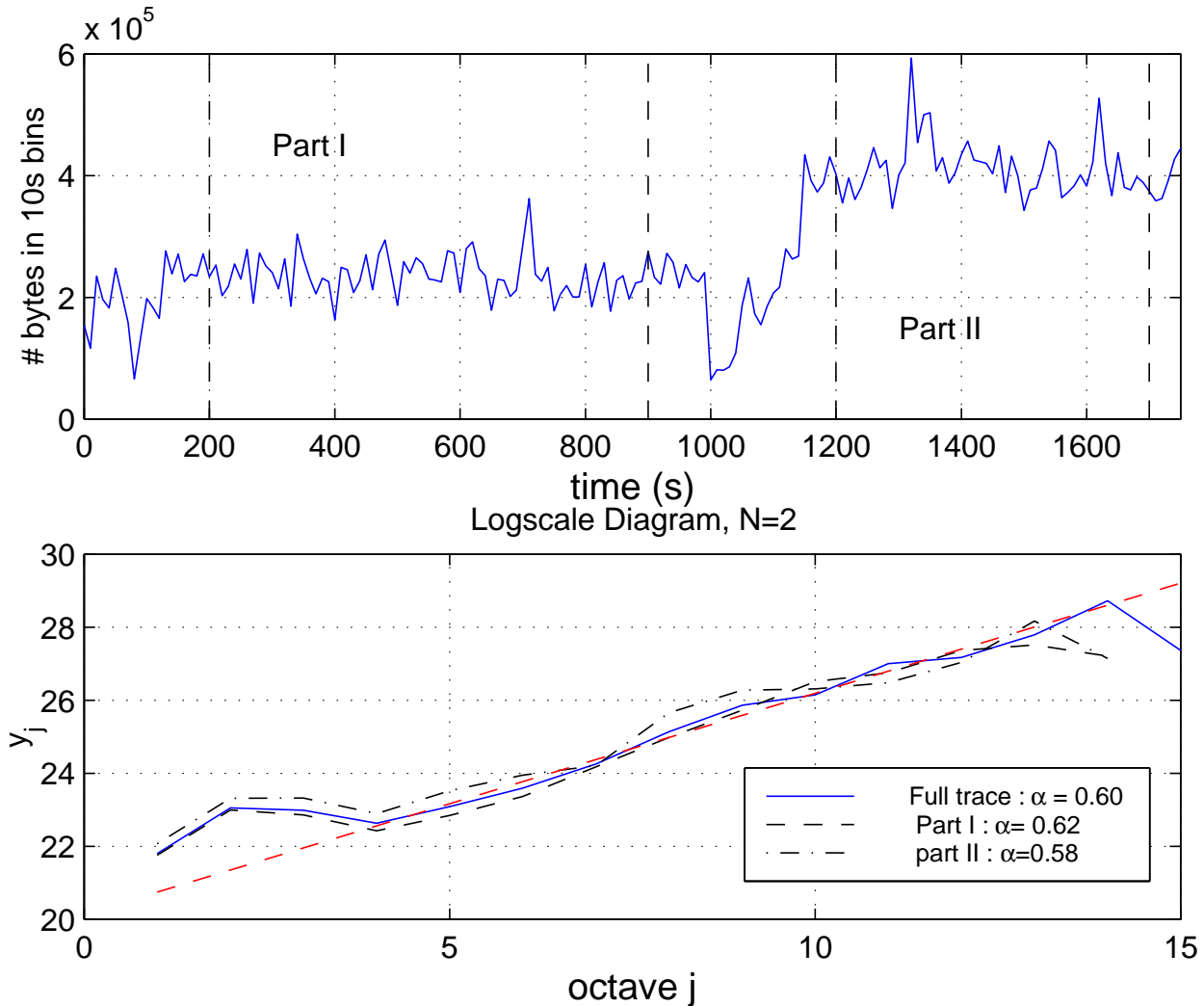
- If $T(t)$ Polynomial of degree P , then $d_T \equiv 0$ when $N > P$,
- If $T(t)$ smooth trend, then the d_T decrease as N increases.



Vary N !

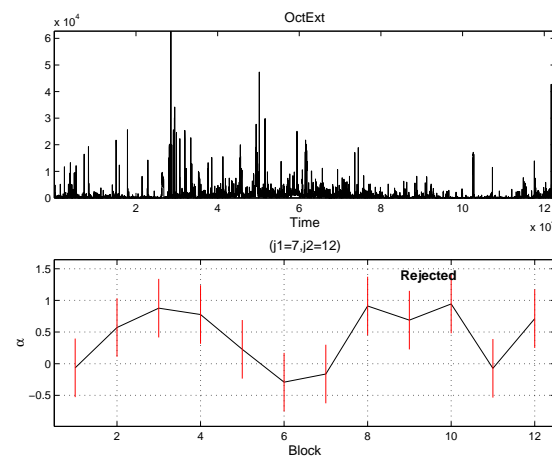
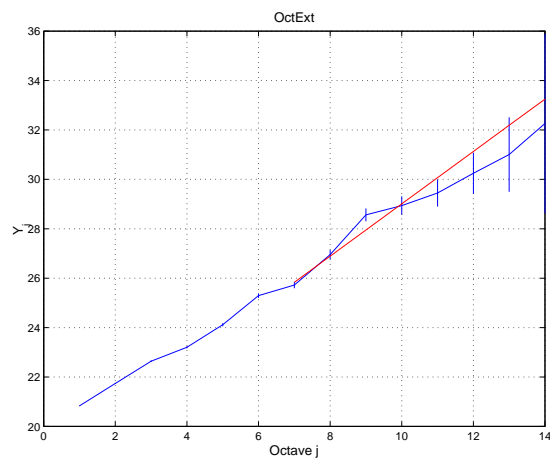
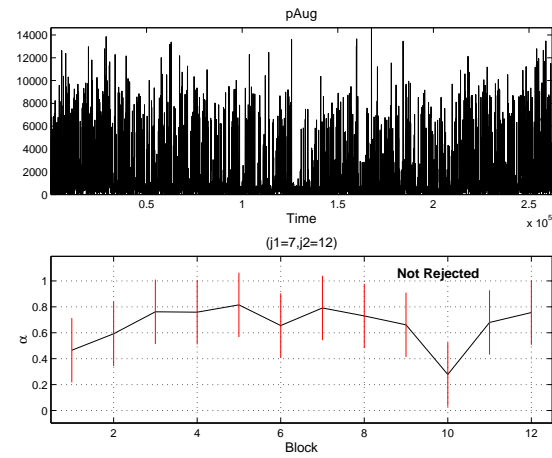
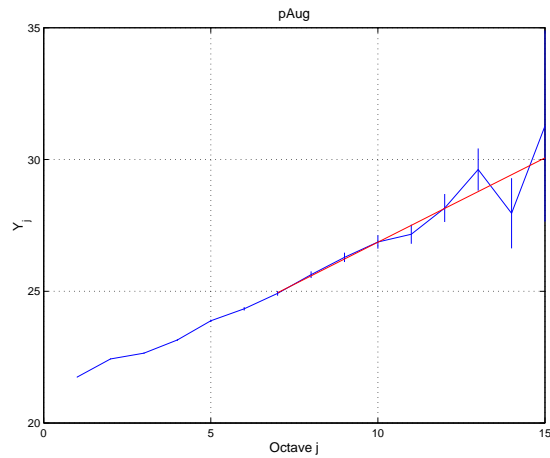
WAV. AND 2ND-ORDER SCALING: ROBUSTNESS

Superimposed Trends - Ethernet Data (Veitch, Abry)



WAV. AND 2ND-ORDER SCALING: ROBUSTNESS

Constancy along time of Scaling laws (Veitch, Abry)



SELF-SIMILARITY

- SELF-SIMILARITY:

$$\mathbf{E}|d_X(j, k)|^q = C_q(2^j)^{qH}$$

- Power Laws,
- $\forall 2^j$ (for all scales),
- $\forall q / \mathbf{E}|d_X(j, k)|^q < \infty$,
- A single parameter H
- Additive Structure.

- ?

- ?

BEYOND SELF-SIMILARITY

- SELF-SIMILARITY:

$$\mathbf{E}|d_X(j, k)|^q = C_q(2^j)^{qH}$$

- Power Laws,
- $\forall 2^j$ (for all scales),
- $\forall q / \mathbf{E}|d_X(j, k)|^q < \infty$,
- A single parameter H
- Additive Structure.

- MULTIFRACTAL

$$\mathbf{E}|d_X(j, k)|^q = C_q(2^j)^{\zeta(q)}$$

- Power Laws,
- $\forall 2^j < L$, (for fine scales only, in the limit $2^j \rightarrow 0$),
- $\forall q$?
- A whole collection of scaling parameter $\zeta(q)$
- Multiplicative Structure.

- ?

OUTLINE

I. INTUITIONS, MODELS, TOOLS

- I.1 INTUITIONS, DEFINITION, APPLICATIONS
- I.2 STOCHASTIC MODELS: SELF-SIMILARITY VS MULTIFRACTAL
- I.3 MULTIREOLUTION TOOLS, AGGREGATION, INCREMENTS
- I.4 WAVELETS, CONTINUOUS, DISCRETE

II. SECOND ORDER ANALYSIS, SELF SIMILARITY AND LONG MEMORY

- II.1 RANDOM WALKS, SELF SIMILARITY, LONG MEMORY,
- II.2 2ND ORDER WAVELET STATISTICAL ANALYSIS,
- II.3 ESTIMATION, ESTIMATION PERFORMANCE,
- II.4 ROBUSTNESS AGAINST NON STATIONARITIES,

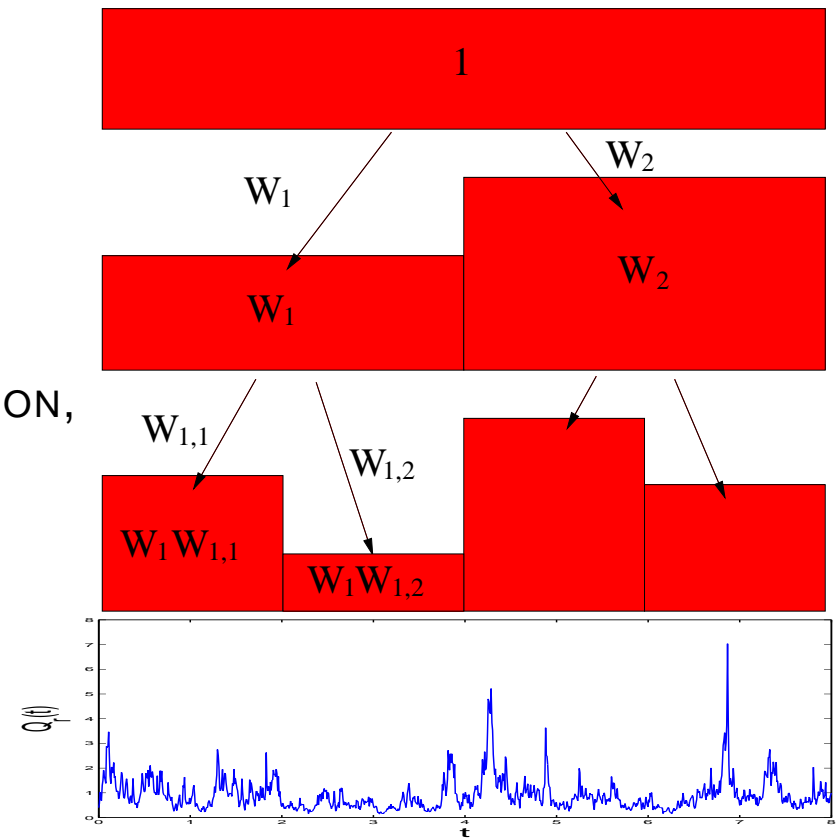
III. HIGHER ORDER ANALYSIS, MULTIFRACTAL PROCESSES

- III.1 MULTIPLICATIVE CASCADES, MULTIFRACTAL PROCESSES,
- III.2 HIGHER ORDER WAVELET STATISTICAL ANALYSIS,
- III.3 FINITENESS OF MOMENTS,
- III.4 ESTIMATION, ESTIMATION PERFORMANCE,
- III.5 NEGATIVE ORDERS,
- III.6 BEYOND POWER LAWS.

MODELLING TOOL 2: MULTIPLICATIVE CASCADES

● DEFINITION:

- SPLIT DYADIC INTERVALS $I_{j,k}$ INTO TWO,
- I.I.D. MULTIPLIERS $W_{j,k}$
- $Q_J(t) = \prod_{\{(j,k): 1 \leq j \leq J, t \in I_{j,k}\}} W_{j,k}$,
- FROM DENSITY TO MEASURE...
- ... AND TO FRACTIONAL BROWNIAN MOTION,
- IN MULTIFRACTAL TIME.



● IMPLICATIONS:

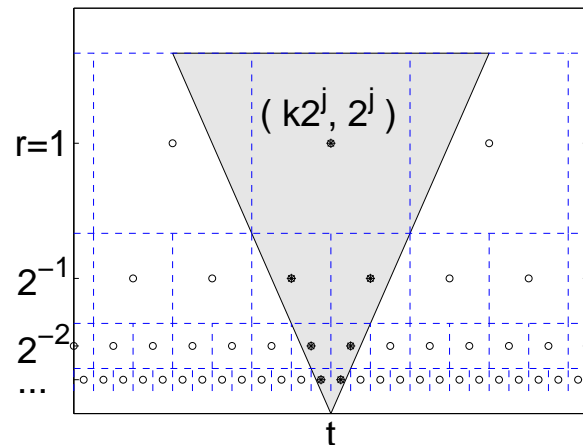
- NON STATIONARITY,
- LOCAL HOLDER EXPONENT,
- MULTIFRACTAL SAMPLE PATHS, MULTIFRACTAL SPECTRUM $D(h)$
- CASCADES, MULTIPLICATIVE STRUCTURE,
- $\mathbb{E} \left(\frac{1}{a} \int_t^{t+a\tau_0} X(u) du \right)^q = C_q |a|^{\zeta_q}$,
- MULTIPLE EXPONENTS ζ_q , FINITE RANGE OF q ,
- $\zeta_q = -\log_2 \mathbb{E} W^q$, NON LINEAR IN q ,
- FINE SCALES $a \rightarrow 0$, $a \ll L$ INTEGRAL SCALE,
- NO CHARACTERISTIC SCALE OF TIME BEYOND AN INTEGRAL SCALE.

MODELLING TOOL 2: MULTIPLICATIVE CASCADES

YAGLOM, MANDELBROT

MANDELBROT'S
CASCADE (CMC)

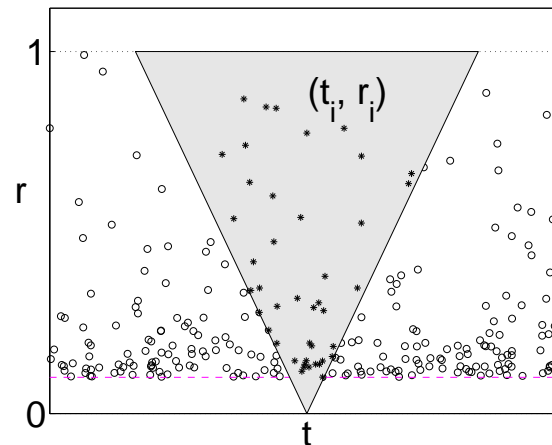
- IID W ,
- DYADIC GRID,



BARRAL, MANDELBROT

COMPOUND POISSON
CASCADE (CPC)

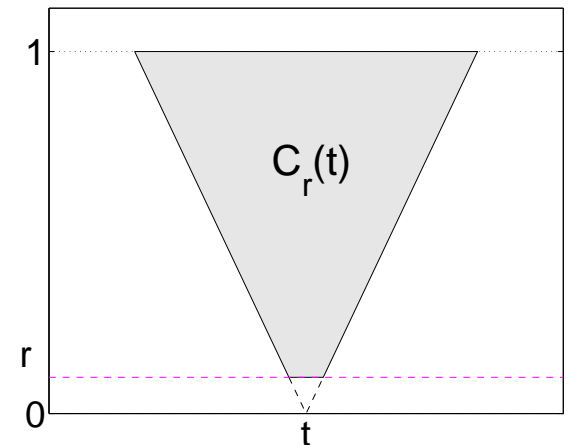
- IID W ,
- POINT PROCESS,



SCHMITT ET AL.,
BACRY ET AL., CHAINAIS ET AL.

INFINITELY DIVISIBLE
CASCADE (IDC)

- CONTINUOUS INFINITELY
DIVISIBLE MEASURE M ,



$$Q_r(t) = \prod W_{j,k}, \quad \prod W_{j,k}, \quad \exp \int dM(t', r'),$$

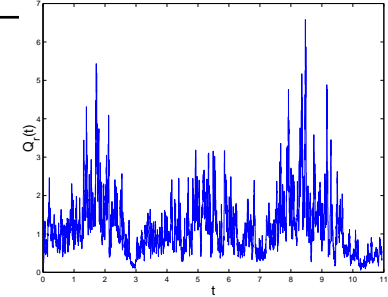
$$\varphi(q) = -\log_2 \mathbf{E}W^q, \quad = -q(1 - \mathbf{E}W) + 1 - \mathbf{E}W^q, \quad = \rho(q) - q\rho(1),$$

$$A(t) = \lim_{r \rightarrow 0} \int_0^t Q_r(u) du,$$

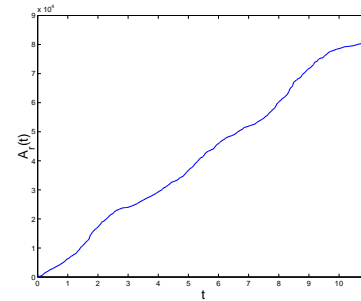
FOR A RANGE OF q S, $\mathbf{E}|A(t + a\tau_0) - A(t)|^q = c_q |a|^{q+\varphi(q)}$,
RESOLUTION DEPTH < SCALE < INTEGRAL SCALE, $a_m = r < a < a_M = L$.

MULTIFRACTAL PROCESSES

DENSITY: $Q_r(t) = \Pi W_{j,k}$
 $\mathbb{E} \left(\frac{1}{a} \int_t^{t+a\tau_0} Q_r(u) du \right)^q = c_q a^{\varphi(q)},$



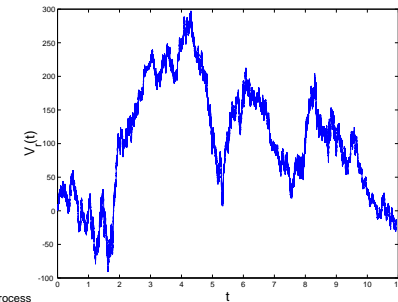
MEASURE: $A(t) = \lim_{r \rightarrow 0} \int_0^t Q_r(u) du,$
 $\mathbb{E} |A(t + a\tau_0) - A(t)|^q = c_q |a|^{q+\varphi(q)},$



FRACTIONAL BROWNIAN MOTION IN MULTIFRACTAL TIME:

$$V_H(t) = B_H(A(t)),$$

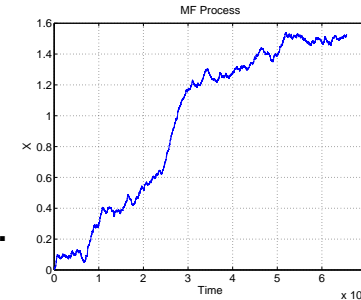
$$\mathbb{E} |V_H(t + a\tau_0) - V_H(t)|^q = c_q |a|^{qH+\varphi(qH)},$$



MULTIFRACTAL RANDOM WALK:

$$Y_H(t) = \int^t Q_r(s) dB_H(s),$$

$$\mathbb{E} |Y_H(t + a\tau_0) - Y_H(t)|^q = c_q |a|^{qH+\varphi(q)}.$$



MATLAB SYNTHESIS ROUTINES : CHAINAIS, ABRY

TEST FOR THE FINITENESS OF MOMENTS

GONÇALVÈS, RIEDI

TEST FOR THE FINITENESS OF MOMENTS

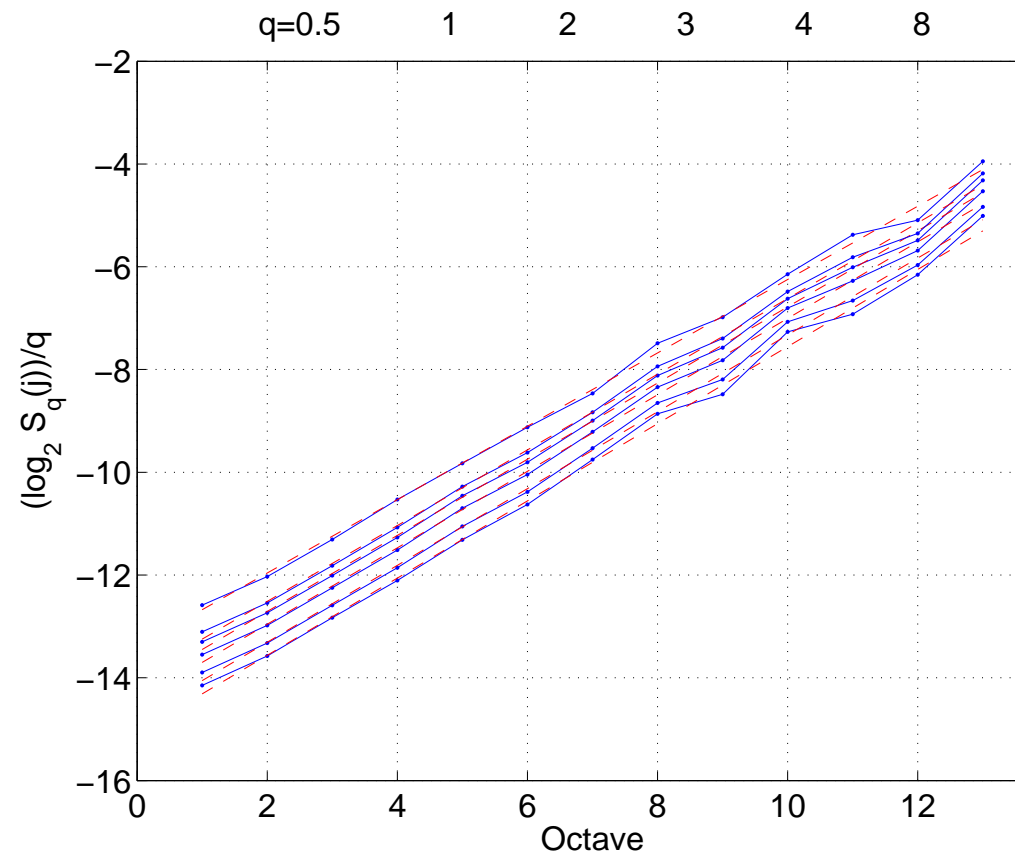
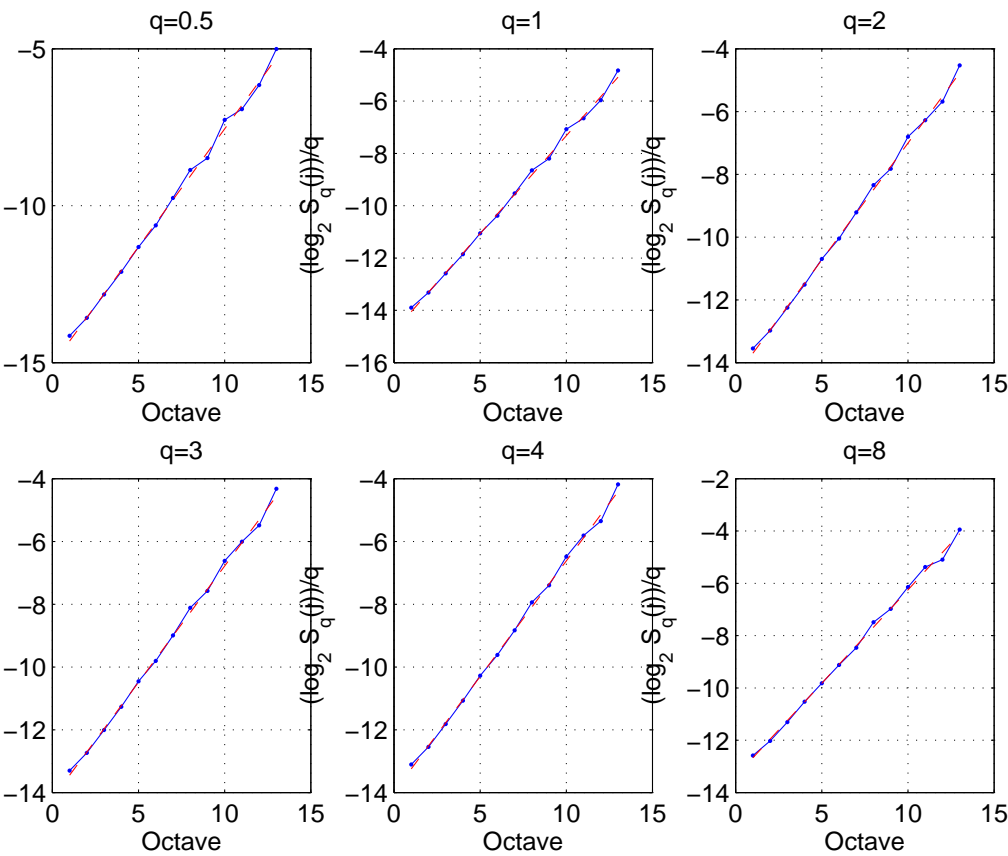
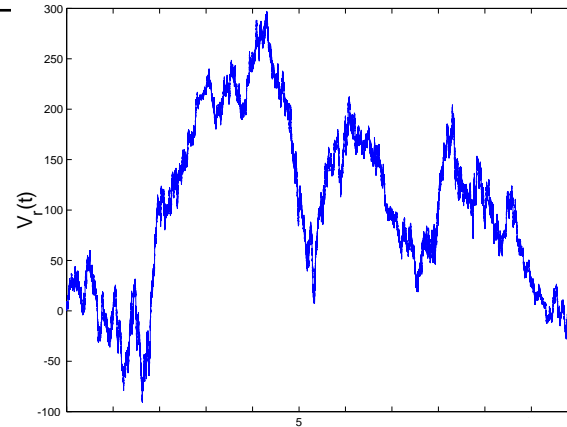
HIGHER-ORDER WAVELET STATISTICAL ANALYSIS

PRINCIPLES :

- IDEAS : **P1** $\Rightarrow \mathbf{E}|d_X(j, k)|^q = \mathbf{E}|d_X(0, k)|^q 2^{j\zeta_q}$
 $\Rightarrow \log_2 \mathbf{E}|d_X(j, k)|^q = j\zeta_q + \beta_q,$
- PROBLEMS: ESTIMATE $\mathbf{E}|d_X(j, k)|^q$ FROM A SINGLE FINITE LENGTH OBSERVATION ?
- SOLUTION : **P2 et P3** \Rightarrow STATISTICAL AVERAGES \Rightarrow TIME AVERAGES,
 $S_q(j) = (1/n_j) \sum_{k=1}^{n_j} |d_X(j, k)|^q$

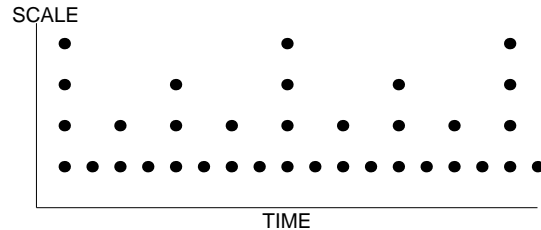
LOG-SCALE DIAGRAMS: $\log_2 S_q(j)$ vs $\log_2 2^j = j$

LOGSCALE DIAGRAMS - MULTIFRACTAL PROC.



WAV. AND HIGHER-ORDER SCALING: ESTIMATION

- **DYADIC GRID** (DISCRETE WAVELET TRANSFORM): $a_j = 2^j$, $t_{j,k} = k2^j$,



- **STRUCTURE FUNCTIONS (TIME AVERAGE):**

$$S_q(j) = (1/n_j) \sum_{k=1}^{n_j} |d_X(j, k)|^q$$

- **DEFINITION:**

$$Y_{j,q,n} = \log_2 S_n(2^j, q; f_0) \text{ VERSUS } \log_2 2^j = j,$$

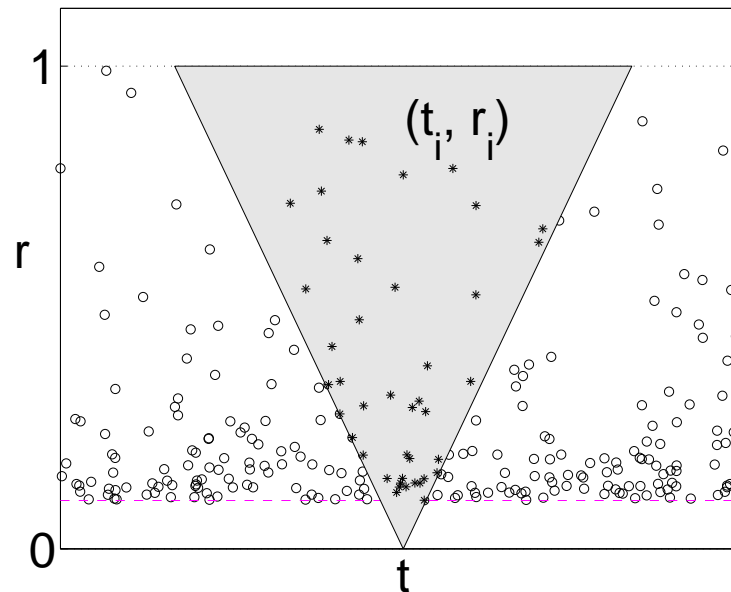
$$\hat{\zeta}(q, n) = \sum_{j=j_1}^{j_2} w_{j,q} Y_{j,q,n} .$$

NON WEIGTHED: $a_j = cste$

- **WHAT ARE THE PERFORMANCE OF SUCH ESTIMATORS ?**
WHEN APPLIED TO MULTIFRACTAL PROCESSES

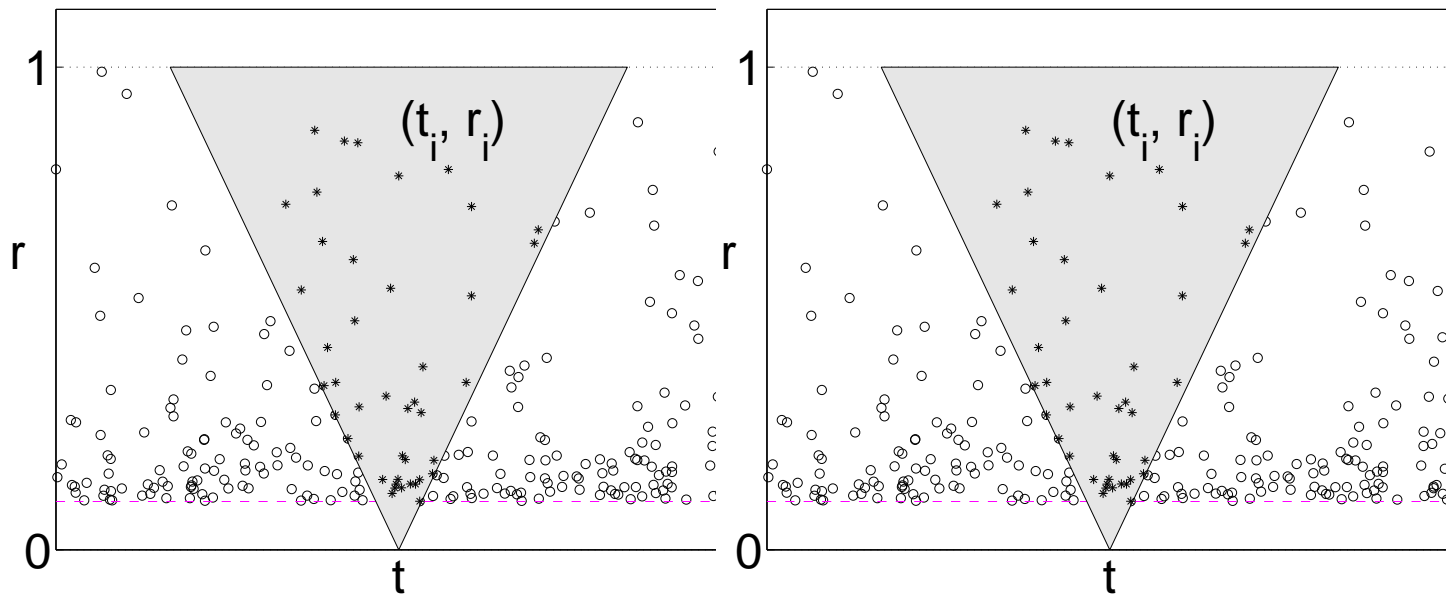
METHODOLOGY

- **NUMERICAL SYNTHESIS OF PROCESSES:**
 - ACCUMULATE n_{breal} NUMERICAL REPLICATIONS WITH LENGTH n SAMPLES.
- **APPLY SCALING EXPONENTS ESTIMATORS:**
 - COMPUTE $\hat{\zeta}(q, n)_{(l)}$ FOR EACH REPLICATION,
 - AVERAGE OVER REPL. TO OBTAIN THE STATISTICAL PERFORMANCE OF $\hat{\zeta}(q, n)$
- **ASYMPTOTIC BEHAVIOURS:**
 - THE CASCADE DEPTH INCREASES FOR A GIVEN NUMBER OF INTEGRAL SCALES.
 - ... ,



METHODOLOGY

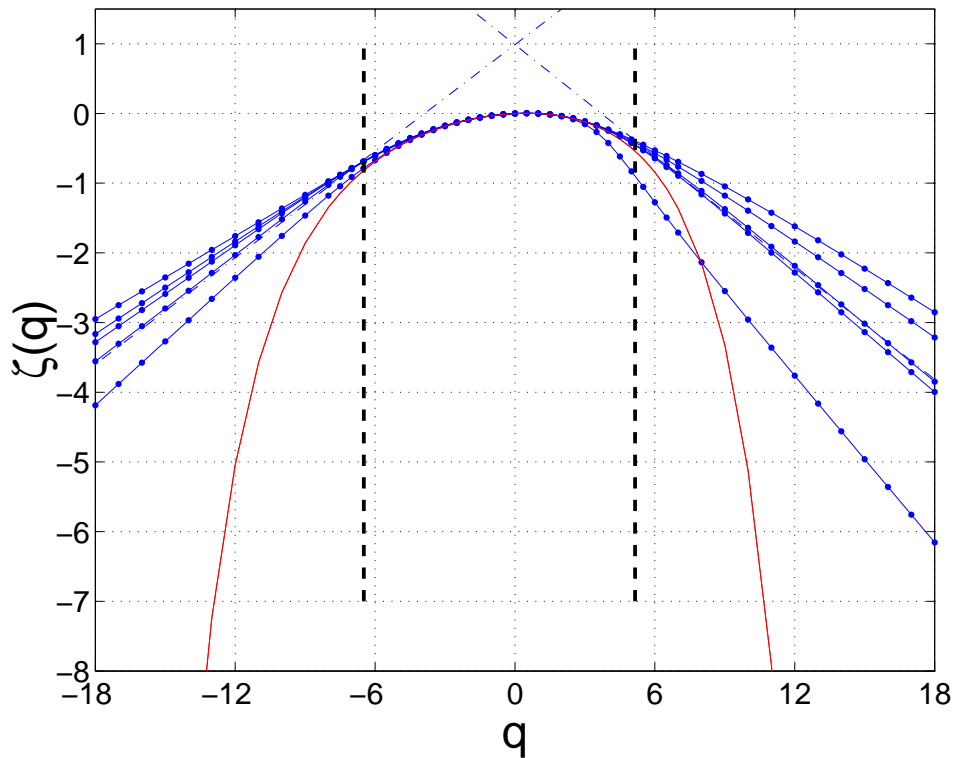
- **NUMERICAL SYNTHESIS OF PROCESSES:**
 - ACCUMULATE n_{real} NUMERICAL REPLICATIONS WITH LENGTH n SAMPLES.
- **APPLY SCALING EXPONENTS ESTIMATORS:**
 - COMPUTE $\hat{\zeta}(q, n)_{(l)}$ FOR EACH REPLICATION,
 - AVERAGE OVER REPL. TO OBTAIN THE STATISTICAL PERFORMANCE OF $\hat{\zeta}(q, n)$
- **ASYMPTOTIC BEHAVIOURS:**
 - THE CASCADE DEPTH INCREASES FOR A GIVEN NUMBER OF INTEGRAL SCALES.
 - THE NUMBER OF INTEGRAL SCALES INCREASES FOR A GIVEN CASCADE DEPTH,



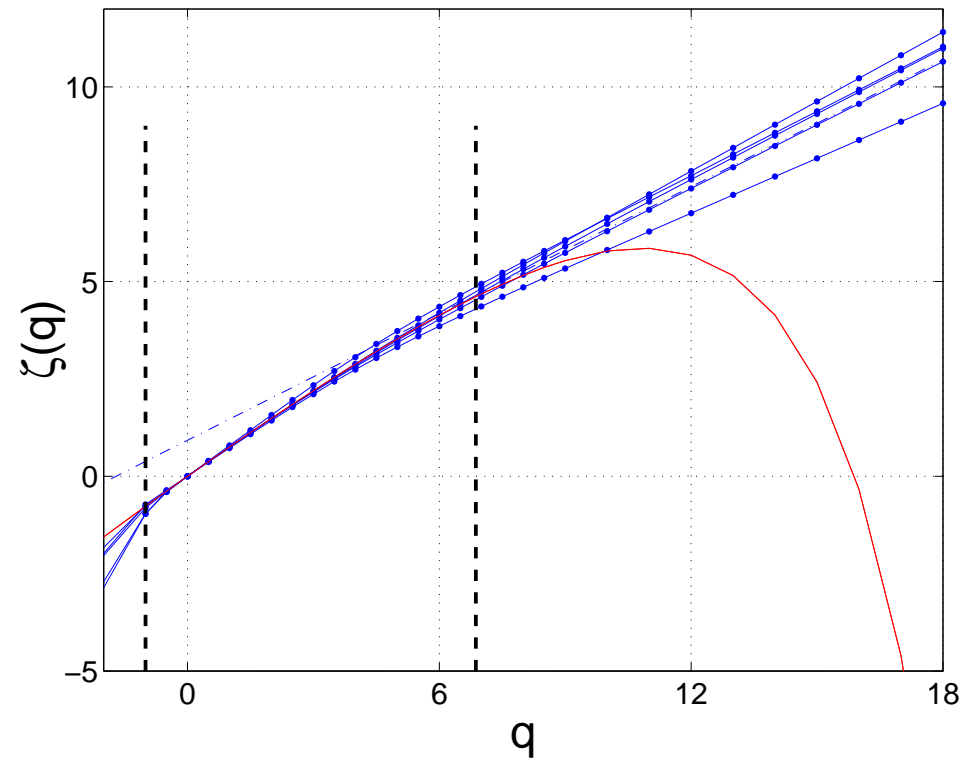
LINEARISATION EFFECT: $\hat{\zeta}(q)$

LASHERMES, ABRY, CHAINAIS

CPC $Q_r EI(1)$



CPC $V_H EIII(3)$

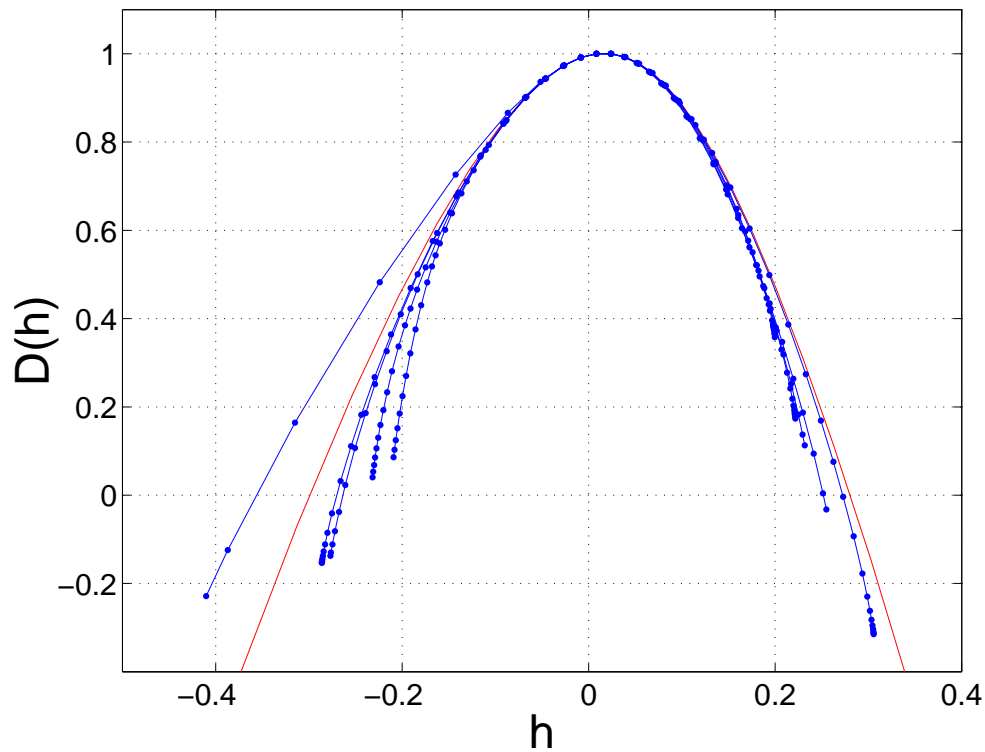


$q > q_o, \hat{\zeta}(q, n) = \alpha_o + \beta_o q, q_o, \alpha_o, \beta_o$ ARE RV.

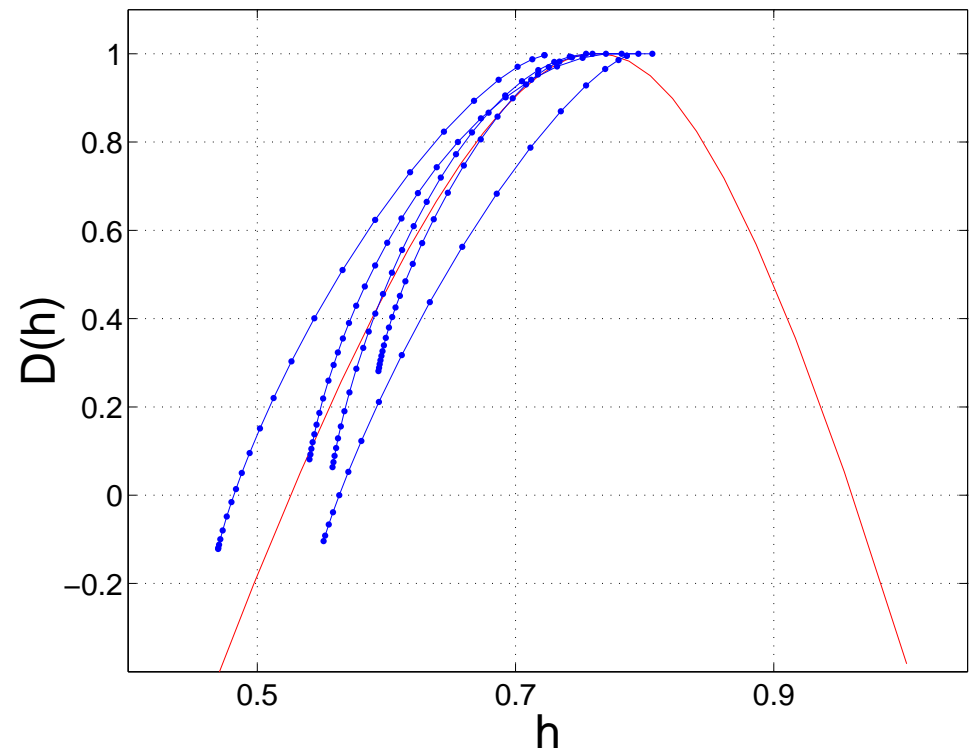
LINEARISATION EFFECT: LEGENDRE TRANSFORM

$$D(h) = d + \text{MIN}_q(qh - \zeta(q)), \text{ (} d \text{ EUCLIDIEN DIMENSION OF SPACE).}$$

CPC Q_r $EI(1)$



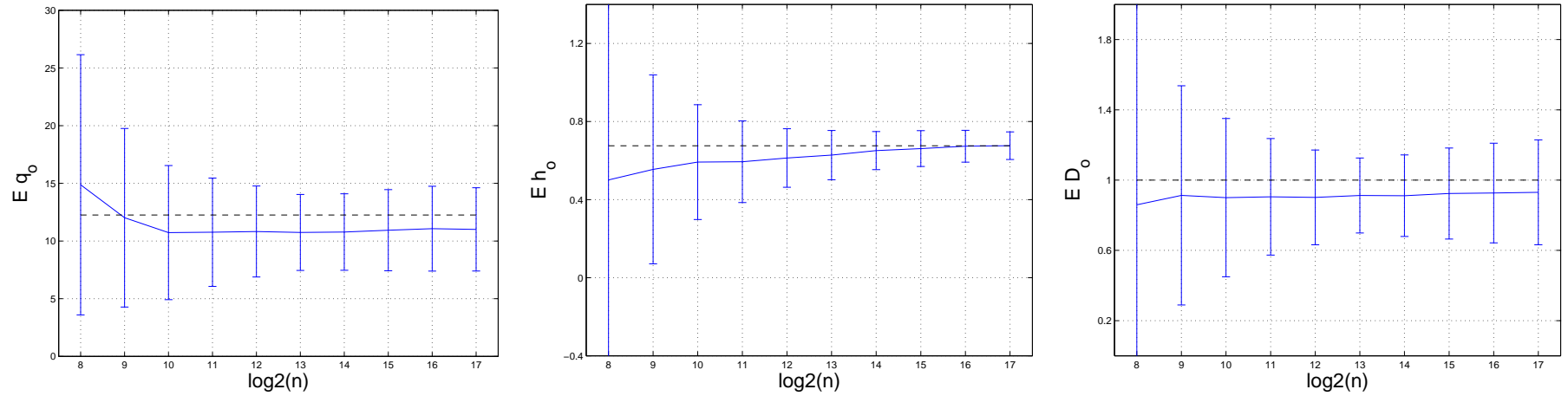
CPC V_H $EIII(3)$



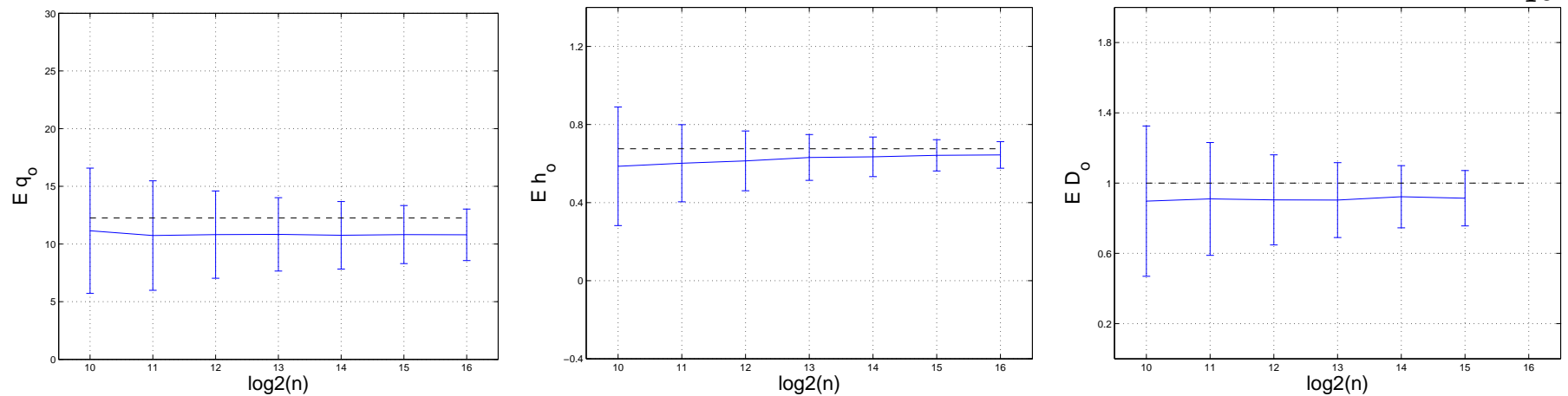
ACCUMULATION POINTS : $D_o(h_o)$, WITH $D_o = d - \alpha_o$, $h_o = \beta_o$,
 D_o, h_o ARE RV.

LIN. EFFECT: ASYMPTOTIC BEHAVIOURS

- GIVEN RESOLUTION, INCREASING NUMBER OF INTEGRAL SCALES,



- GIVEN NUMBER OF INTEGRAL SCALES, INCREASING RESOLUTION,



LINEARISATION EFFECT: CONJECTURE

• CRITICAL POINTS:

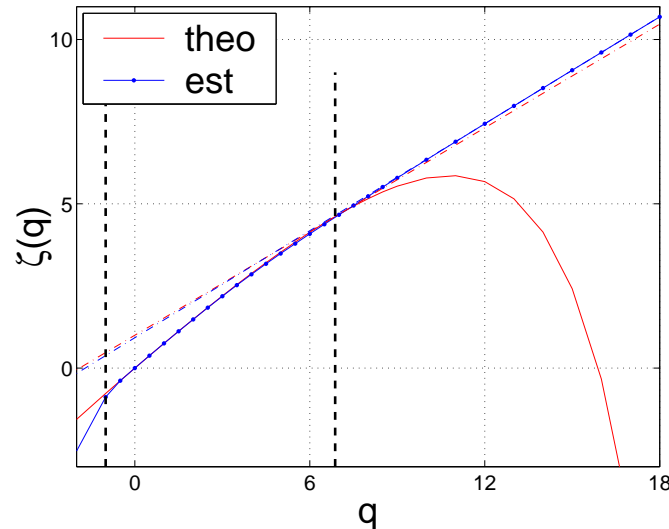
$$\begin{cases} D_*^\pm & = 0, \\ D(h_*^\pm) & = 0, \\ h_*^\pm & = (d\zeta(q)/dq)_{q=q_*^\pm}. \end{cases}$$

• RESULTS:

$$EI : \begin{cases} \hat{\zeta}(q, n) = d - D_o^- + h_o^- q & \rightarrow d - D_*^- + h_*^- q, & q \leq q_*^-, \\ \hat{\zeta}(q, n) & \rightarrow \zeta(q), & q_*^- \leq q \leq q_*^+, \\ \hat{\zeta}(q, n) = d - D_o^+ + h_o^+ q & \rightarrow d - D_*^+ + h_*^+ q, & q_*^+ \leq q. \end{cases}$$

$$EII \& III : \begin{cases} \hat{\zeta}(q, n) & \rightarrow \zeta(q), & 0 < q \leq q_*^+, \\ \hat{\zeta}(q, n) = d - D_o^+ + h_o^+ q & \rightarrow d - D_*^+ + h_*^+ q, & q_*^+ \leq q. \end{cases}$$

• ILLUSTRATION:

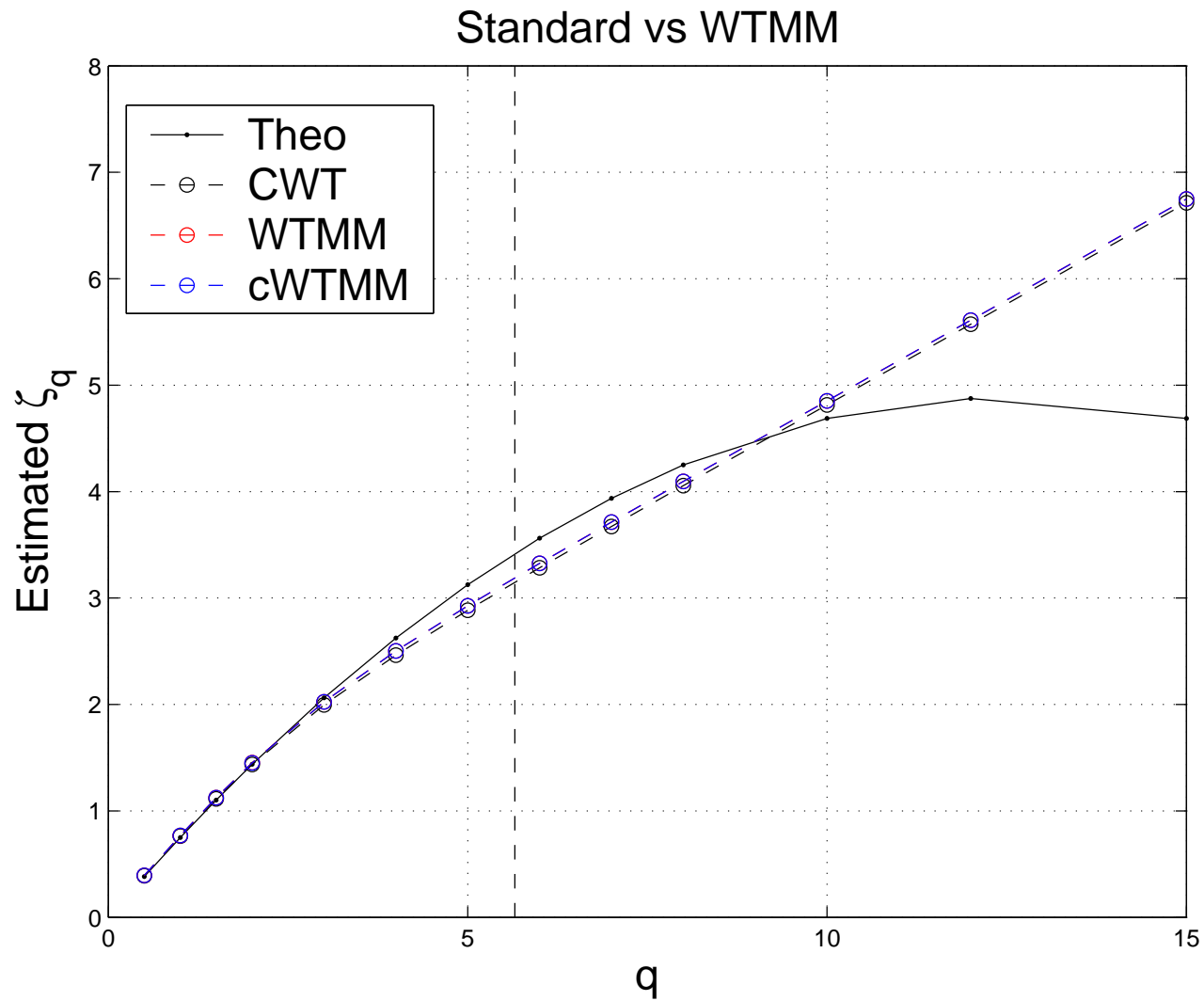


LINEARISATION EFFECT: COMMENTS

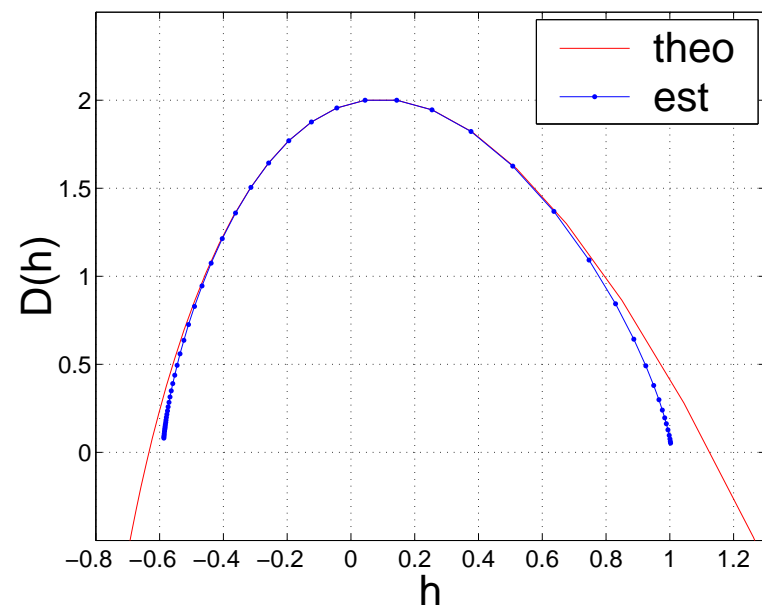
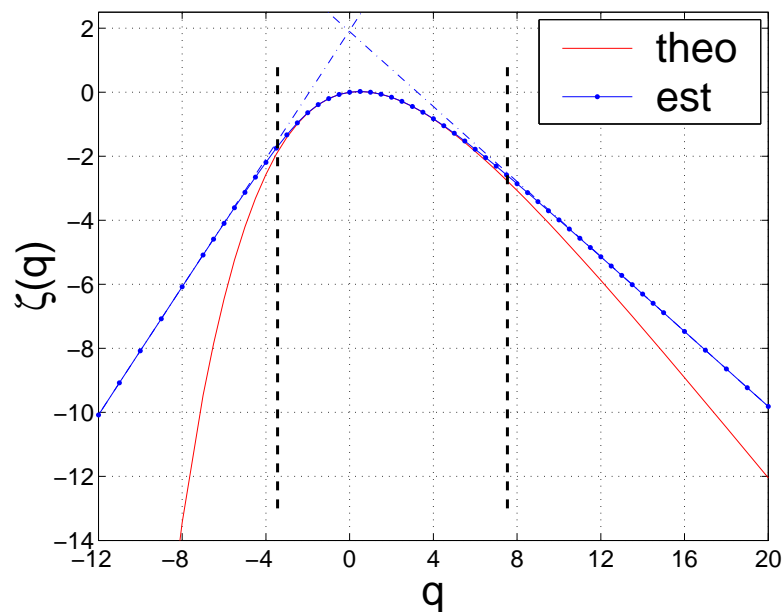
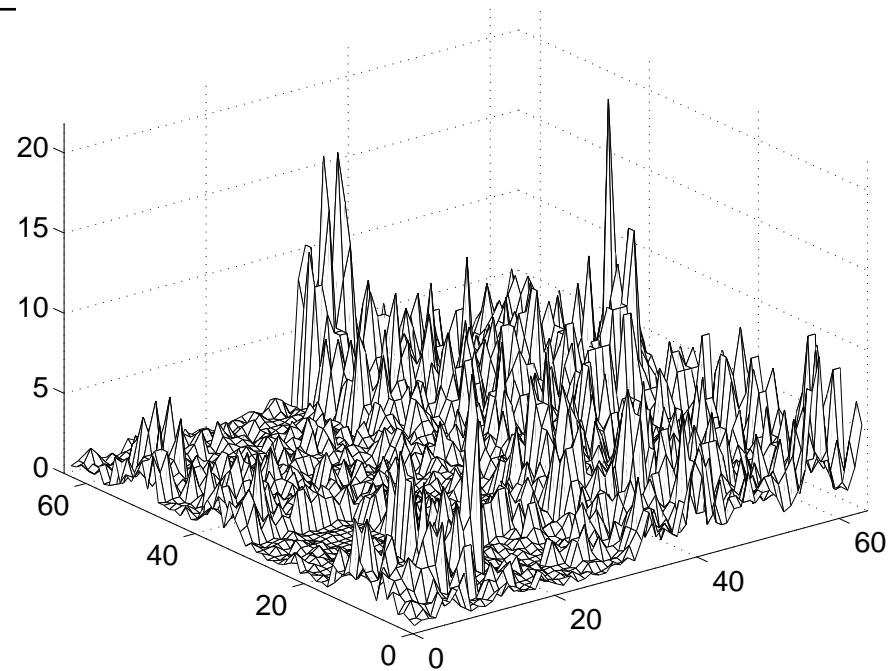
WHEN DOES THE LINEARISATION EFFECT EXIST ?

- FOR ALL TYPES OF CASCADES: CMC, CPC, IDC,
- FOR ALL TYPES OF PROCESSES: $Q_r, A, V_H, Y_H,$
- FOR ALL NUMBERS OF VANISHING MOMENTS: $N \geq 1,$
- FOR ALL MRA-BASED ESTIMATORS: WAVELETS, INCREMENTS, AGGREGATION,
- CAN BE WORKED OUT FOR $q < 0,$
- EXTENDS TO DIMENSION HIGHER THAN $d > 1.$

EXTENSION: STANDARD WT VERSUS WTMM (1/3).

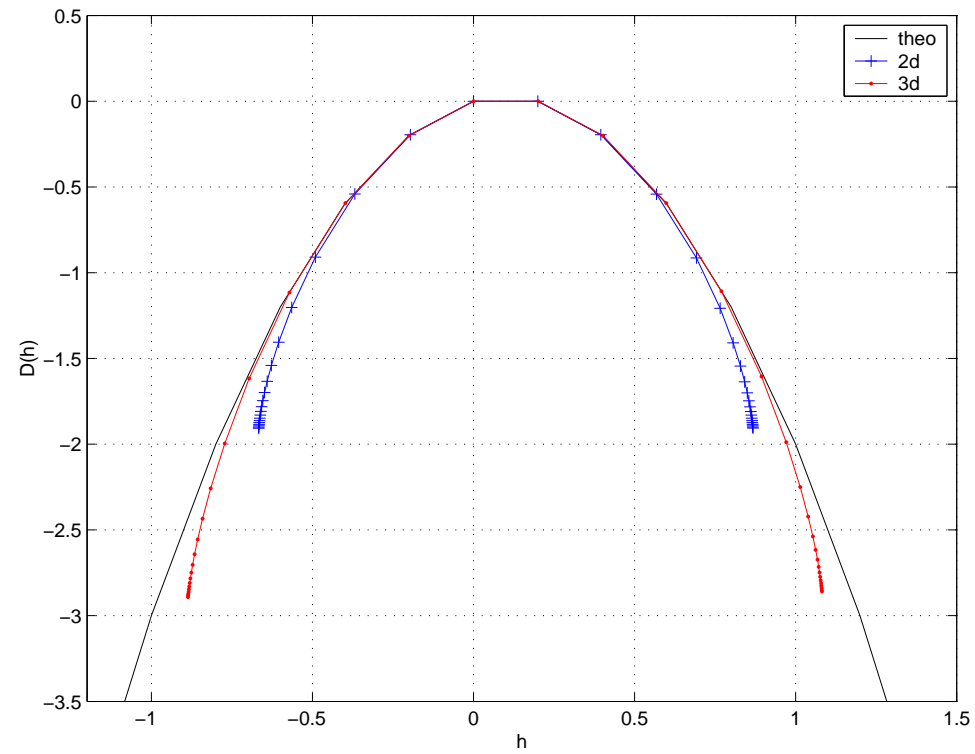
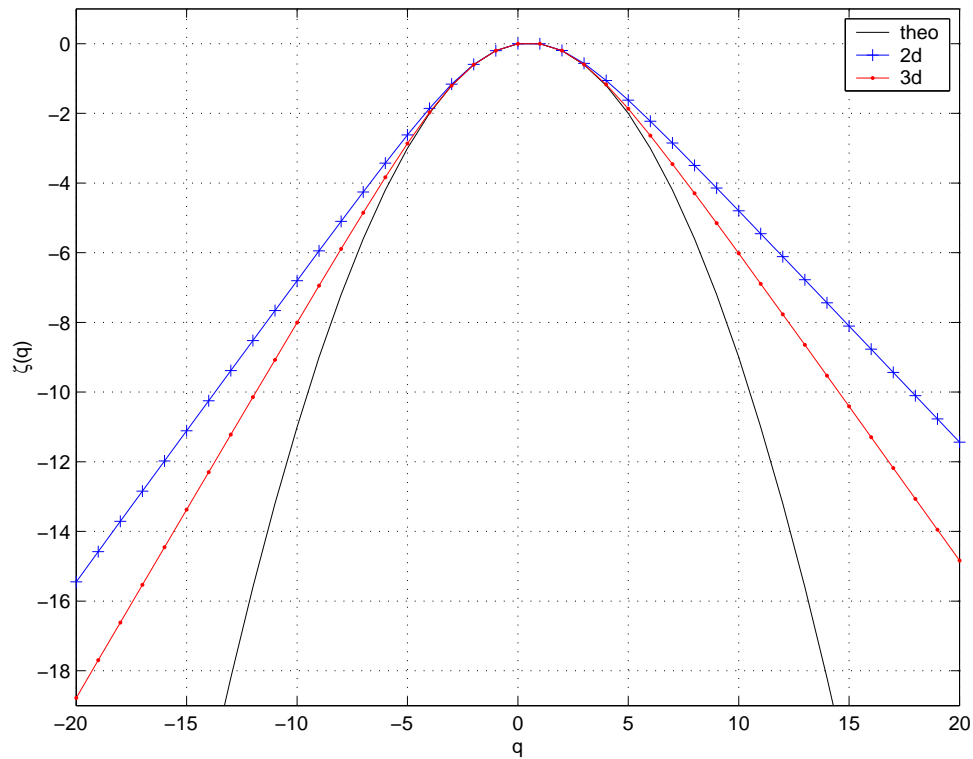


EXTENSION: 2D MULTIPLICATIVE CASCADE (2/3).



EXTENSION: 3D MULTIPLICATIVE CASCADE (3/3).

3D CMC (LOG NORMAL), EI(1) COMPARED TO A 2D SLICE.



LINEARISATION EFFECT: COMMENTS

WHEN DOES THE LINEARISATION EFFECT EXIST ?

- FOR ALL TYPES OF CASCADES: CMC, CPC, IDC,
- FOR ALL TYPES OF PROCESSES: $Q_r, A, V_H, Y_H,$
- FOR ALL NUMBERS OF VANISHING MOMENTS: $N \geq 1,$
- FOR ALL MRA-BASED ESTIMATORS: WAVELETS, INCREMENTS, AGGREGATION,
- CAN BE WORKED OUT FOR $q < 0,$
- EXTENDS TO DIMENSION HIGHER THAN $d > 1.$

WHAT THE LINEARISATION EFFECT IS NOT:

- A LOW PERFORMANCE ESTIMATION EFFECT.
- A FINITE SIZE EFFECT : THE CRITICAL PARAMETERS DO NOT DEPEND ON $n,$
BE IT THE NUMBER OF INTEGRAL SCALES,
OR THE DEPTH (OR RESOLUTION) OF THE CASCADES.
- A FINITENESS OF MOMENTS EFFECT,
 - $q_c^- < 0 < 1 < q_c^+, q - 1 + \varphi(q) = 0,$
 - $q_c^- < q_*^- < 0 < 1 < q_*^+ < q_c^+,$

WHAT THE LINEARISATION EFFECT MIGHT BE:

- MULTIPLICATIVE MARTINGALES ?
- OSSIANDER, WAYMIRE 00, KAHANE, PEYRIÈRE 75, BARRAL, MANDELBROT 02.

LINEARISATION EFFECT: PICTURE

- TWO POWER-LAWS, TWO FUNCTIONS OF q :

- BARE CASCADE:

$$\mathbb{E}Q_r(t)^q = r^{\varphi(q)}, \quad q \in \mathcal{R}.$$

- DRESSED CASCADE:

$$\left. \begin{aligned} \mathbb{E}T_{Q_0}(t, \mathbf{a}; \beta_0)^q &= c_q |\mathbf{a}|^{\zeta(q)}, & q \in [q_c^-, q_c^+], \\ \mathbb{E}T_{Q_0}(t, \mathbf{a}; \beta_0)^q &= \infty, & \text{ELSE,} \end{aligned} \right\}$$

WITH:

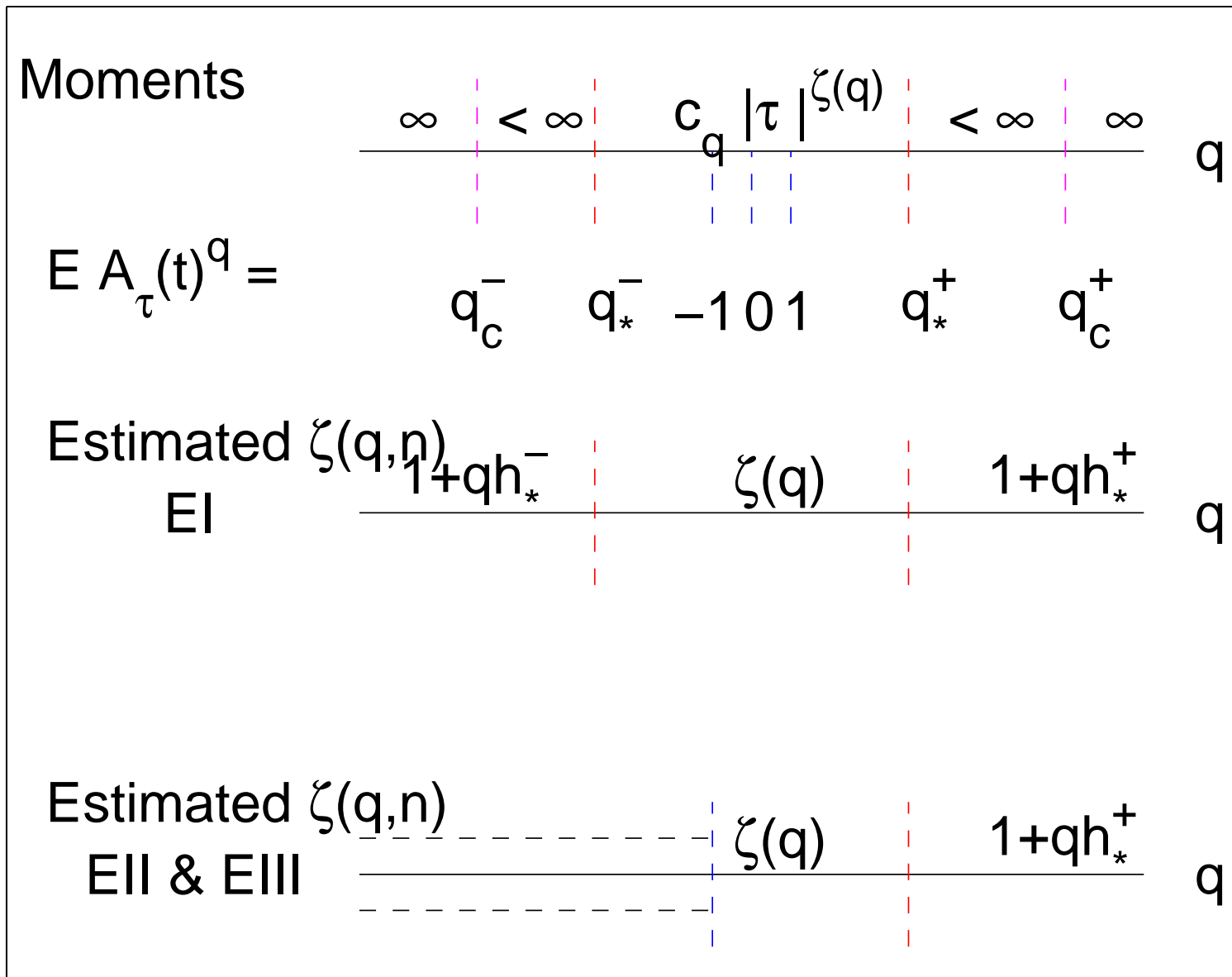
$$\left. \begin{aligned} \zeta(q) &= 1 + qh_*^-, & q \in [q_c^-, q_*^-], \\ \zeta(q) &= \varphi(q), & q \in [q_*^-, q_*^+], \\ \zeta(q) &= 1 + qh_*^+, & q \in [q_*^+, q_c^+]. \end{aligned} \right\}$$

- CONFUSION BETWEEN $\varphi(q)$ AND $\zeta(q)$:

- MULTIPLICATIVE CASCADE: $\varphi(q)$, $q \in \mathcal{R}$,

- SCALING EXPONENTS: $\zeta(q)$, $q \in [q_c^-, q_c^+]$.

LINEARISATION EFFECT: SKETCHED VIEWS



LINEARISATION EFFECT: IMPACTS AND IMPORTANCE

CONSEQUENCES: RECAST THE USUAL GOALS :

- ESTIMATE THE INTEGRAL SCALE AND THE RESOLUTION OF THE CASCADE,
⇒ I.E., FIND A SCALING RANGE $[a_m, a_M]$
- ESTIMATE THE CRITICAL PARAMETERS $D_*^\pm, h_*^\pm, q_*^\pm$,
- ESTIMATE THE $\zeta(q)$ FOR $q \in [q_*^-, q_*^+]$,
→ VISIT B. LASHERMES'S POSTER.

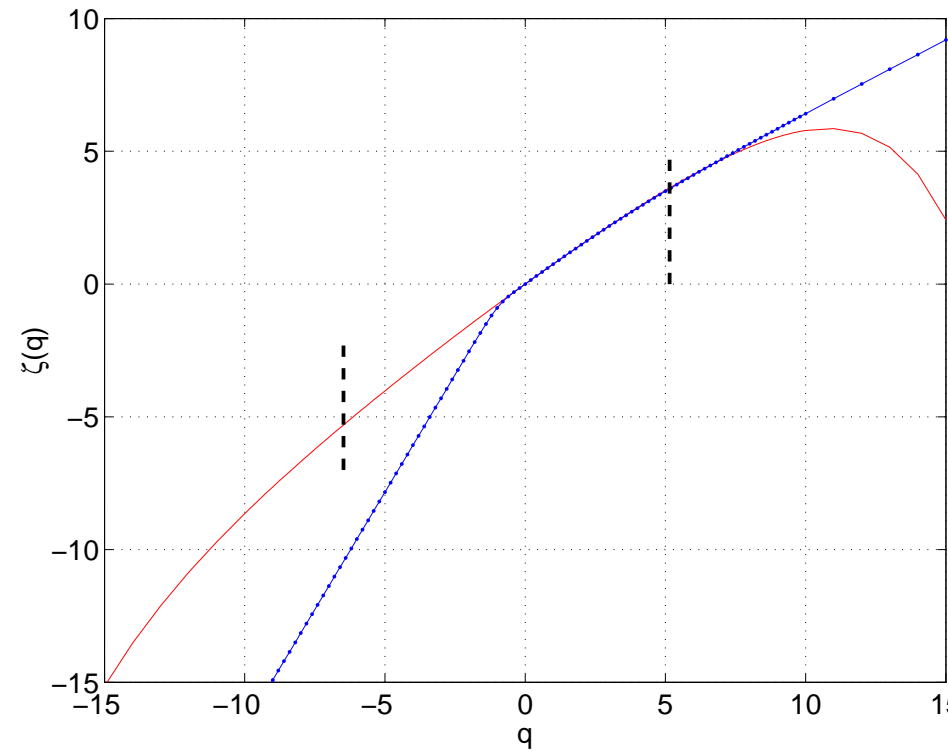
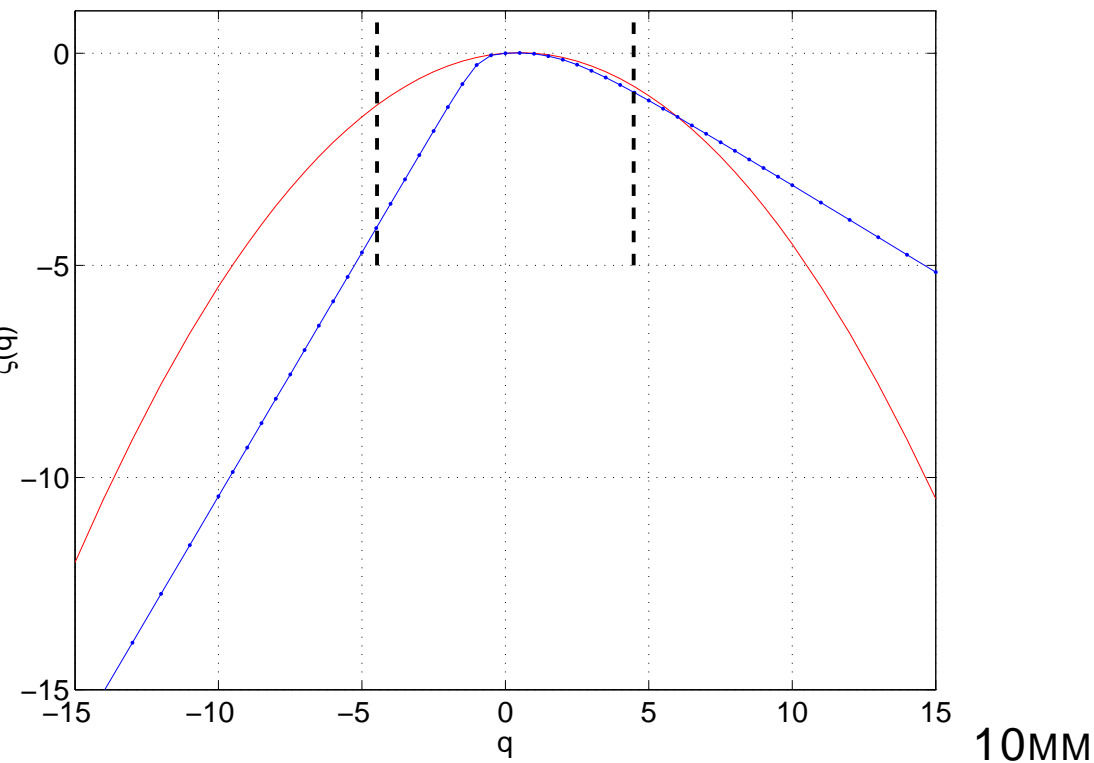
IMPORTANCE OF THE LINEARISATION EFFECT:

- DISCRIMINATION OF MF MODELS BASED ON $\hat{\zeta}(q, n)$,
- DISCRIMINATION BETWEEN MONOFRACTAL AND MULTIFRACTAL,

NEGATIVE VALUES OF q

DIFFICULTIES ?

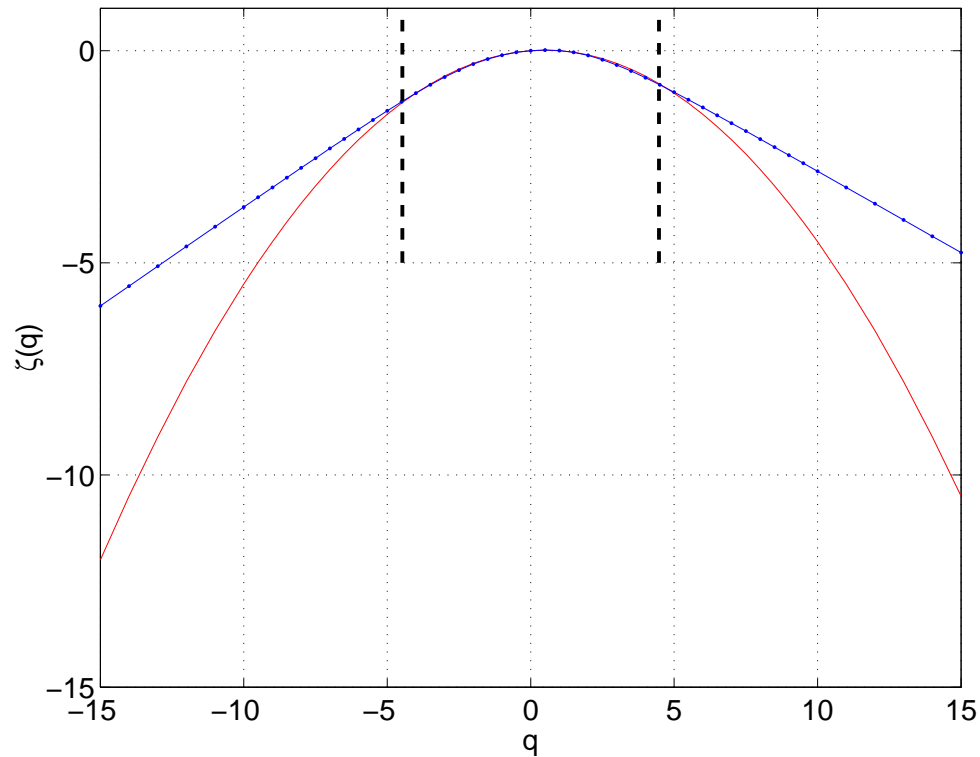
- FINITENESS ? $S_q(j) = (1/n_j) \sum_{k=1}^{n_j} |d_X(j, k)|^q < \infty$?
- NUMERICAL INSTABILITY ? $d_X(j, k) \simeq 0 \rightarrow |d_X(j, k)|^q = \infty$
- THEORY ? WEAK HÖLDER EXPONENT VS EXACT HÖLDER EXPONENT



SOLUTIONS ?

NEGATIVE VALUES OF q S - SOLUTION 1

AGGREGATION: $T_X(a, t) = \frac{1}{aT_0} \int_t^{t+aT_0} X(u) du$



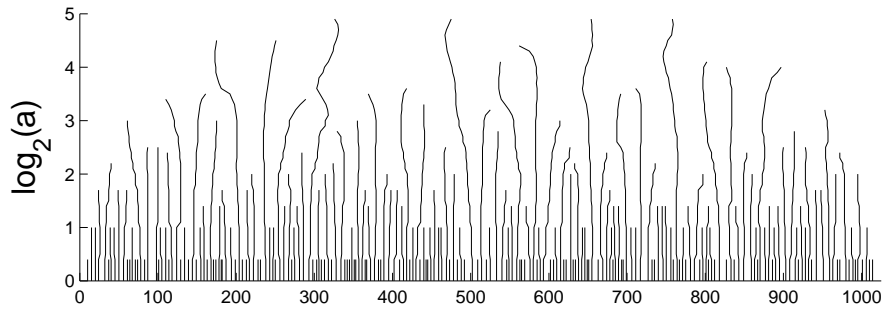
APPLIES ONLY TO POSITIVE DATA (MEASURE)

NEGATIVE VALUES OF q S - SOLUTION 2

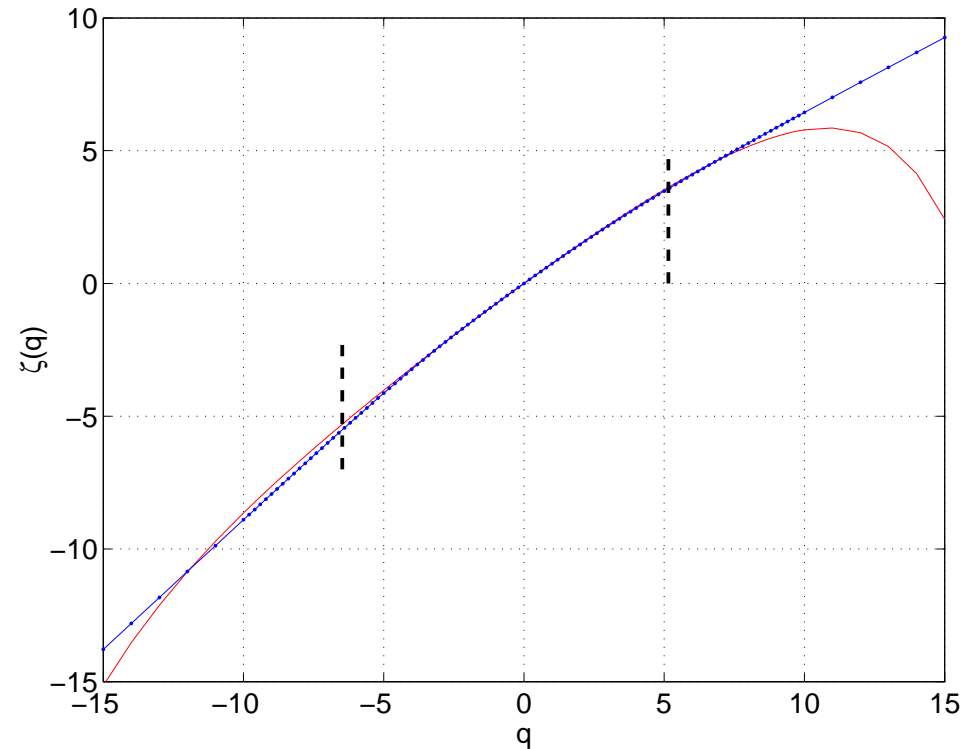
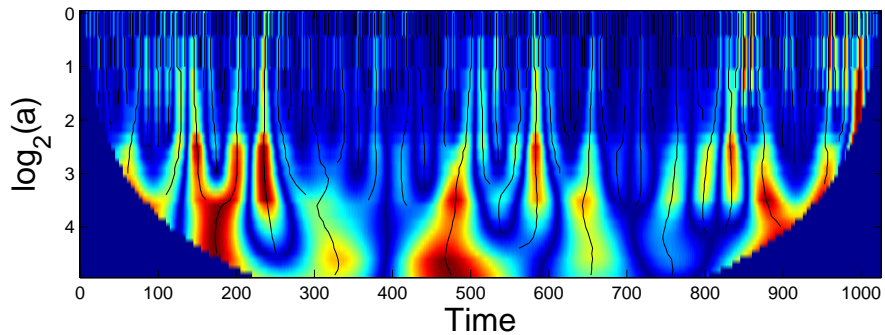
WT MODULUS MAXIMA (ARNEODO ET AL.)

$$L_X(a, t_k) = \text{SUP}_{a' < a} |T_X(a', t_k(a'))|$$

WTMM



Skeleton of the Wavelet Transform

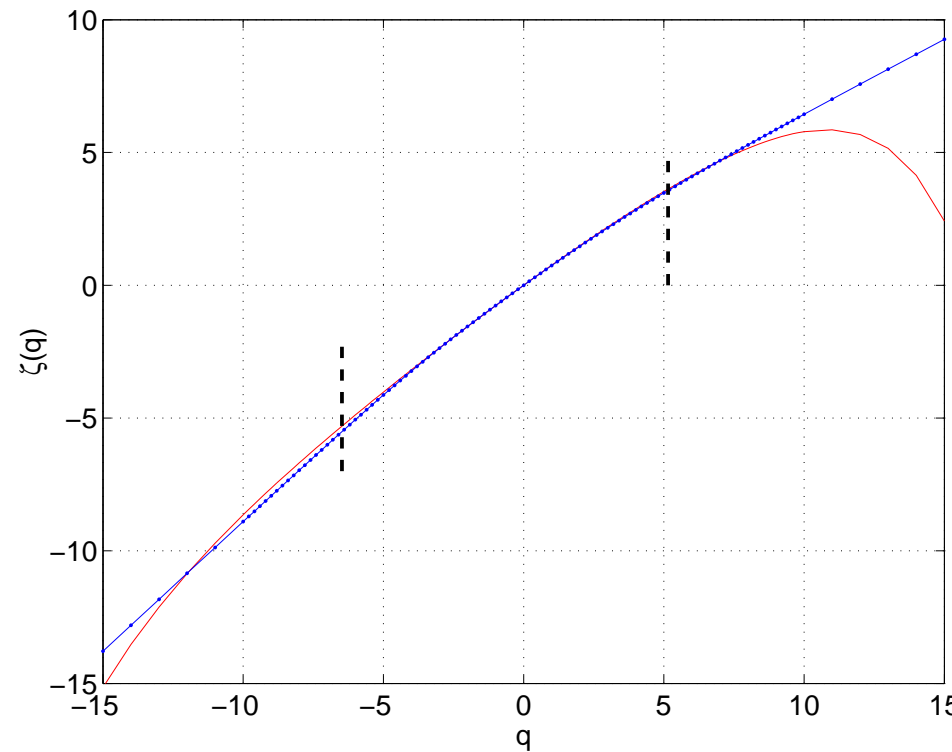
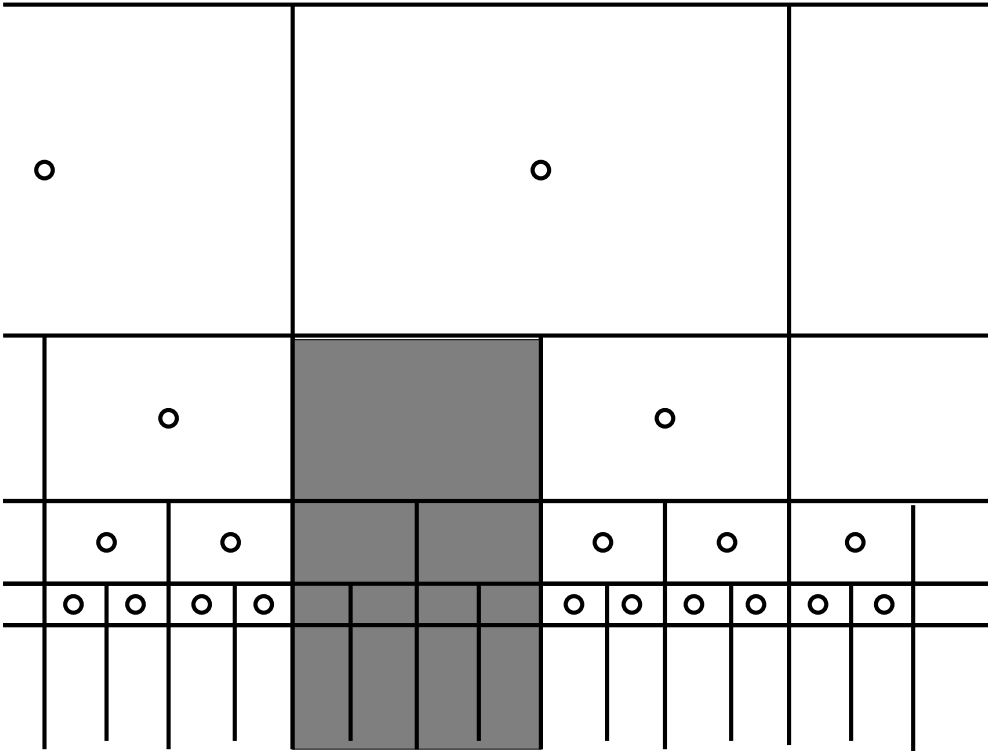


COMPUTATIONALLY EXPENSIVE

NEGATIVE VALUES OF q S - SOLUTION 3

WAVELET LEADERS: (JAFFARD ET AL.)

$$d_X(j, k) \rightarrow L_X(j, k) = \sup_{j' < j} d_X(j', 2^{-j'})$$



COMPUTATIONALLY EFFICIENT AND EXCELLENT STATISTICAL PERFORMANCE

BEYOND POWER LAWS

- SELF-SIMILARITY:

$$\mathbb{E}|d_X(j, k)|^q = C_q(2^j)^{qH} = C_q \exp(qH \ln 2^j)$$

- POWER LAWS,
- $\forall 2^j$ (FOR ALL SCALES),
- $\forall q / \mathbb{E}|d_X(j, k)|^q < \infty$,
- A SINGLE PARAMETER H
- ADDITIVE STRUCTURE.

- MULTIFRACTAL

$$\mathbb{E}|d_X(j, k)|^q = C_q(2^j)^{\zeta(q)} = C_q \exp(\zeta(q) \ln 2^j)$$

- POWER LAWS,
- $\forall 2^j < L$, (FOR FINE SCALES ONLY, IN THE LIMIT $2^j \rightarrow 0$),
- $\forall q?$
- A WHOLE COLLECTION OF SCALING PARAMETER $\zeta(q)$
- MULTIPLICATIVE STRUCTURE.

- WARPED INF. DIV. CASCADES :

$$\mathbb{E}|d_X(j, k)|^q = C_q(2^j)^{\zeta(q)} = C_q \exp(\zeta(q)n(2^j))$$

\Rightarrow VISIT PIERRE CHAINAIS'S POSTER.

CONCLUSIONS AND REFERENCES

ANALYSING SCALING IN DATA ?

- THINK WAVELET
 - EFFICIENCY,
 - PRACTICAL AND CONCEPTUAL ADEQUATION AND SIMPLICITY,
 - ROBUSTNESS AGAINST NON STATIONARITIES,
 - EASY TO USE, LOW COAST, REAL TIME ON LINE.

MODELLING SCALING IN DATA ?

- THINK SELF SIMILARITY VERSUS MULTIPLICATIVE CASCADES,
- AND POSSIBLY ADD LONG MEMORY.
- ALSO SCALING MAY NOT BE POWER LAWS

REFERENCES AND RESOURCES, VISIT :

- perso.ens-lyon.fr/patrice.abry
- www.cubinlab.ee.mu.oz.au/~darryl
- fraclab
- www.isima.fr/~chainais

WANT TO TALK TO US ?

**Patrice.Abry@ens-lyon.fr
perso.ens-lyon.fr/patrice.abry**

**Paulo.Goncalves@inria.fr
www.inrialpes.fr/is2**