IMPACT OF DATA QUANTIZATION ON EMPIRICAL MULTIFRACTAL ANALYSIS

Herwig Wendt, Stéphane G. Roux, Patrice Abry

Physics Lab., CNRS UMR 5672, Ecole Normale Supérieure de Lyon, France.
Herwig.Wendt@ens-lyon.fr, Stephane.Roux@ens-lyon.fr, Patrice.Abry@ens-lyon.fr

SUMMARY

PROBLEM **Multifractal Analysis**: Commonly used in real-life data analyses, involved in standard signal processing tasks such as detection, identification or classification. In a number of situations, in particular in Image Processing, the **data** available for the analyses **severely quantized**.

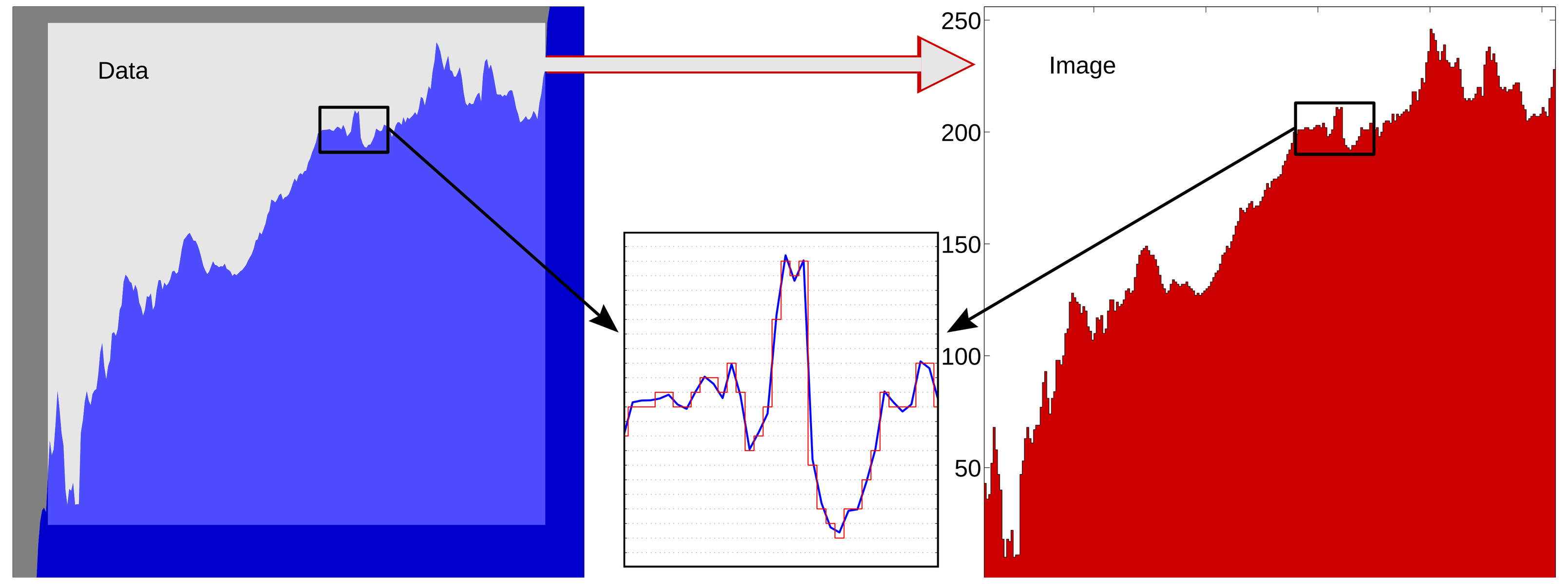
GOAL Analyzing **robustness of standard multifractal estimation procedures against quantization**.

RESULTS
- **Immunity against quantization**: **restricting range of scales** involved in multifractal parameter estimation to the **largest ones**.
- **Wavelets** bring **robustness against quantization** when **increments do not**.

EXAMPLE 1: IMAGE

- High acquisition rate for video capture $\Longrightarrow$ low image resolution
- Fracture: Analysis of roughness of boundaries

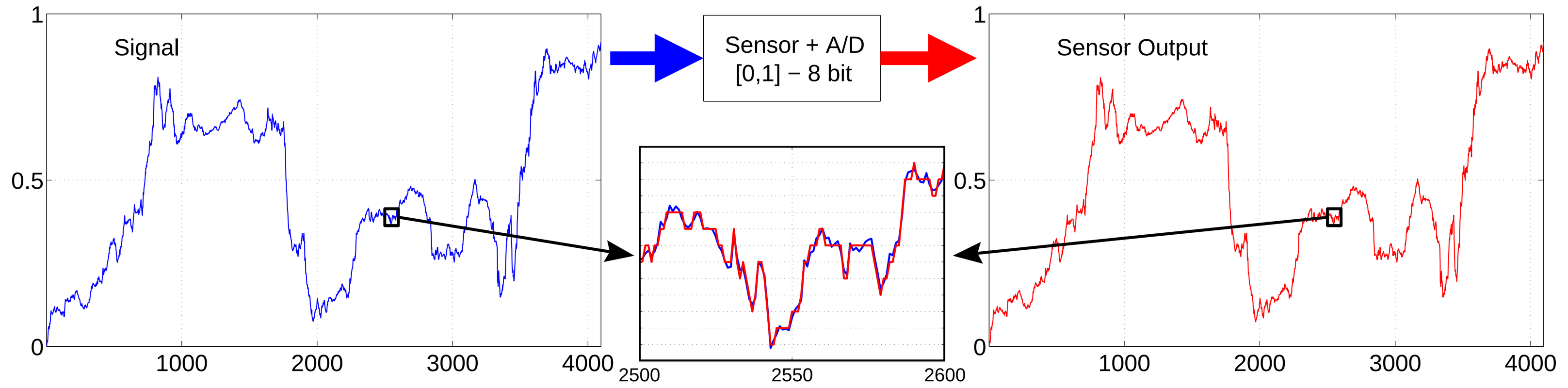
Image 256×256 $\Longrightarrow$ Quantization (8 bit) in amplitude



EXAMPLE 2: SIGNAL

Sampling in Amplitude

Analyzed signal $\longrightarrow$ quantized



REFERENCES

P. Flandrin, "Wavelet analysis and synthesis of fractional Brownian motions," *IEEE Trans. on Info. Theory*, vol. 38, pp. 910–917, 1992.

S. J. Mitchell, "Discontinuities in self-affine functions lead to multiaffinity," *Phys. Rev. E*, vol. 72, pp. 4065103(R), 2005.

S. Santucci et al., "Gaussian statistics of fracture surfaces," *http://lanl.arxiv.org/cond-mat/pdf/0607/0607372.pdf*, 2006.

N. Mallick, P. P. Cortet, S. Santucci, S. G. Roux, and L. Vanel, "Discrepancy between subcritical and fast rupture roughness: A cumulant analysis," *preprint*, 2006.

A.M. Zoubir and D.R. Iskander, *Bootstrap Techniques for Signal Processing*, Cambridge University Press, 2004.

E. Bacry, J. Delour, and J.F. Muzy, "Multifractal random walk," *Phys. Rev. E*, vol. 64, pp. 026103, 2001.

D. Veitch, M. Taqqu, and P. Abry, "Meaningful MRA initialisation for discrete time series," *Signal Processing, Elsevier Science*, vol. 8, pp. 1971–1983, 2001.

J. Delour, J.F. Muzy, and A. Arneodo, "Intermittency of 1d velocity spatial profiles in turbulence: A magnitude cumulant analysis," *The Euro. Phys. Jour. B*, vol. 23, pp. 243–248, 2001.

S. Jaffard, "Wavelet techniques in multifractal analysis," in *Fractal Geometry and Applications: A Jubilee of Benoît Mandelbrot, M. Lapidus and M. van Frankenhuijsen Eds., Proceedings of Symposia in Pure Mathematics*. 2004, vol. 72(2), pp. 91–152, AMS.

P. Abry, P. Flandrin, M. Taqqu, and D. Veitch, "Wavelets for the analysis, estimation and synthesis of scaling data," in *Self-similar Network Traffic and Performance Evaluation,* K. Park and W. Willinger, Eds. spring 2000, Wiley.

MULTIFRACTAL ANALYSIS

MULTIFRACTAL ANALYSIS AND MULTIRESOLUTION QUANTITIES

Scale Invariance

$$\boxed{\frac{1}{n_a}\sum_{k=1}^{n_a} |T_X(a,ka)|^q \simeq c_q |a|^{\zeta(q)}}$$

for statistical orders $q \in [q_*^-, q_*^+]$
for scales $a = 2^j \in [a_m, a_M]$, $a_M/a_m >> 1$

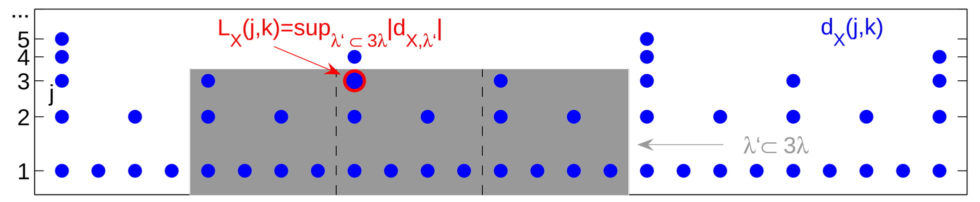
$X(t)$, $t \in [0,n)$ - Process under analysis
$T_X(a,ka)$ - Multiresultion quantities of X jointly depend on analysis scale a and time position t
$\zeta(q)$ - Scaling exponents.

Multiresolution Quantities

Increment of order 1
$$T_X^{(I_1)}(2^j,t) = X(t+2^j\tau_0) - X(t)$$
Increment of order 2
$$T_X^{(I_2)}(2^j,t) = X(t+2\cdot 2^j\tau_0) - 2X(t+2^j\tau_0) + X(t)$$
Wavelet Coefficients
$$T_X^{(W)}(2^j,t) = d_X(j,k) = \langle \psi_{j,k} | X \rangle$$
Wavelet Leaders
$$T_X^{(L)}(2^j,t) = L_X(j,k) = \sup_{\lambda' \in 3\lambda_{j,k}} |d_{\lambda'}|$$

Empirical Multifractal Analysis

- $\zeta(q)$ closely tied to theoretical multifractal spectrum $D(h)$ of X
- Goal: Obtain $\zeta(q)$ from single realization of X



Definition of Wavelet Leaders

LOG-CUMULANTS AND ESTIMATION

Log-Cumulants

$\boxed{\zeta(q) = \sum_{p=1}^{\infty} c_p \frac{q^p}{p!}}$ polynomial expansion

- $\forall p \geq 1$: $C_p^j = c_p^0 + c_p \ln 2^j$
  C_p^j: p-th cumulant of $\ln |T_X(2^j,\cdot)|$
- Measurements of $\zeta(q)$ (and $D(h)$):
  $\longrightarrow$ measurements of **log-cumulants** c_p

Estimation Procedures

1. Scale j: calculate n_j coefficients $T_X(2^j, k\cdot 2^j)$
2. Discard $|T_X|$ below threshold (10^{-10})
3. Calculate $\hat{C}_p^j$: standard estimate of p-th cumulant of $\ln |T_X(2^j,\cdot)|$
4. Linear regressions: $\hat{c}_p = \log_2 e \sum_{j=j_1}^{j_2} w_j \hat{C}_p^j$

QUANTIZATION AND NUMERICAL STUDY

PROCESSES

Multifractal Random Walk (MRW)

$X(k) = \sum_{k=1}^{n} G_H(k) e^{\omega(k)}$
- Non Gaussian processes with stationary increments
- Multifractal properties mimic Mandelbrot's multiplicative LN cascades
- $\zeta(q) = (H+\lambda^2)q - \lambda^2 q^2/2$ for $q \in \left[-\sqrt{2/\lambda^2}, \sqrt{2/\lambda^2}\right]$
  $c_1 = (H+\lambda^2)$, $\quad c_2 = -\lambda^2$

$G_H(k)$: increments of FBM with parameter H.
ω : independent of G_H, Gaussian, with non trivial covariance:
$\mathrm{cov}(\omega(k_1),\omega(k_2)) = \lambda^2 \ln\left(\frac{L}{|k_1-k_2|+1}\right)$ for $|k_1-k_2| < L$
FBM: Fractional Brownian Motion: Only Gaussian self-similar process with stationary increments
$X(t) \overset{\text{fdd}}{=} a^H X(t/a)$ for all $a > 0$

QUANTIZATION AND MONTE CARLO SIMULATION

Quantization

$$\boxed{X^{\Delta}(n) = [X(n)/\Delta]\cdot\Delta, \quad b = -\log_2 \Delta}$$

$X(n)$, $n = 1,\cdots,N$ - Realization of process X
$[\cdot]$ - Rounding operation
Δ - Quantization interval width
b - quantization level (in bits)

$\longrightarrow$ unit interval $[0,1]$ has 2^b quantization levels

Performance Assessment and Simulation Setup

Apply estimation procedure to large number N_{MC} of realizations of MRW at different quantization levels

$$\mathrm{MSE} = \sqrt{\left(\widehat{\mathbb{E}}\hat{c}_p - c_p\right)^2 + \widehat{\mathrm{Var}}\hat{c}_p}$$

$\widehat{\mathbb{E}}$, $\widehat{\mathrm{Var}}$ - sample mean/variance over N_{MC} realizations

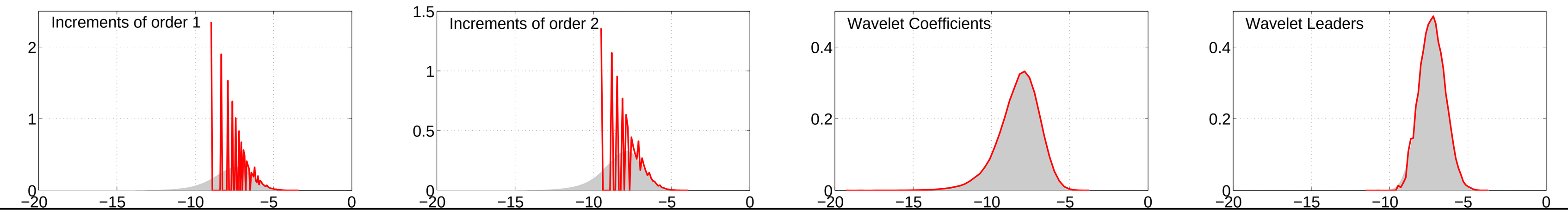
Wavelets: Daubechies2 MRW: $(H,\lambda) = (0.72, \sqrt{0.08})$ $(c_1,c_2) = (0.8, -0.08)$	$N_{MC} = 1000$ $n = 2^{14}$

QUANTIZATION IMPACTS

DISTRIBUTIONS OF $\ln |T_X(2^j,\cdot)|$ ($b = 12$ bits quantization $j = 3$)

Increments:
- Quantized data: Distributions of $\ln|T_X^{(I_1)}(2^j,\cdot)|$ and $\ln|T_X^{(I_2)}(2^j,\cdot)|$: lattice, severely different from non-quantized data.

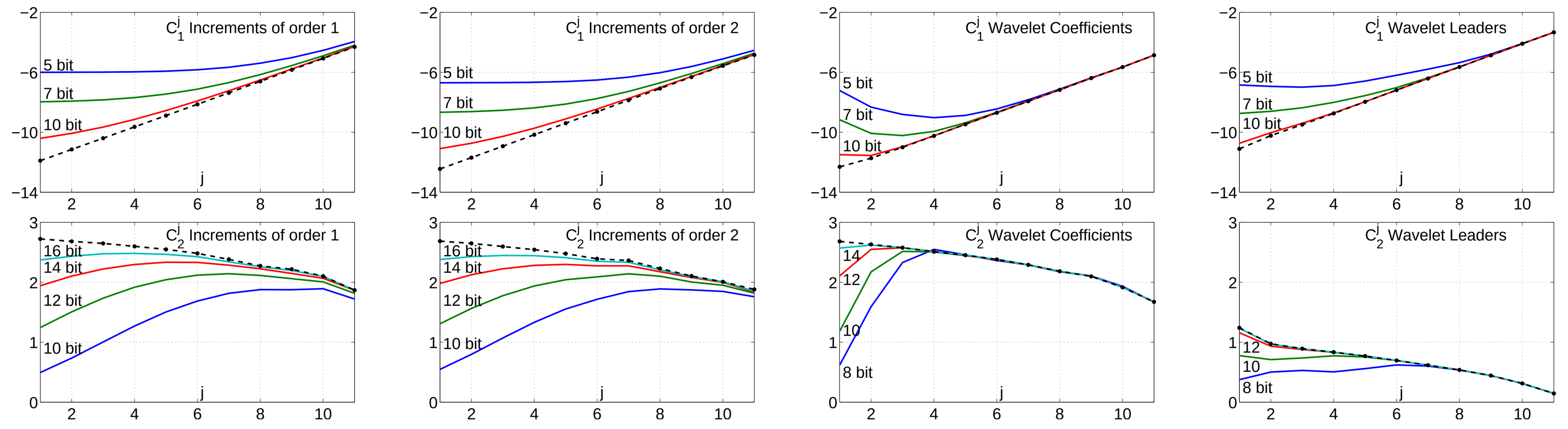
Coefficients and Leaders:
- Quantization has no visible effect on distributions of $\ln|T_X^{(W)}(2^j,\cdot)|$ and $\ln|T_X^{(L)}(2^j,\cdot)|$.



$\hat{C}_p^j = c_p^0 + c_p \ln 2^j$

Increments:
- Strong impact at coarse scales as b decreases.
- No meaningful regression for $b \leq 7$ ($\hat{c}_1$) and $b \leq 12$ ($\hat{c}_2$)

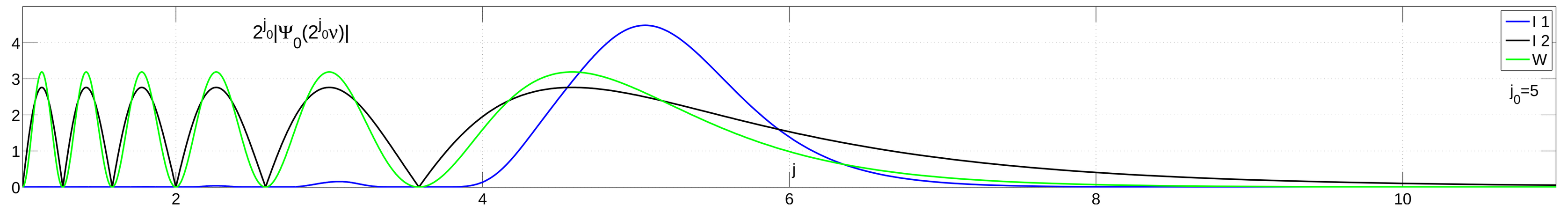
Coefficients and Leaders:
- Impact at coarse scales less dramatic as b decreases.
- Meaningful regression for $b \geq 5$ ($\hat{c}_1$) and $b \geq 8$ ($\hat{c}_2$)



$\Rightarrow$ Impact on $\hat{C}_p^j$ from fine to coarse scales as Δ increases. $\Rightarrow$ Much more dramatic for Increments than for Coefficients and Leaders
$\Rightarrow$ Frequency resolution of the mother wavelets. $\Rightarrow$ Narrow regression range

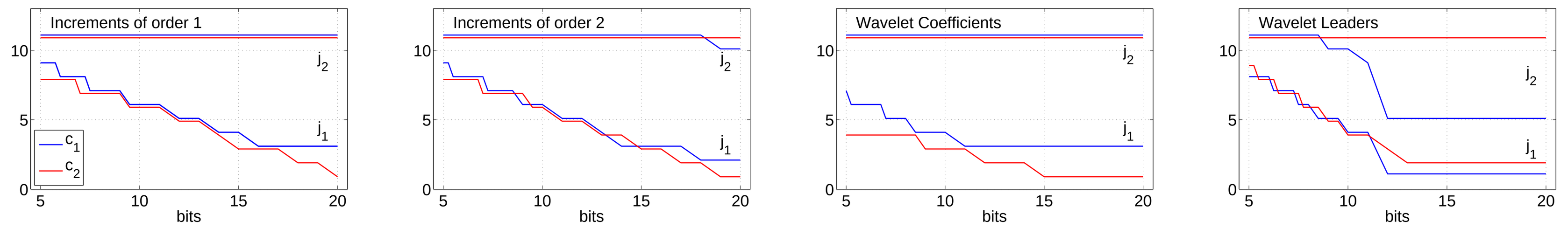
FREQUENCY RESOLUTION

Fourier transforms of mother wavelets. Increments: poor frequency resolution, Wavelets: high frequency resolution



OPTIMAL REGRESSION RANGE

Decreasing b forces j_1 to increase $\Longrightarrow$ Confines regression range to coarser and coarser scales



STATISTICAL PERFORMANCE

Non Quantized Signal: Increments and Coefficients comparable, Leaders significantly better

Quantized Signal - Fixed regression range ($j_1 = 5$, $j_2 = 11$):
- Increments: degrade dramatically fast as $b < 14$
- Coefficients, Leaders: performant still at $b = 8$

Optimal regression range: Estimation improved
- Increments: improvement small, confined to $b > 14$
- Coefficients, Leaders: relevant estimates still at $b = 5$