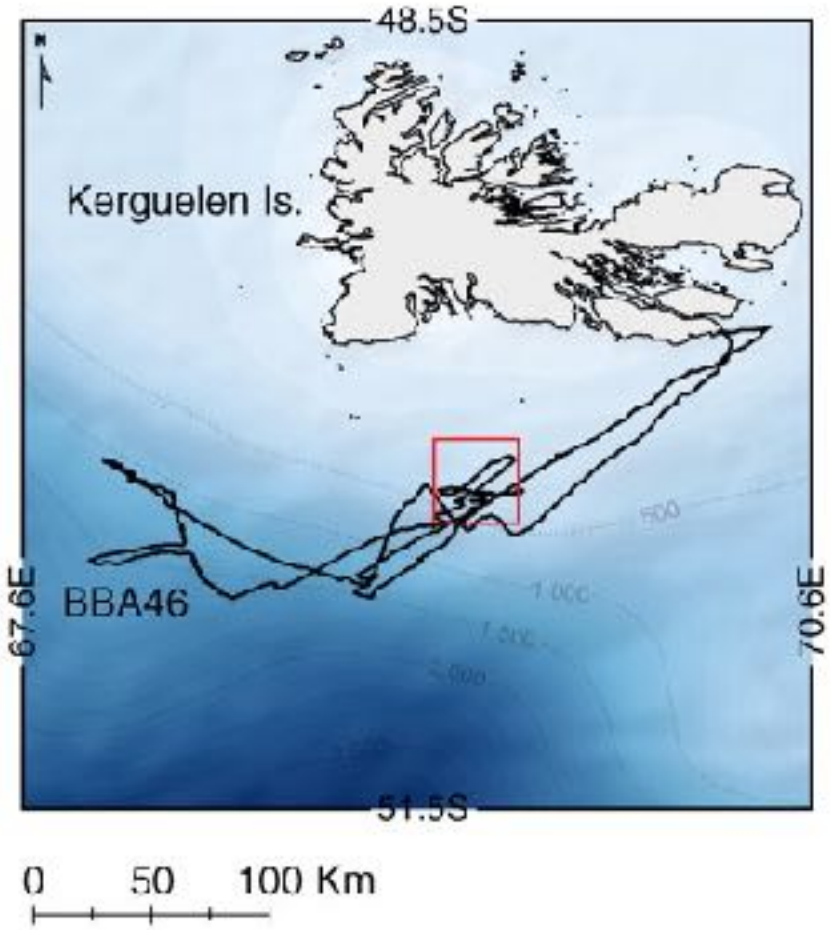
Data analysis techniques for plasmas

Thierry Dudok de Wit
LPC2E, University of Orléans, France

with thanks to many people: Marie Farge, Kai Schneider, Khurom Kiyani, and more

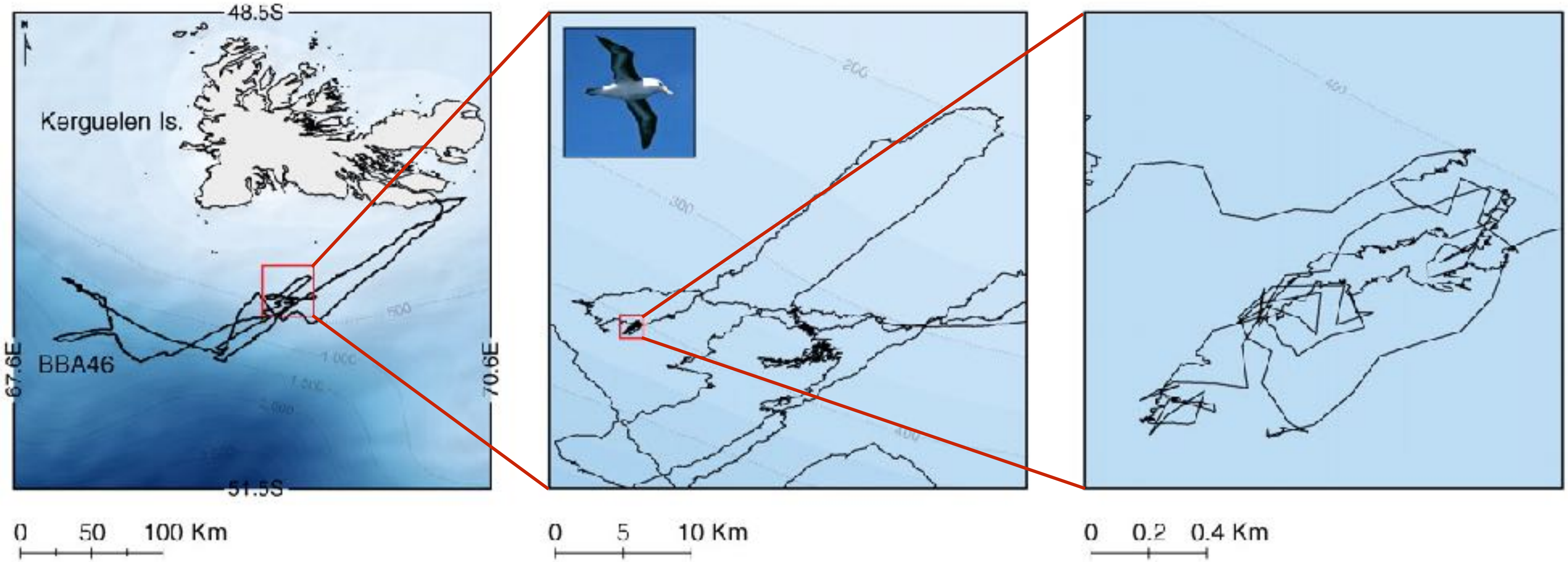
Les Houches, 2-12 May 2017

Foraging albatrosses



Humphries et al. (PNAS, 2012)

Foraging albatrosses



Humphries et al. (PNAS, 2012)

Conclusion

Conclusion 1

- The best strategy for foraging consists of mixing Brownian motion (random walk) and Lévy flights (long stretches)

- The probability distribution $p(\text{distance travelled})$ has long tails = highly non Gaussian

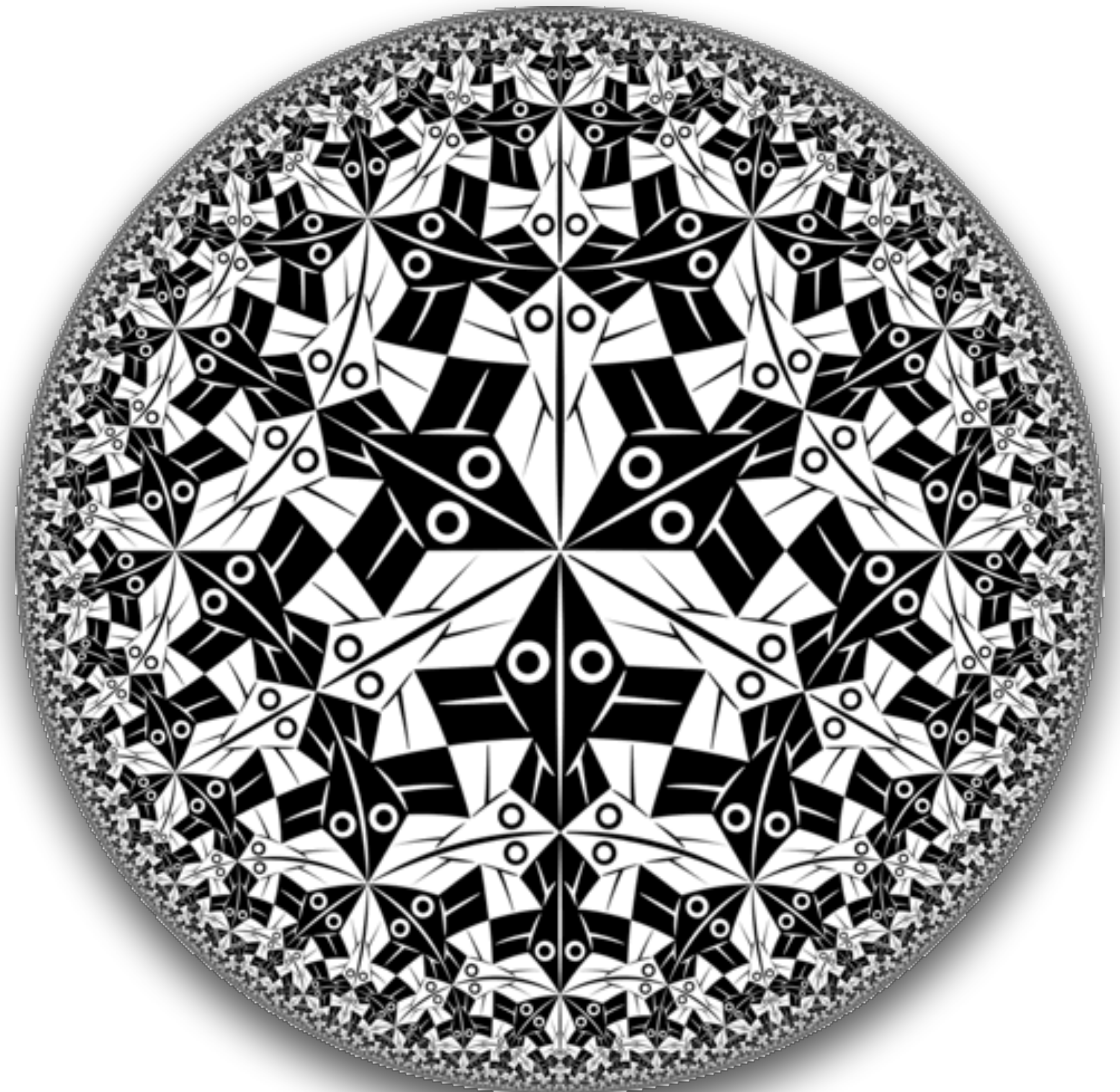
- Strong analogy with charge motion in magnetised plasmas

Such data cannot be analysed with the classical tools we learn in our physics courses!

Conclusion 2

Plasmas tend to be highly

- non-stationary **focus of this lecture**
  - most analysis techniques assume stationarity
- non-Gaussian
  - most analysis techniques assume Gaussianity
- non-linear
  - most analysis techniques assume linearity
- etc.



M. C. Escher

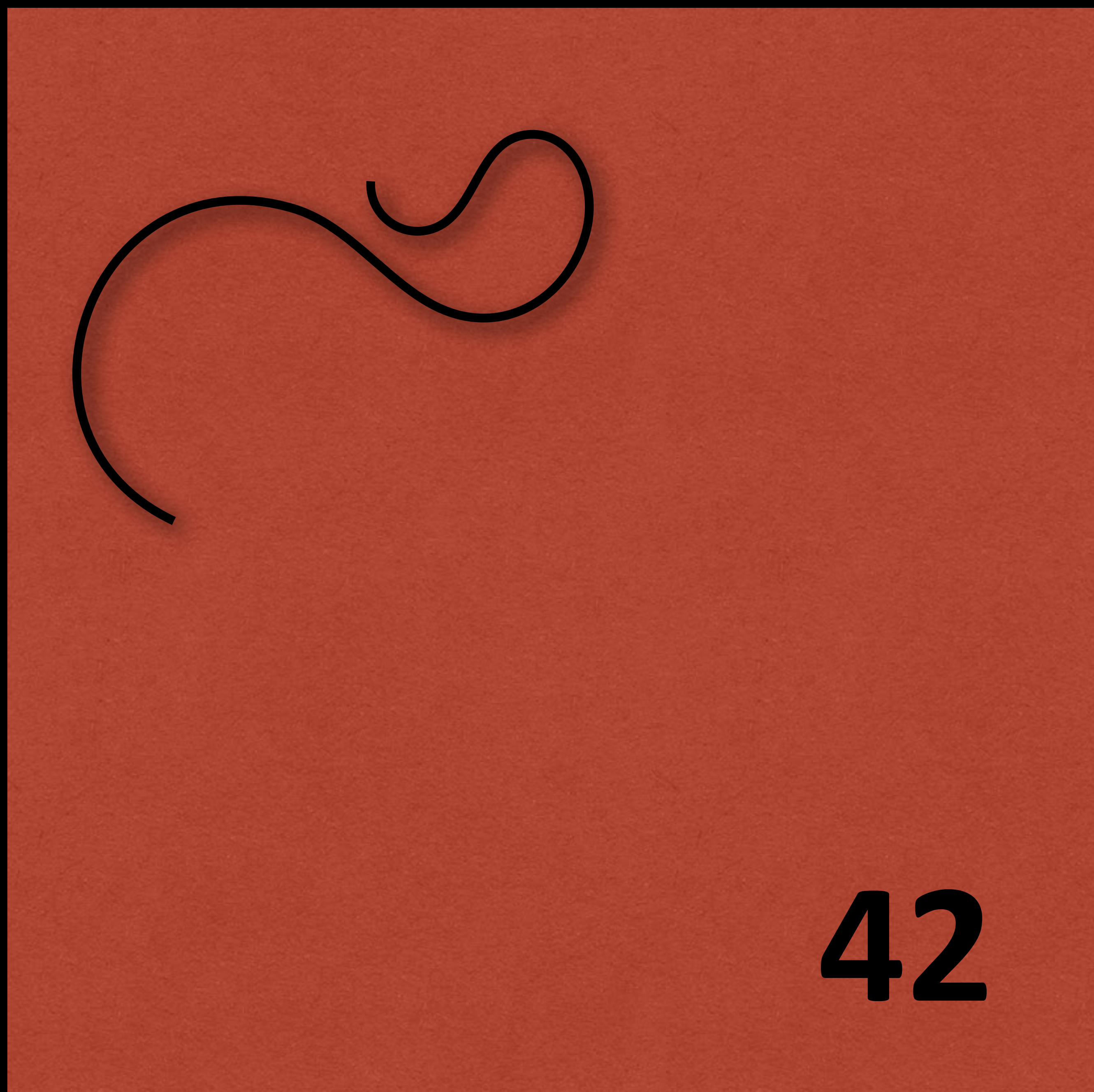
Plasmas through the multiscale (wavelet) lens

Outline

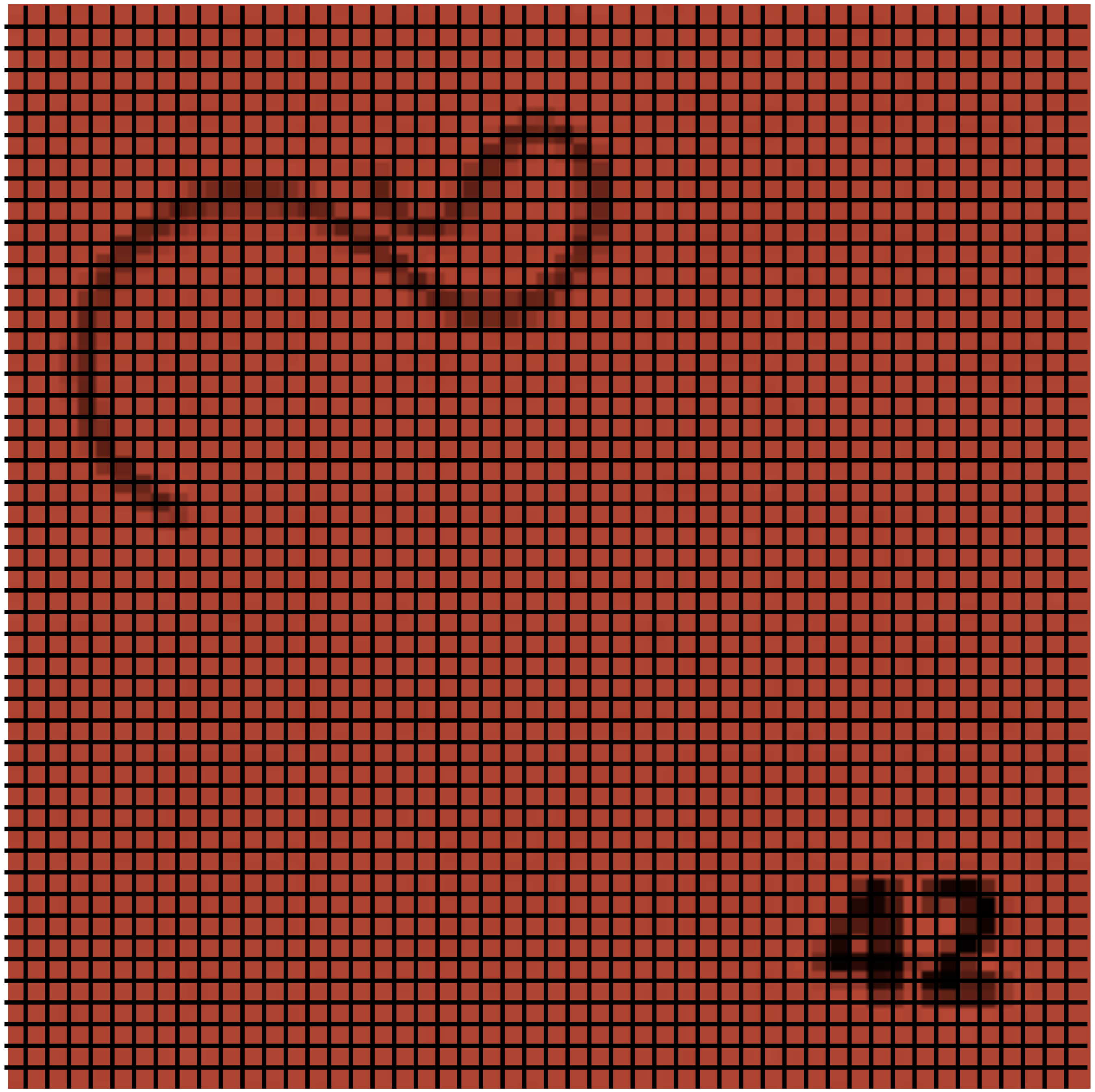
- A wavelet primes
- Some examples from plasmas
- Coherent structures
- Sparsity
- Blind source separation

An exercise

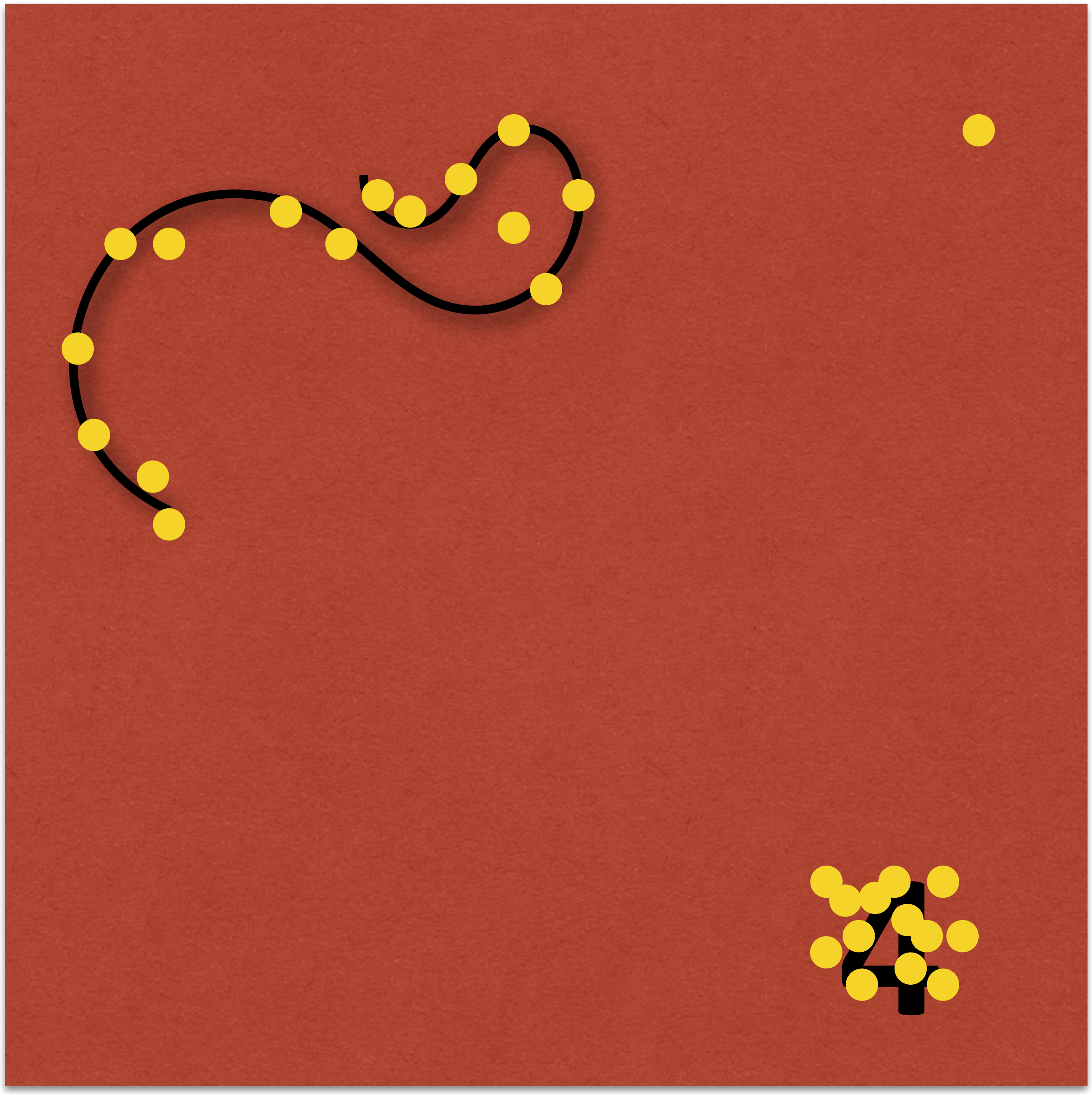
What do you see ?



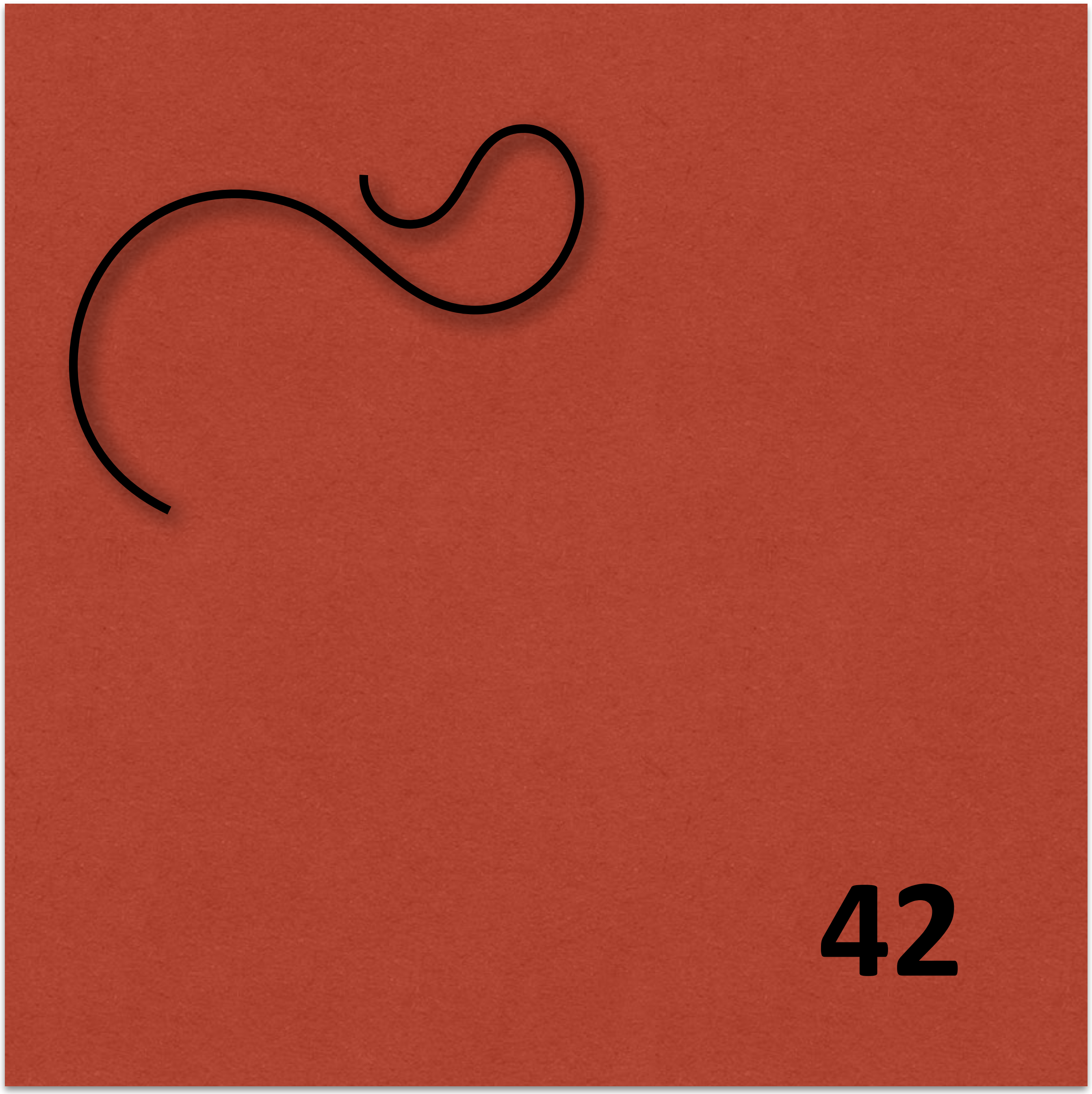
How a CCD camera sees this



How our eye sees this



A more symbolic representation



= red square + lace + Deep thought

A multiscale view

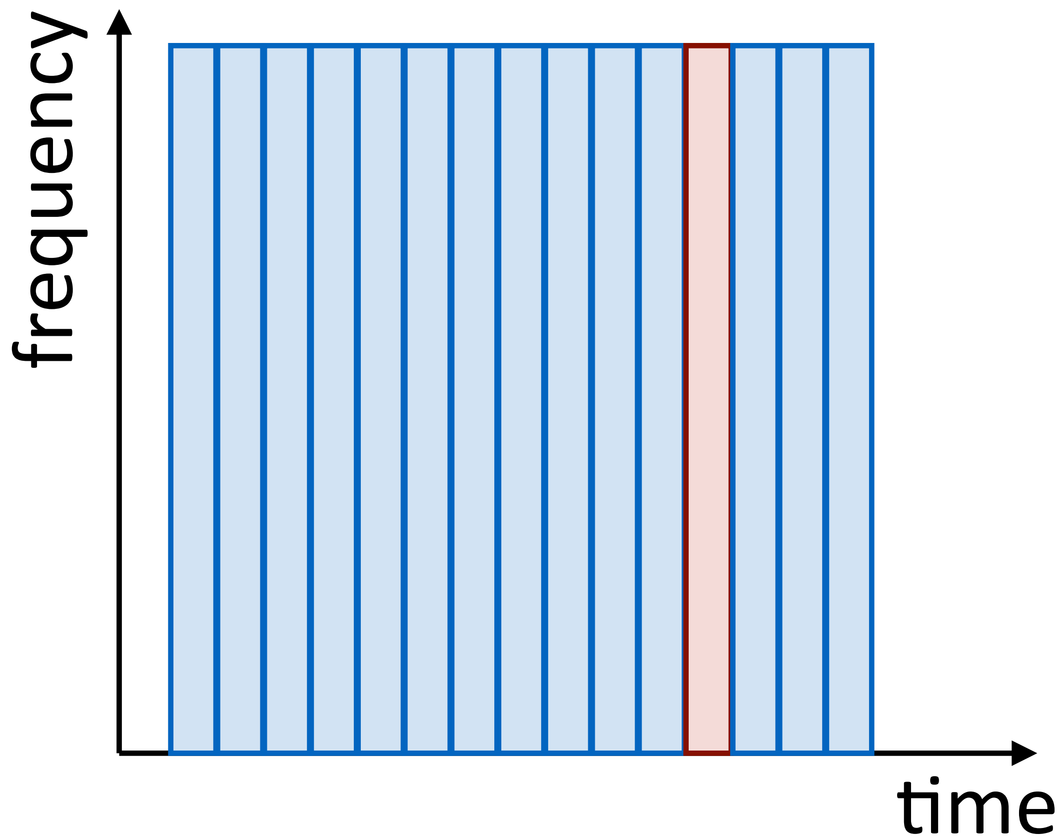
- Our eye extracts **multiple scales** and only keep the most relevant information.

- What is the optimal representation of plasma data ?
  - optimal = to get deeper **insight** (find invariants) ?
  - optimal = find highest **compression** (jpeg) ?
  - optimal = find the simplest code that can generate these data (Kolmogorov **complexity**) ?
  - ...

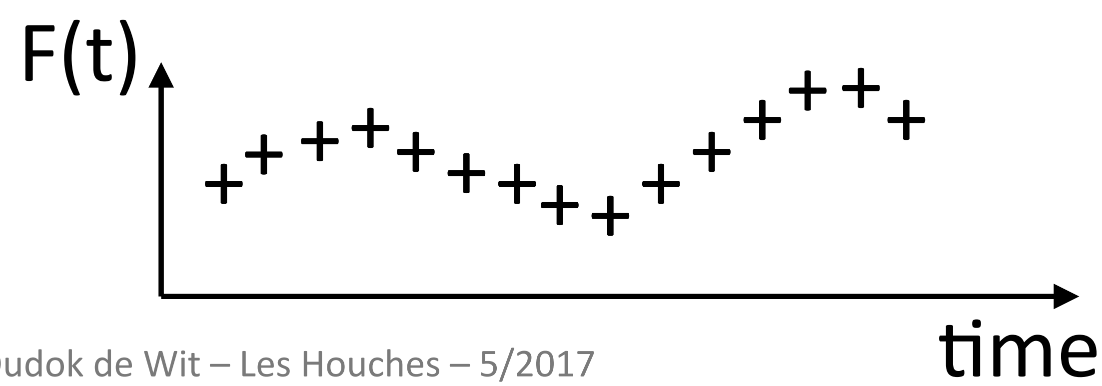
A wavelet primer

Optimal tiling

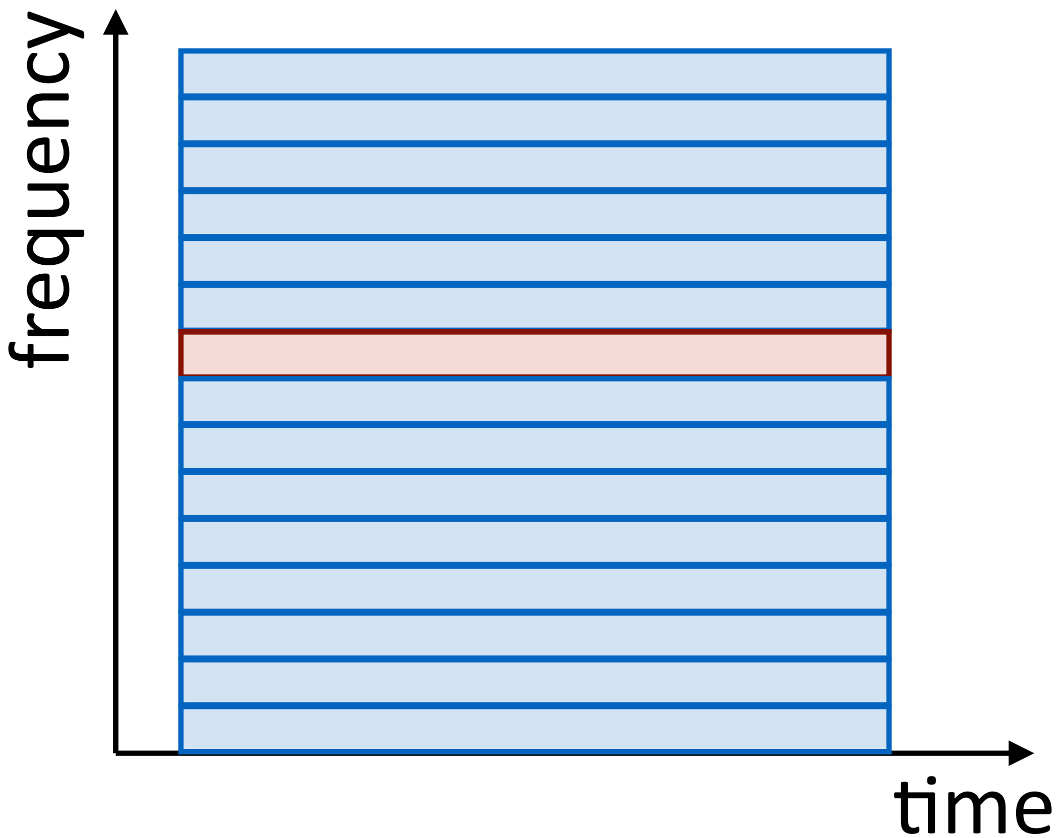
Time series



- Excellent time localisation
- No spectral localisation



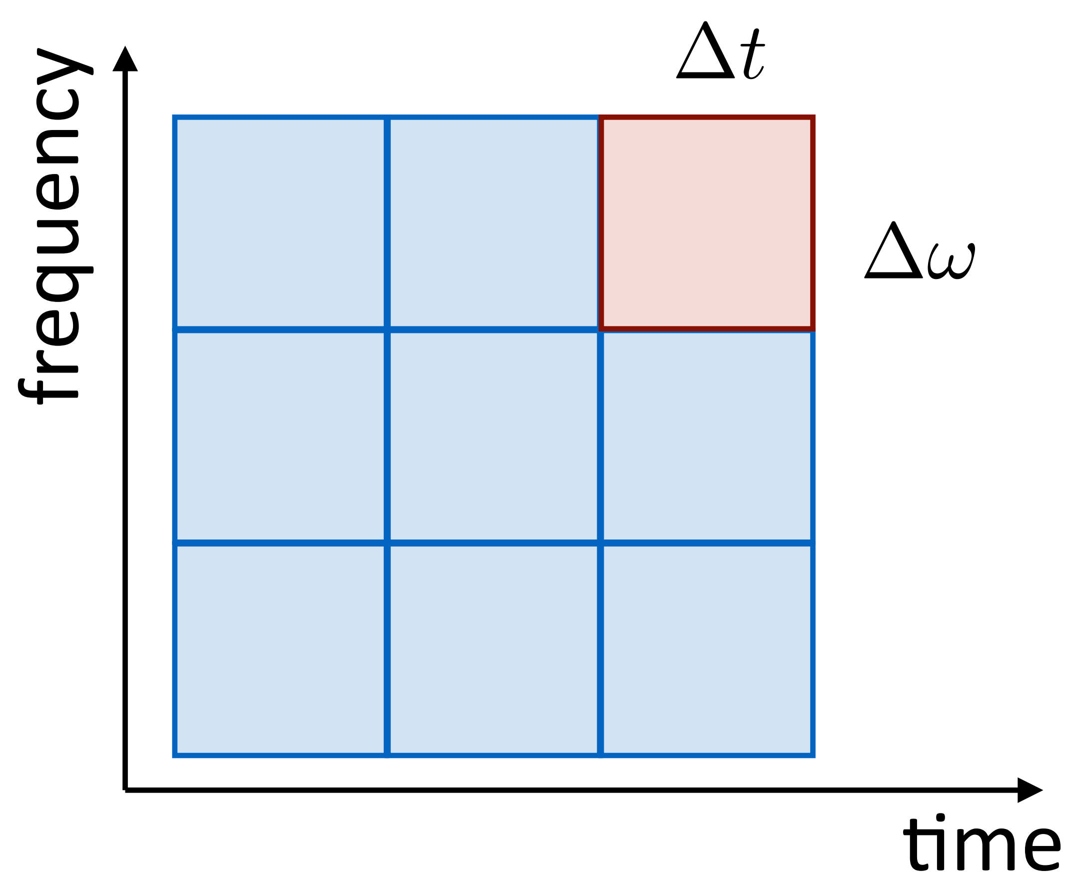
Fourier transform



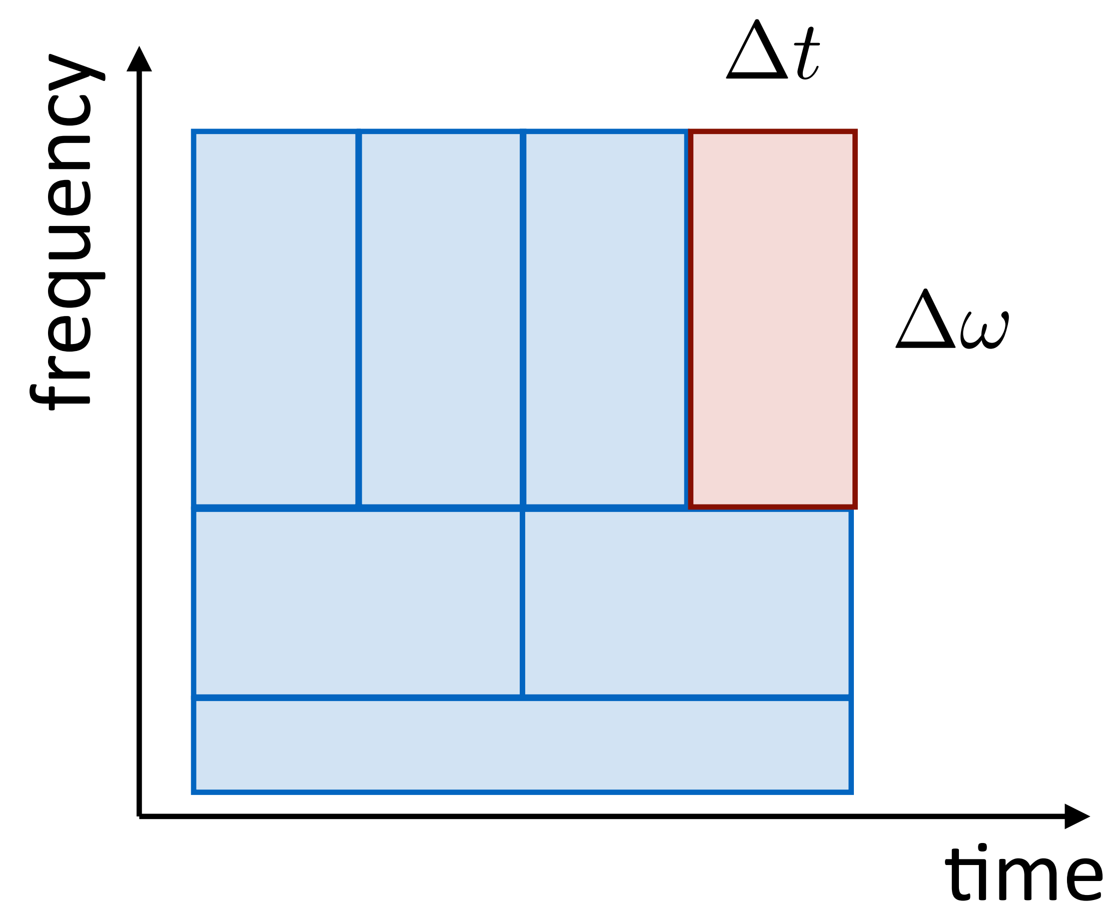
- Excellent spectral localisation
- No time localisation

Optimal tiling

Gabor atoms (1946)



Wavelets (1984)



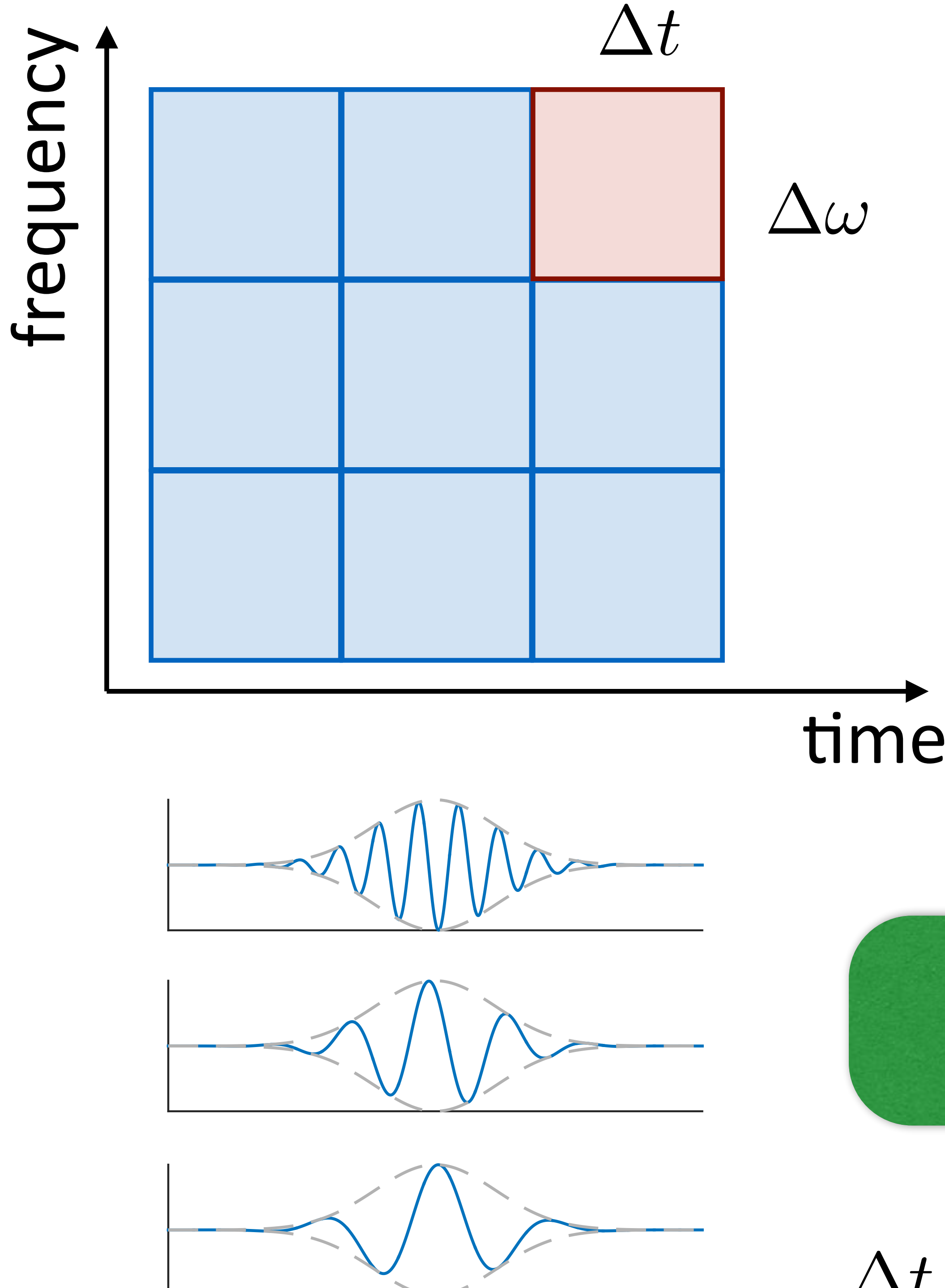
- Fixed time localisation
- Fixed spectral localisation

- Variable time localisation
- Fixed relative spectral localisat.

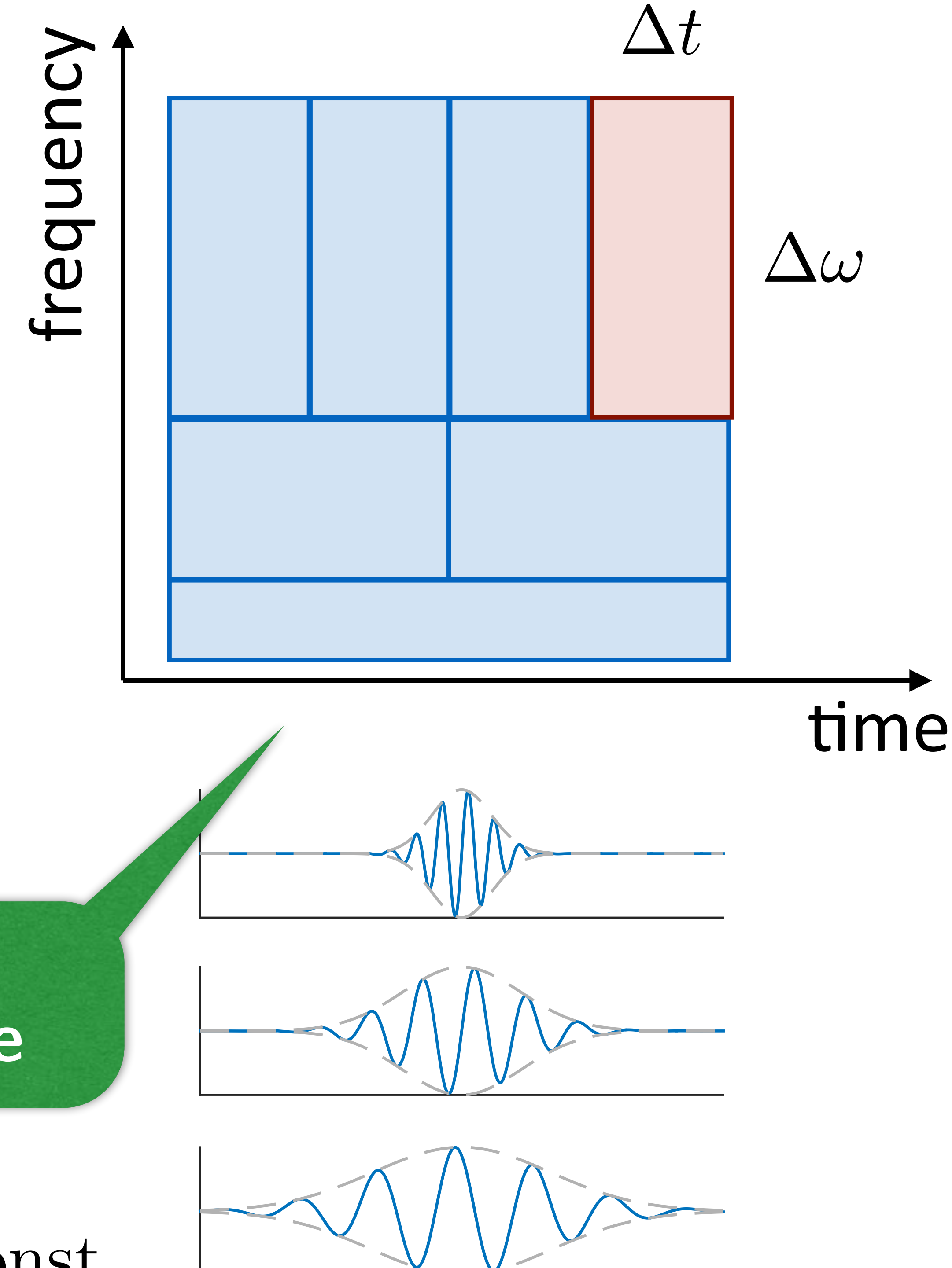
$$\Delta t \cdot \Delta\omega \geq \text{const.}$$

Optimal tiling

Gabor atoms (1946)



Wavelets (1984)



far preferable

$$\Delta t \cdot \Delta\omega \geq \text{const.}$$

Wavelets are born

- Morlet and Grossman (1984) : project *f(t)* on a set of wavelet functions

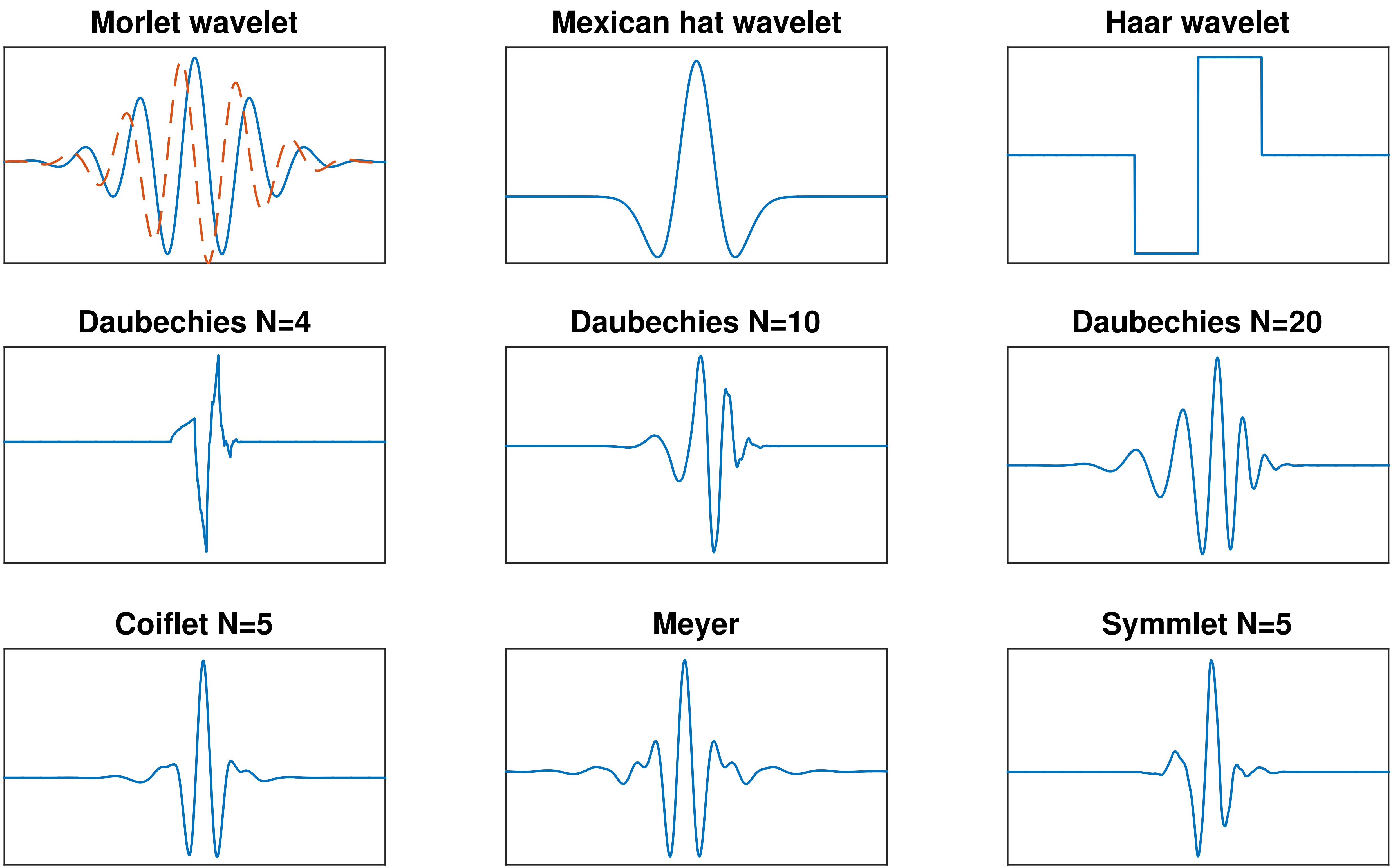
$$\tilde{f}(a,b) = \int_{-\infty}^{+\infty} f(t)\, \psi^*_{a,b}(t)\, \mathrm{d}t$$

with

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right)$$

mother wavelet

Some common wavelets



What is a wavelet ?

- Almost anything can be a wavelet

1) admissibility condition

$$\int_{-\infty}^{+\infty} \psi(t)\,\mathrm{d}t = 0$$

a mother wavelet must be wiggling

2) integrability

$$\int_{-\infty}^{+\infty} |\psi(t)|\,\mathrm{d}t < \infty \qquad \int_{-\infty}^{+\infty} |\psi(t)|^2\,\mathrm{d}t < \infty$$

a mother wavelet must not diverge

Moments

- Wavelets are classified by their order N
  = number of vanishing moments

$$\int_{-\infty}^{+\infty} t^N \, \psi(t) \, \mathrm{d}t = 0$$

Wavelets will be insensitive to polynomial trends of order N (→ non-stationarity)

Wavelets are born

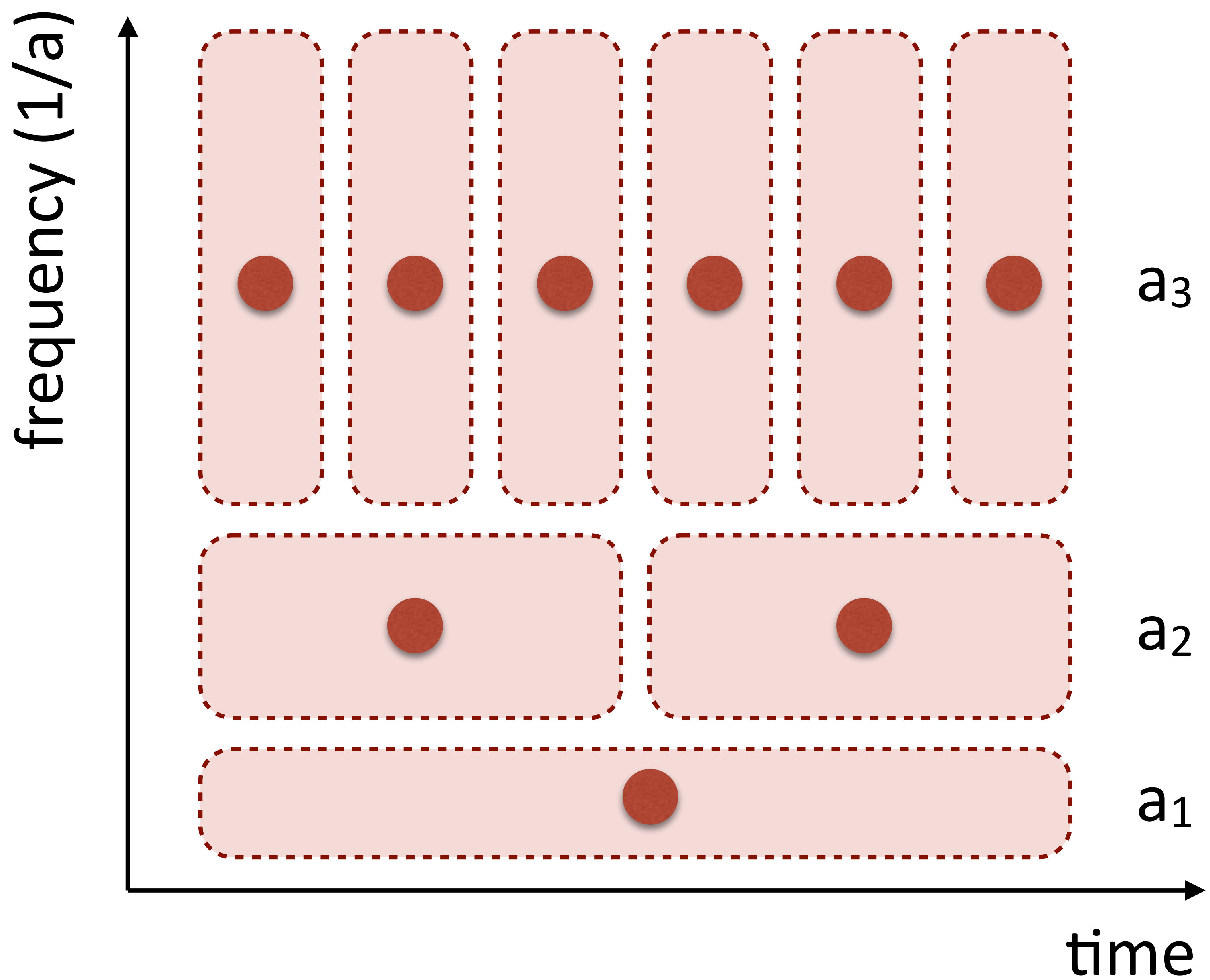
$$\tilde{f}(a,b) = \int_{-\infty}^{+\infty} f(t)\, \psi^*_{a,b}(t)\, \mathrm{d}t$$

$$f(t) \xrightarrow{\text{analysis}} \tilde{f}(a,b)$$

$$f(t) \xleftarrow{\text{synthesis}} \tilde{f}(a,b)$$

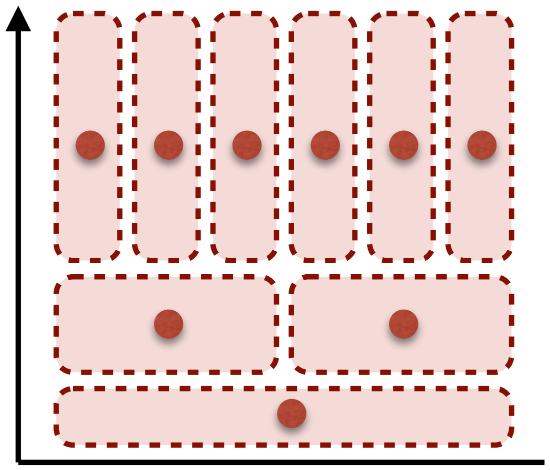
$$f(t) = \frac{1}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} \tilde{f}(a,b)\, \psi_{a,b}(t)\, \frac{\mathrm{d}a\, \mathrm{d}b}{a^2}$$

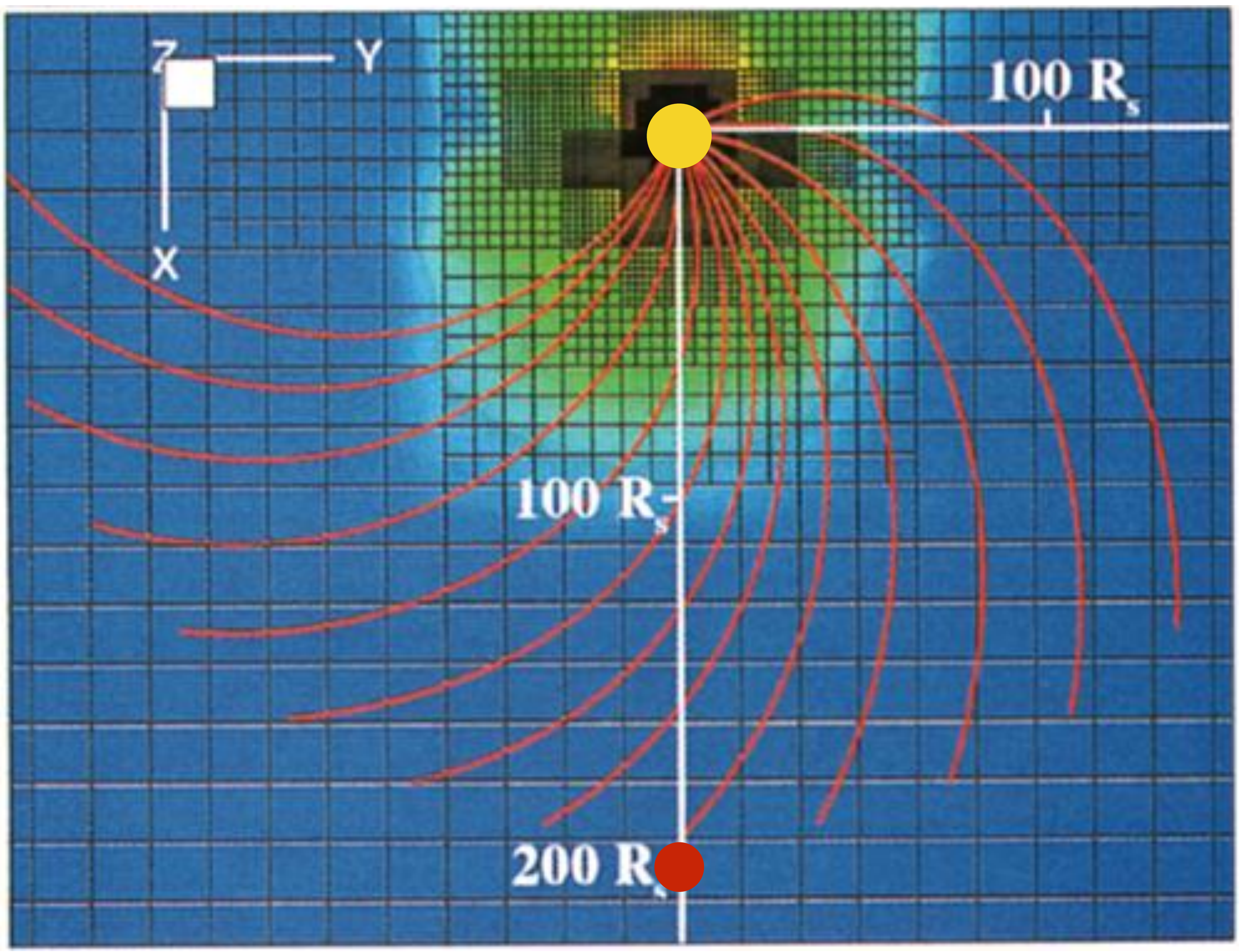
From continuous to discrete wavelets



Idea : choose optimal grid to avoid redundant information = find orthogonal "cells"

Similarities





Adaptive grid (U. Michigan)

Discrete wavelets

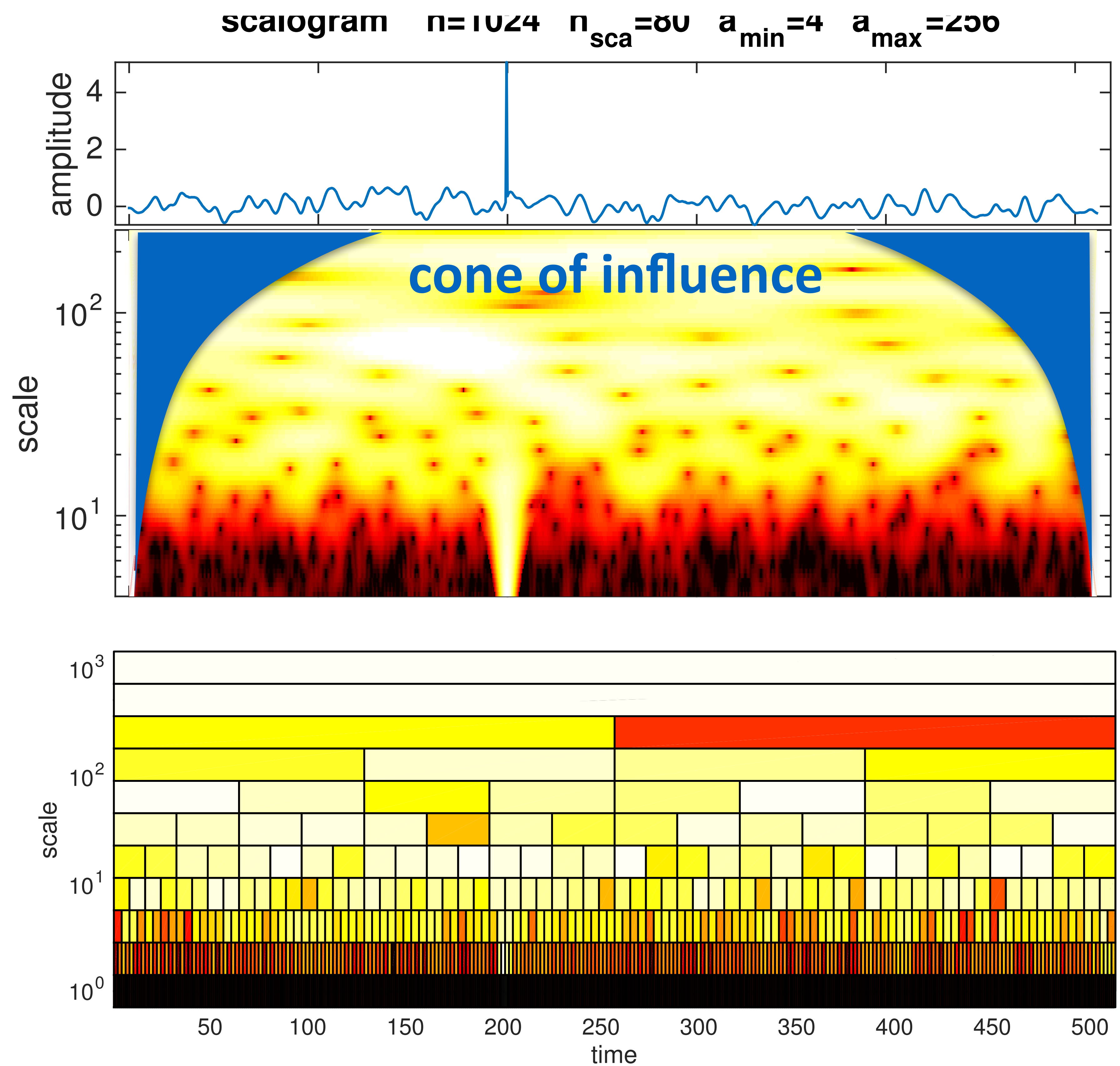
$$\tilde{f}_{ji} = \langle \psi_{ji} | f \rangle$$

$$\begin{tikzcd} f_k \arrow[rr, bend left, "\textbf{analysis}"] & & \tilde{f}_{ji} \arrow[ll, bend left, "\textbf{synthesis}"] \end{tikzcd}$$

$$f = \sum_j \sum_i \langle \psi_{ji} | f \rangle \; \psi_{ji}$$

with $\quad \psi_{ji} = 2^{\frac{j}{2}} \; \psi(2^j t - i)$

Continuous vs discrete



signal

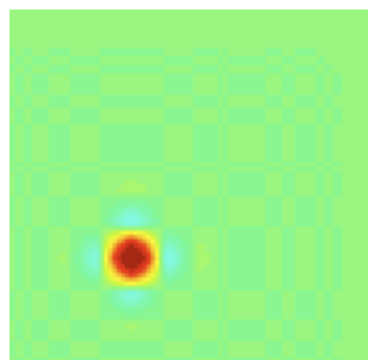
continuous transform (scalogram)

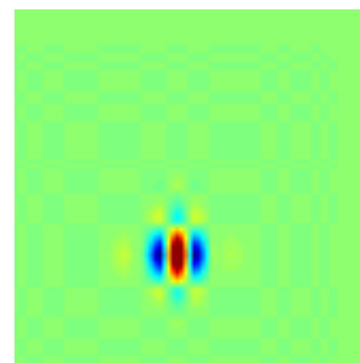
discrete transform

- This formalism can be readily extended to 2D, and beyond

$\phi(x)\ \phi(y)$

coarse approximation



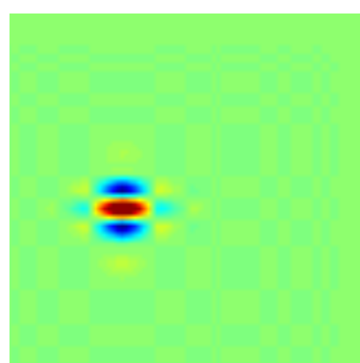


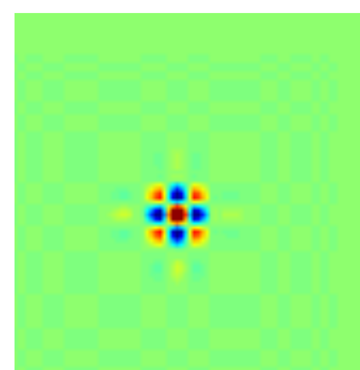
$\psi(x)\ \phi(y)$

horizontal details

$\phi(x)\ \psi(y)$

vertical details



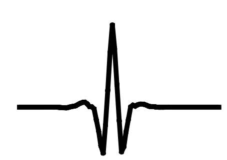


$\psi(x)\ \psi(y)$

diagonal details

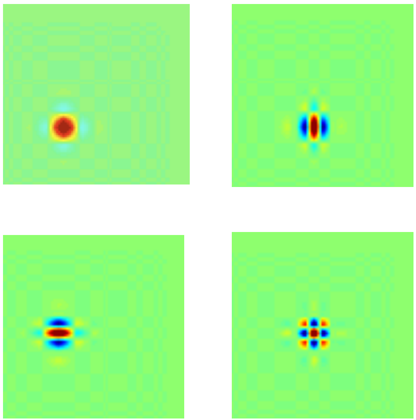
Scaling function (lowpass)

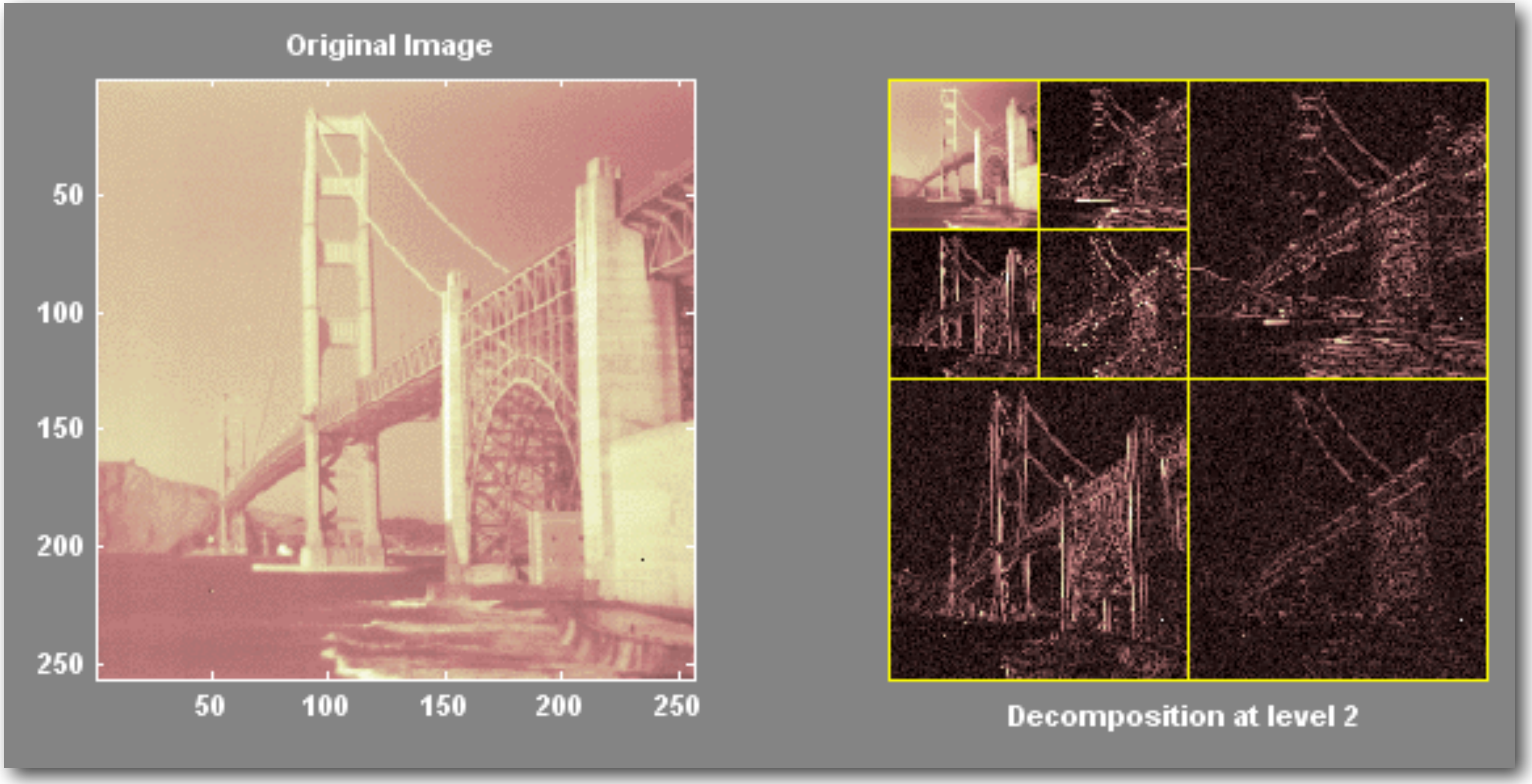
ϕ 

ψ 

Wavelet (highpass)

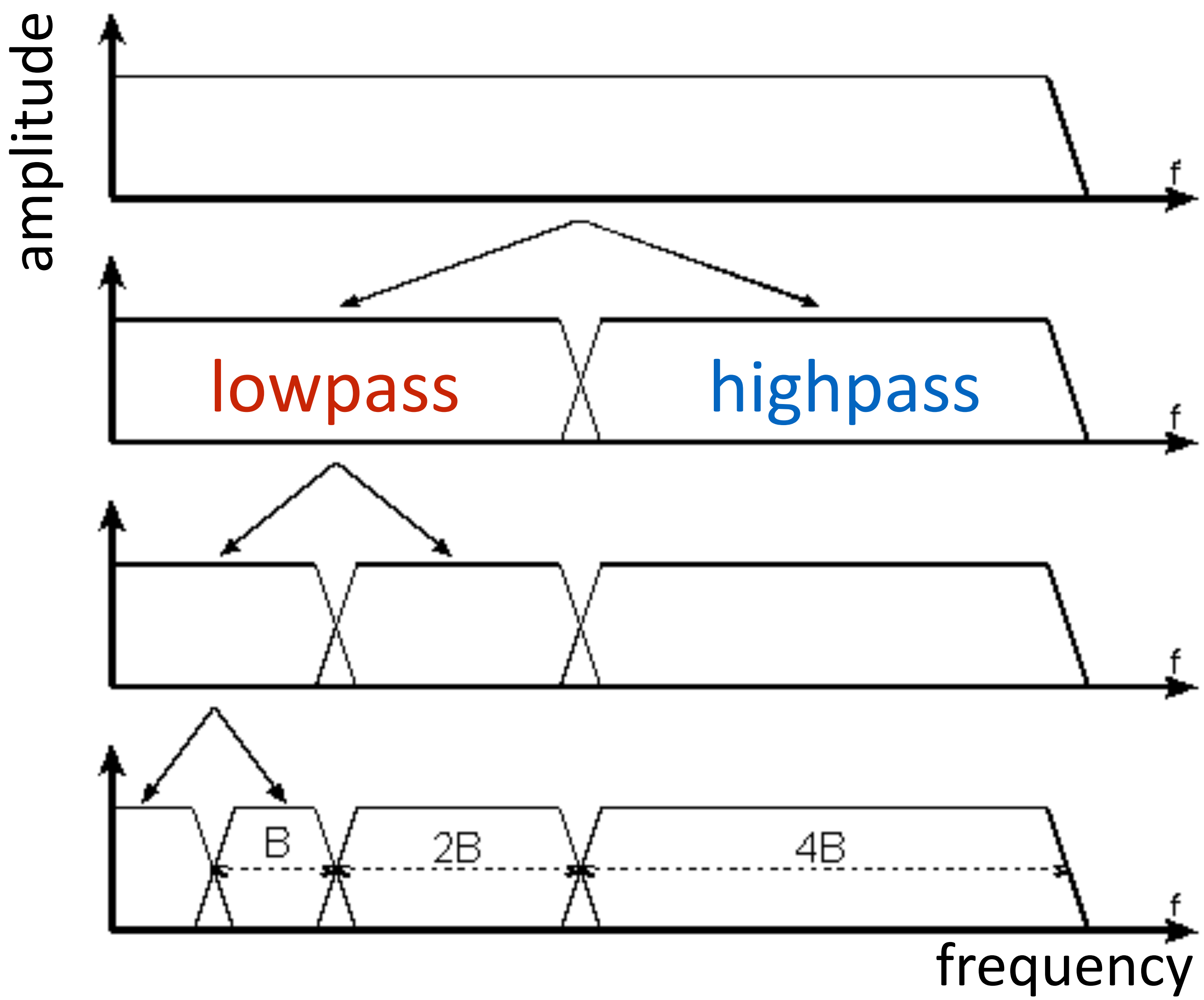
2D decomposition: example

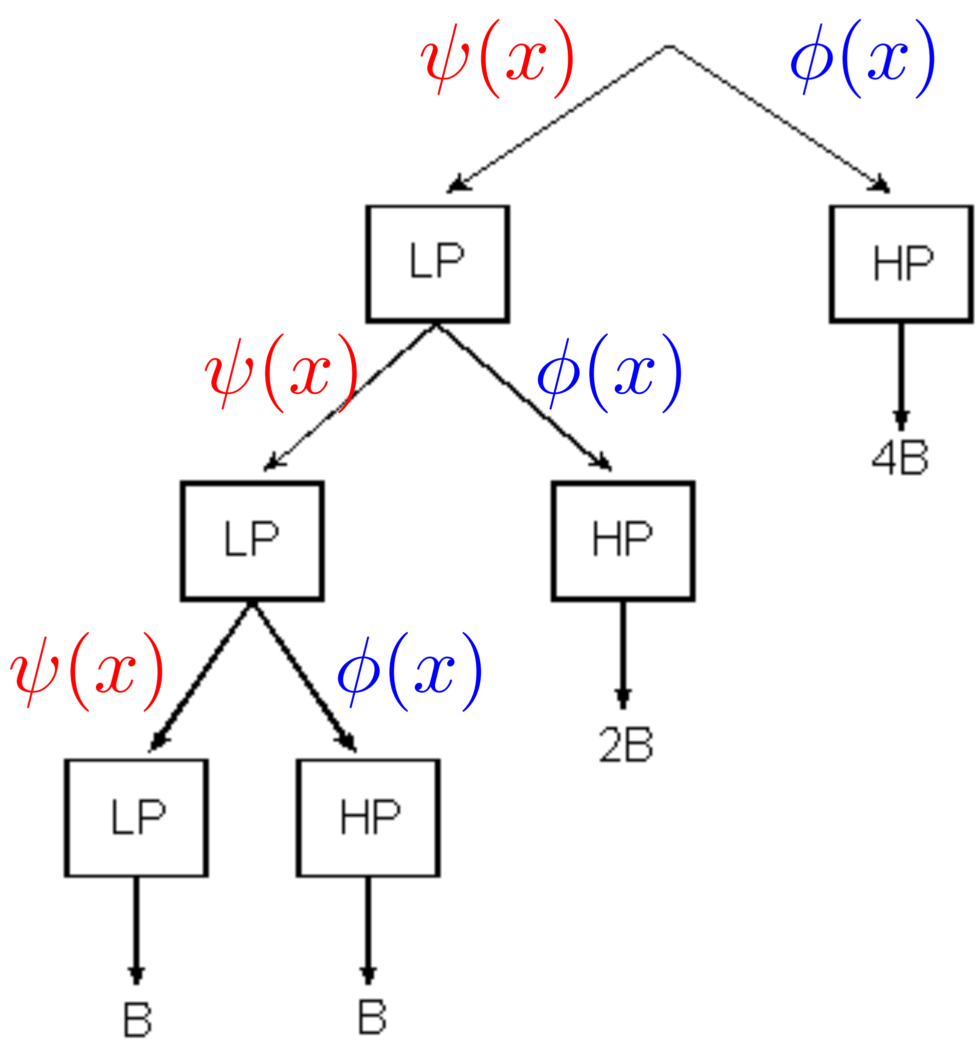




2D algorithmics

- There exist efficient computational schemes (filter banks)





Pros and cons

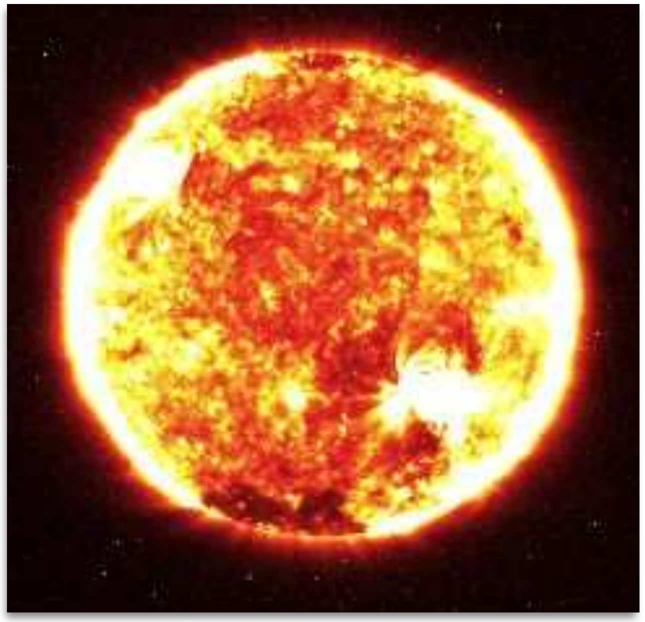
Continuous wavelet transform

- Good for analysis: interpreting time/scale content
- Coefficients are highly redundant
- Computationally expensive
- Not good for synthesis

Discrete wavelet transform

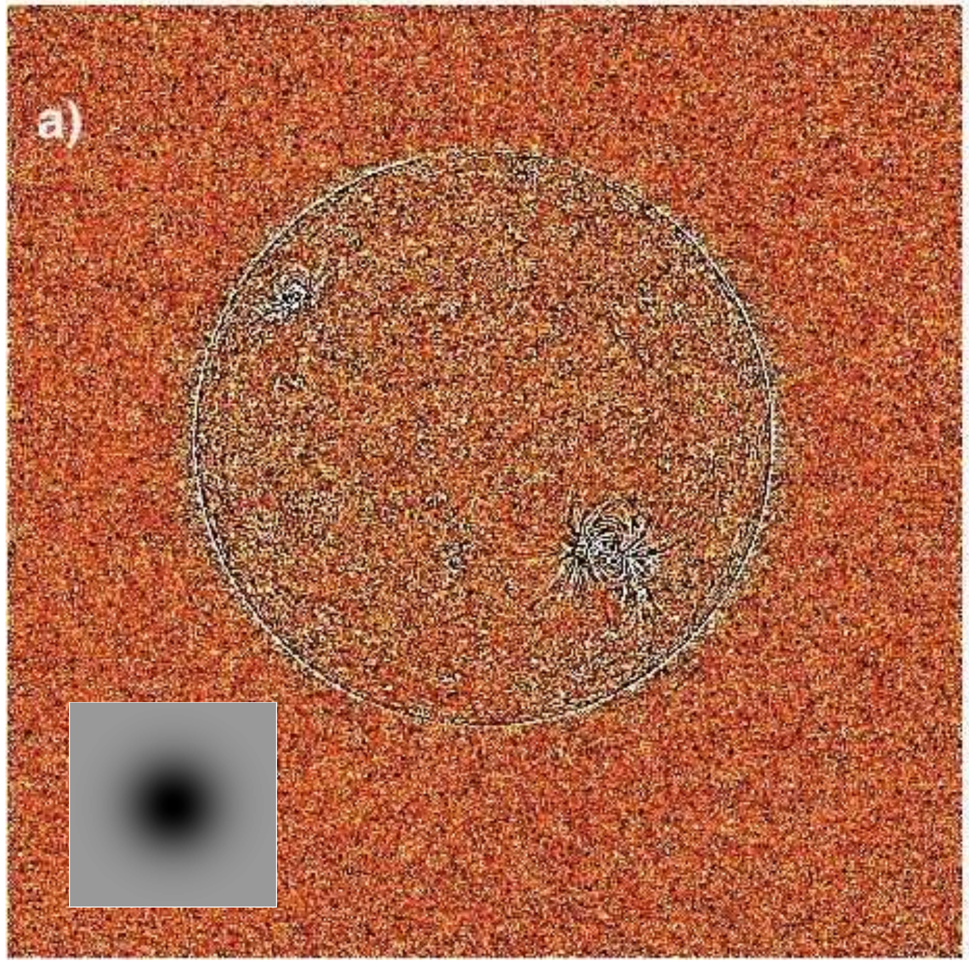
- Interpretation of coefficients is difficult (no translational invariance)
- Orthogonal basis
- Computationally efficient (faster than FFT)
- Good for filtering and compression

What wavelet should I choose ?

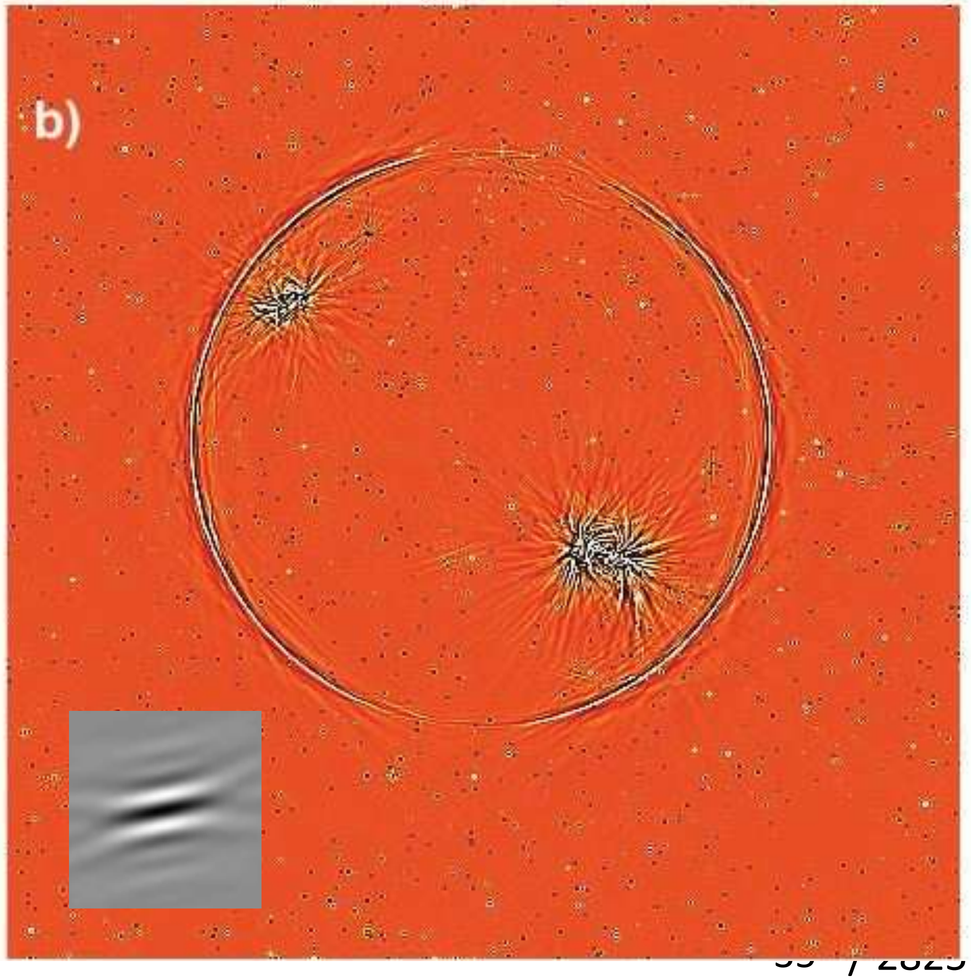


original solar EUV image (SWAP/PROBA2)

enhanced with isotropic wavelets = blurry



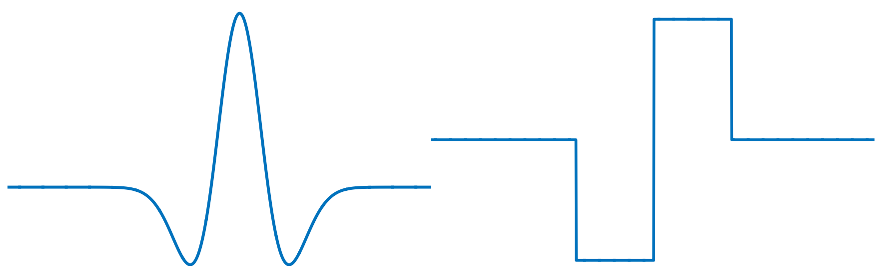
enhanced with curvelets = more appropriate



What wavelet should I choose ?

The choice of the mother wavelet SHOULD be driven by physical considerations

- what kind of structure am I looking for ?
- how regular is it (Hölder regularity) ?

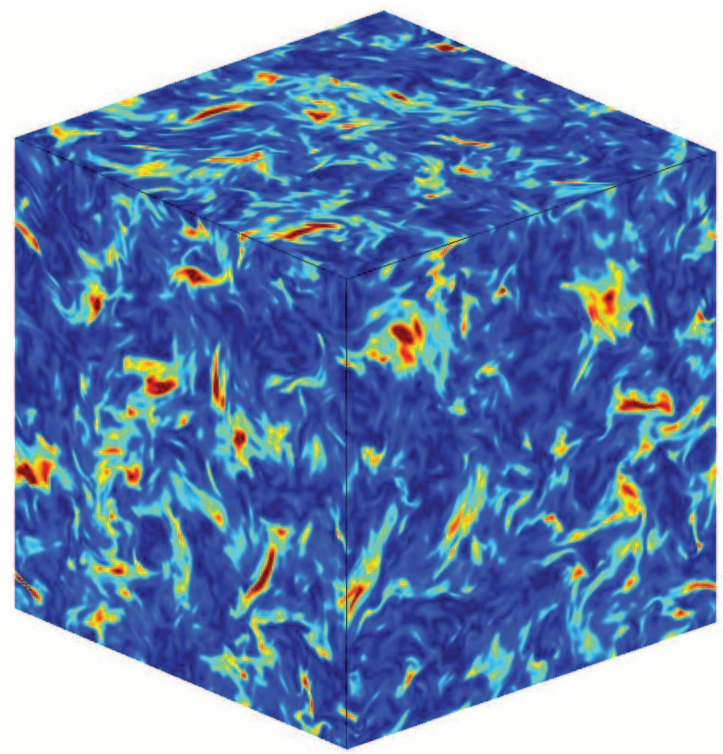


- symmetries ? what shape in 2D ? etc.

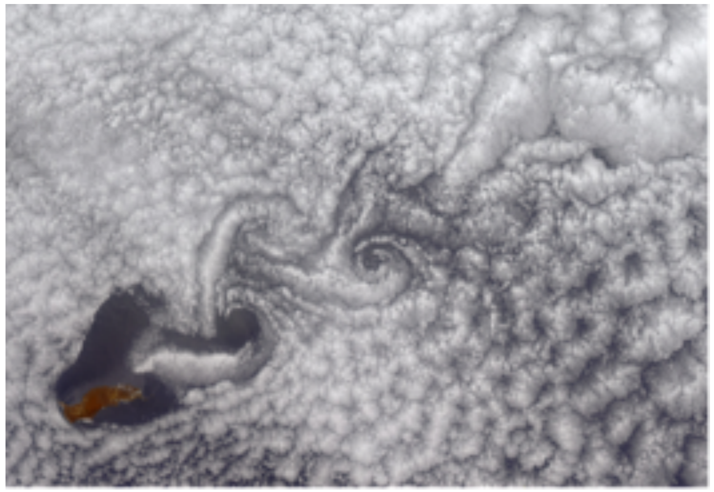
M. C. Escher

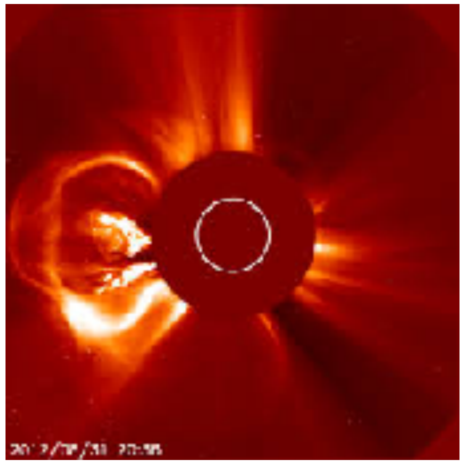
Coherent structures

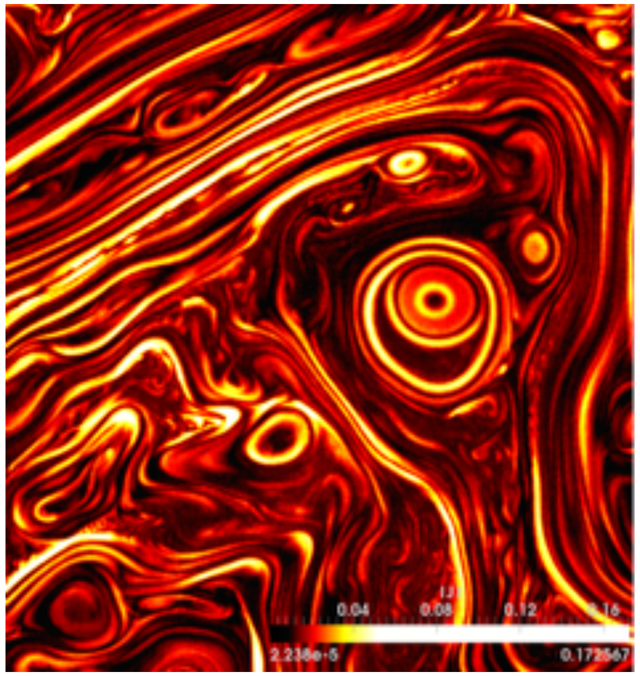
What is a coherent structure ?



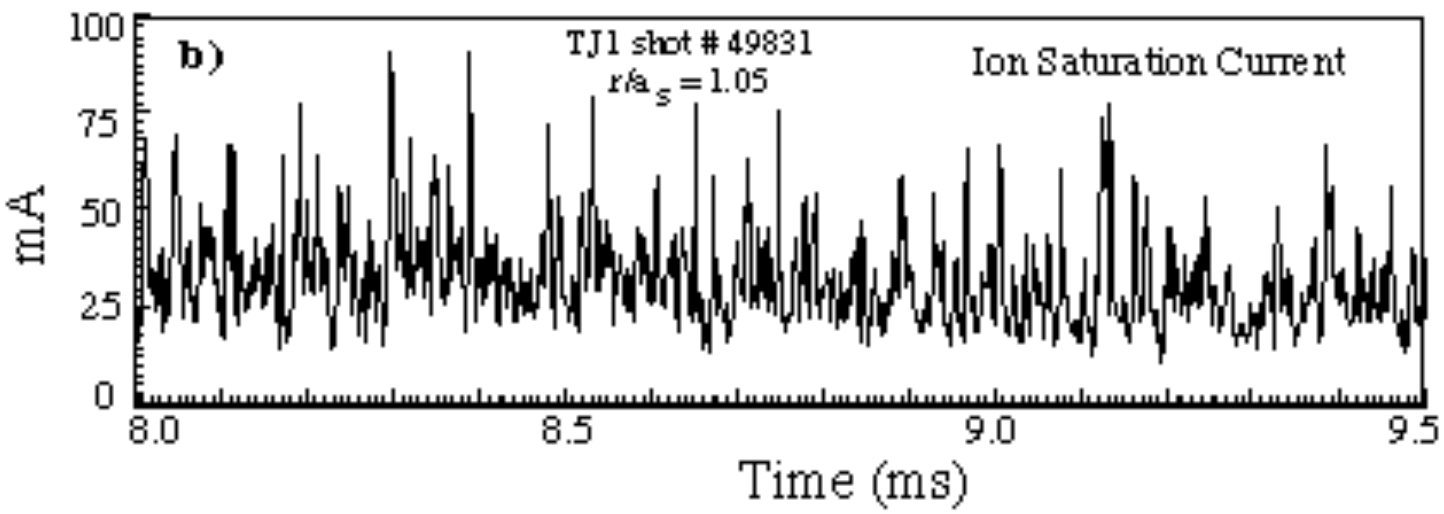
Arzner et al. (2006)







Karimabadi et al. 2013

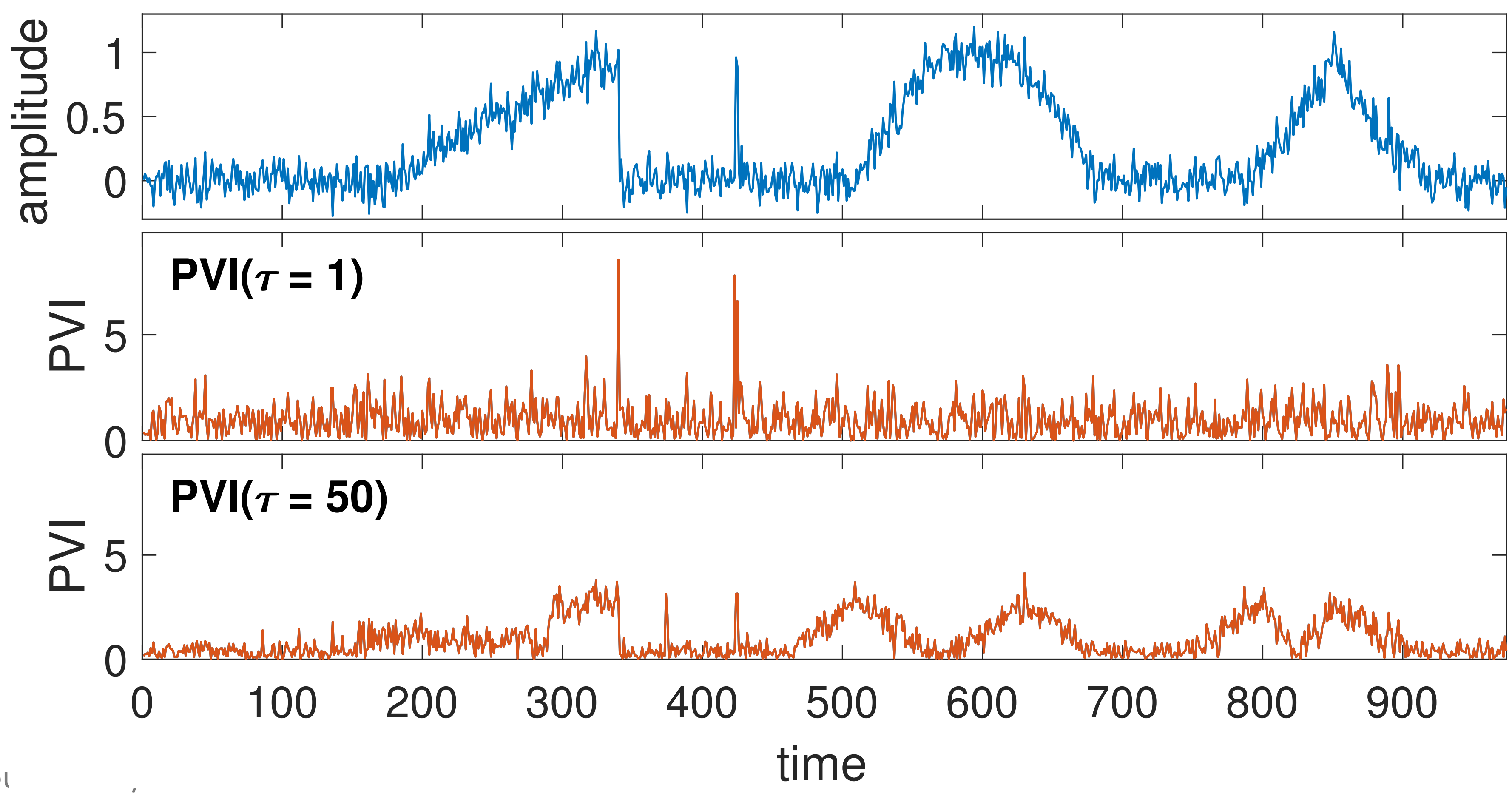


C. Hidalgo et al. (2002)

How do I identify a coherent structure ?

- local accumulation of “energy” in real space
  - PVI : partial velocity increments (Greco et al. 2008)

$$\mathrm{PVI} = \frac{|\Delta f(t)|^2}{\langle |\Delta f(t)|^2 \rangle} \qquad \Delta f(t) = f(t+\tau) - f(t)$$



How do I identify a coherent structure ?

- local accumulation of “energy” in real space
  - PVI : partial velocity increments (Greco et al. 2008)

$$\text{PVI} = \frac{|\Delta f(t)|^2}{\langle |\Delta f(t)|^2 \rangle} \qquad \Delta f(t) = f(t+\tau) - f(t)$$

- local accumulation of “energy” in wavelet space
  - LIM : local intermittency measure (Farge, 1992)

$$\text{LIM} = \frac{|\tilde{f}(t,a)|^2}{\langle |\tilde{f}(t,a)|^2 | \rangle_t}$$

- phase coherence
  - the phases are somehow coupled to each other

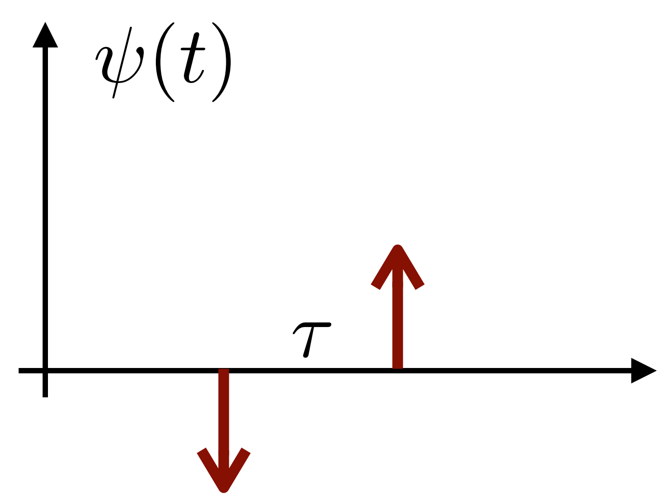
PVI versus LIM

- The PVI is a special case of the LIM, for a wavelet that actually has poor properties...

$$\mathrm{PVI} = \frac{|\Delta f(t)|^2}{\langle|\Delta f(t')|^2\rangle_{t'}}$$

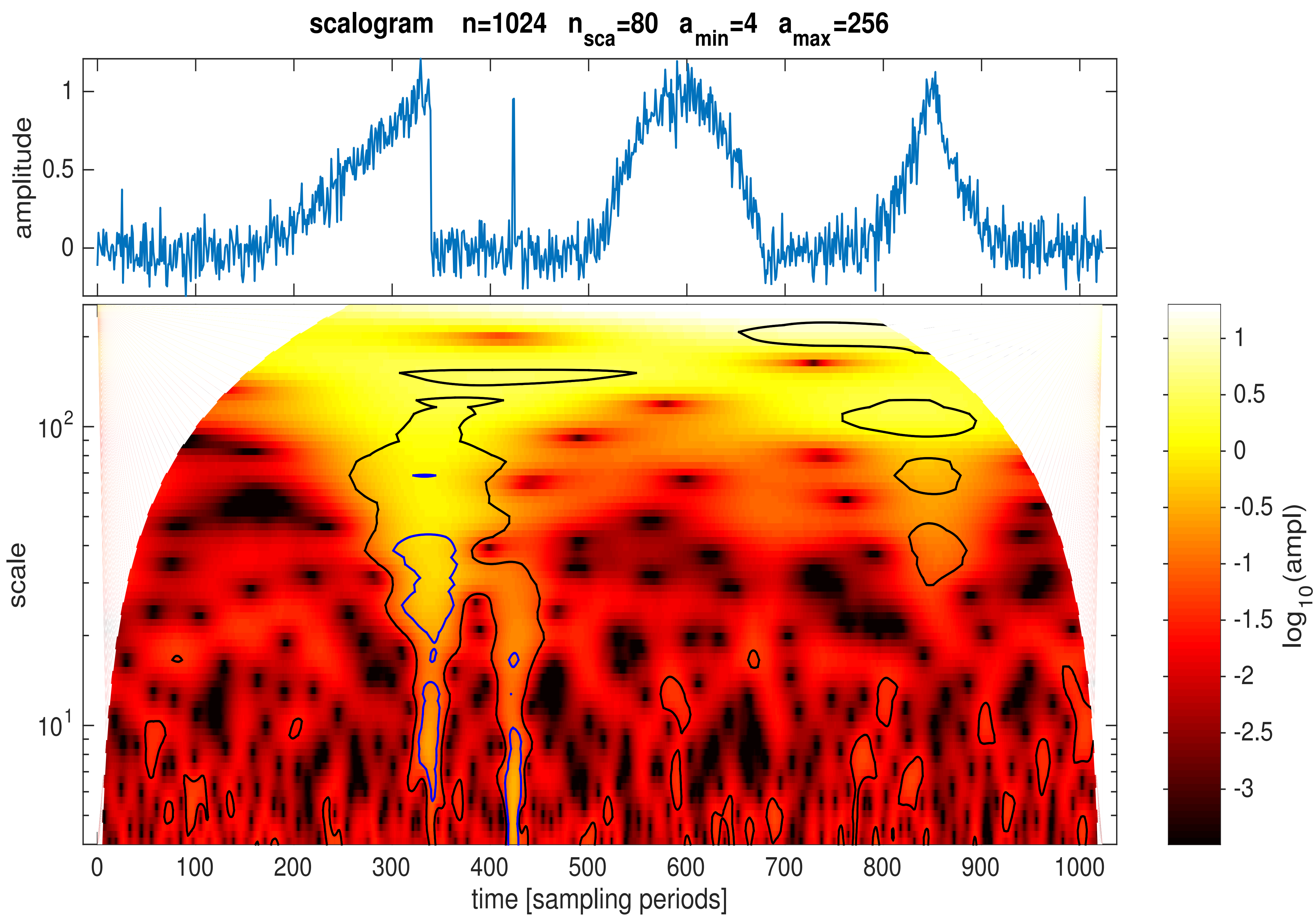
$$\tilde{f}(t,a) = f(t) * \big(\delta(t+\tau) - \delta(t)\big)$$

$$\mathrm{LIM} = \frac{|\tilde{f}(t,a)|^2}{\langle|\tilde{f}(t',a)|^2|\rangle_{t'}}$$



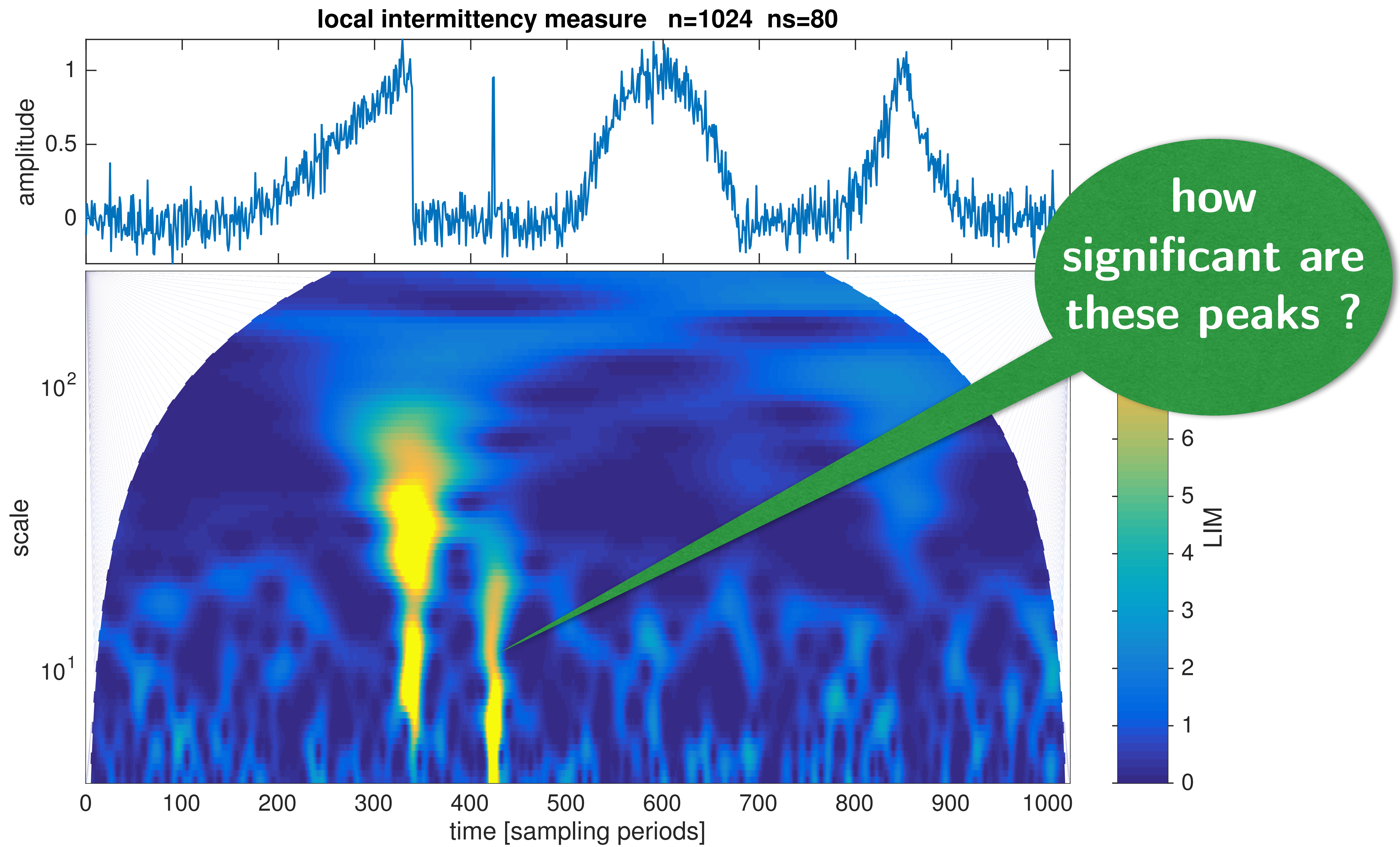
Example

Scalogram



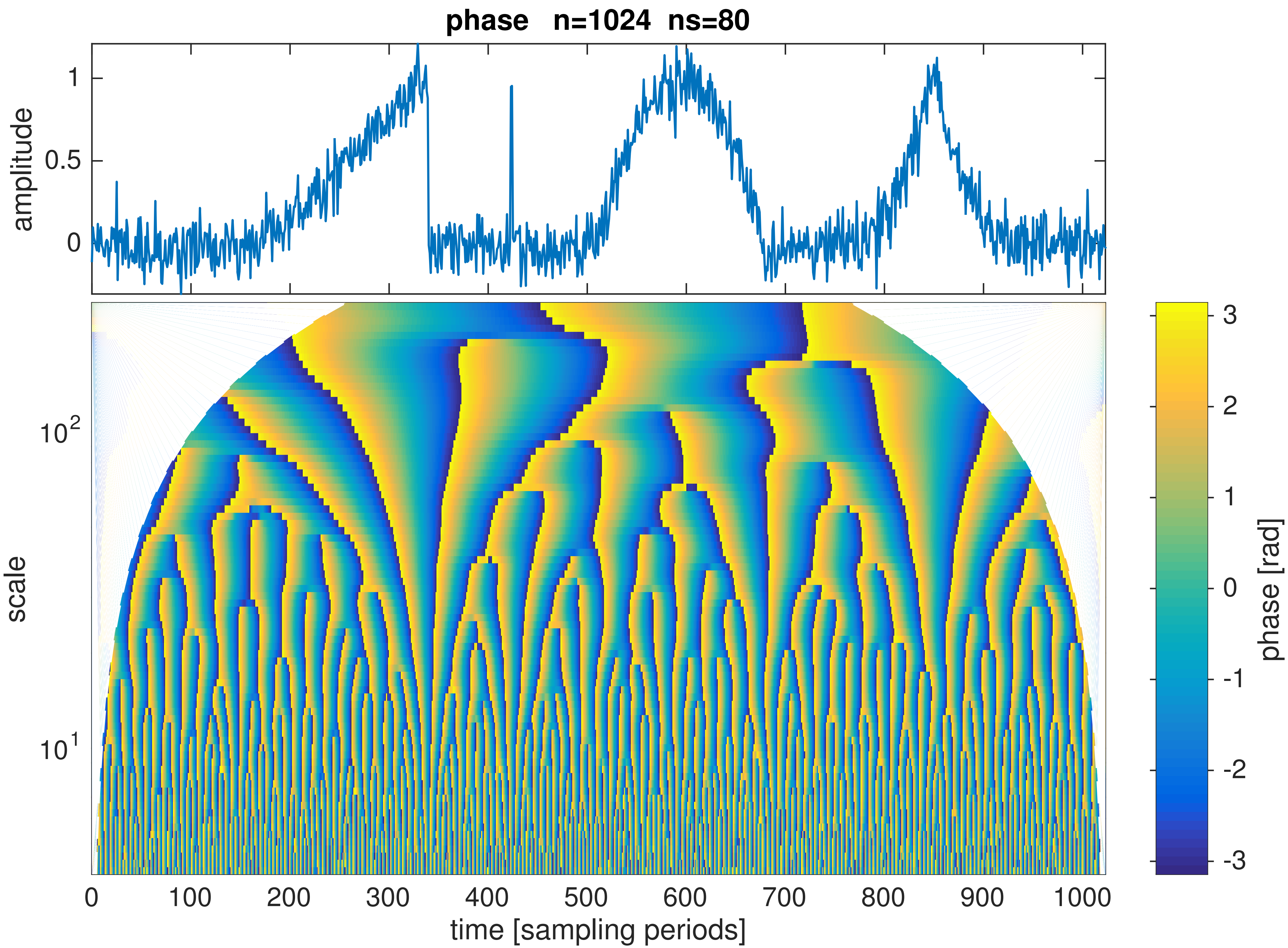
Example

Local Intermittency Measure (LIM)



Example

Phase



Intermediate conclusion (1)

- Energy-based measures (LIM, PVI, ...) are mostly sensitive to sharp gradients (e.g. current sheets)
- Are they able to detect "real" coherent structures (solitons, wave packets, etc.) ?
- Phase information is difficult to exploit

We need a different approach

M. C. Escher

What is a coherent structure ?

What is a coherent structure ?

- Instead of defining what a coherent structure **IS**
- Define what it is **NOT**

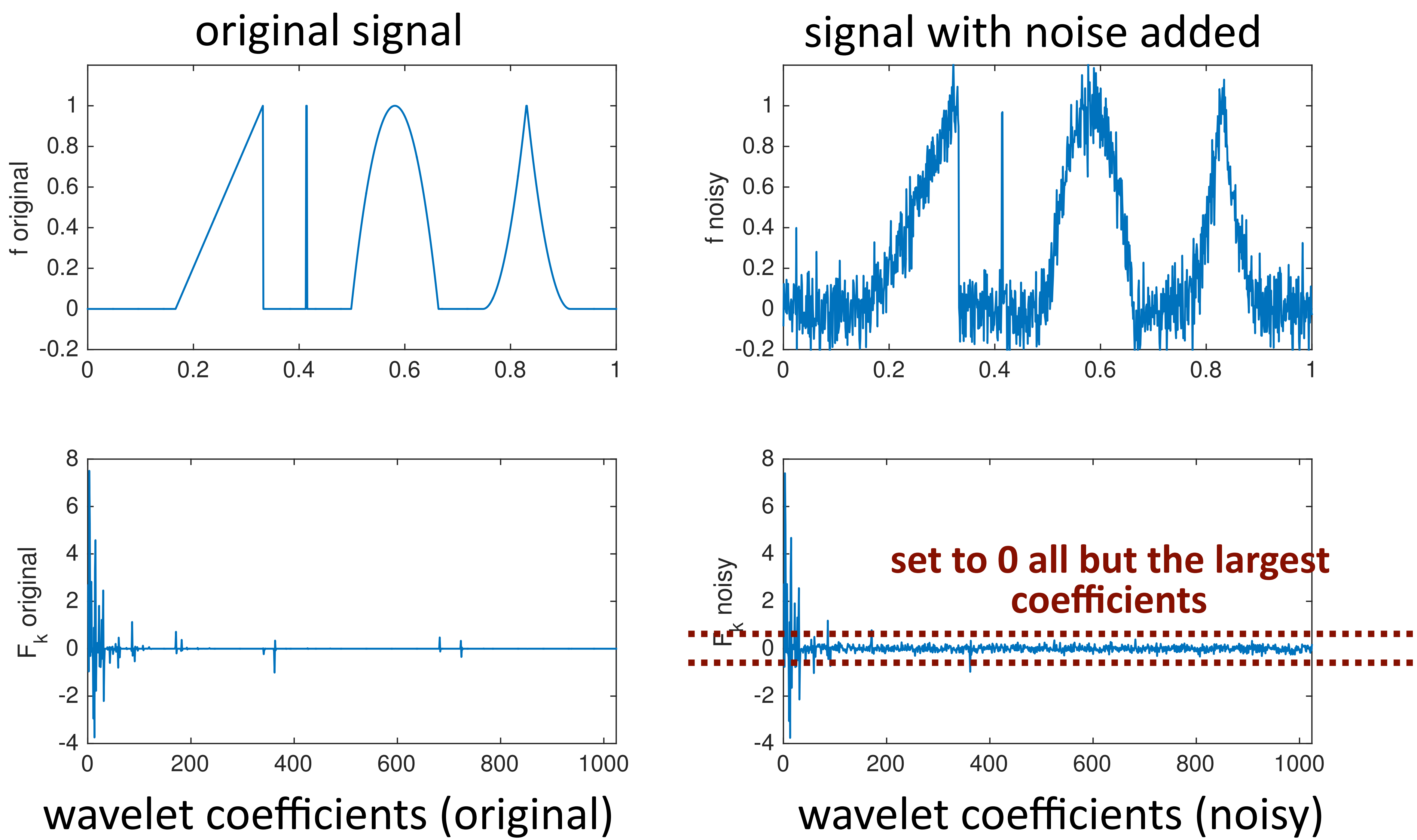
Coherent structure ∩ noise = ∅

"Noise" cannot be properly projected on any basis functions,
(≈ its energy is distributed over many scales, with no phase coherence)

Example : is there a coherent structure ?

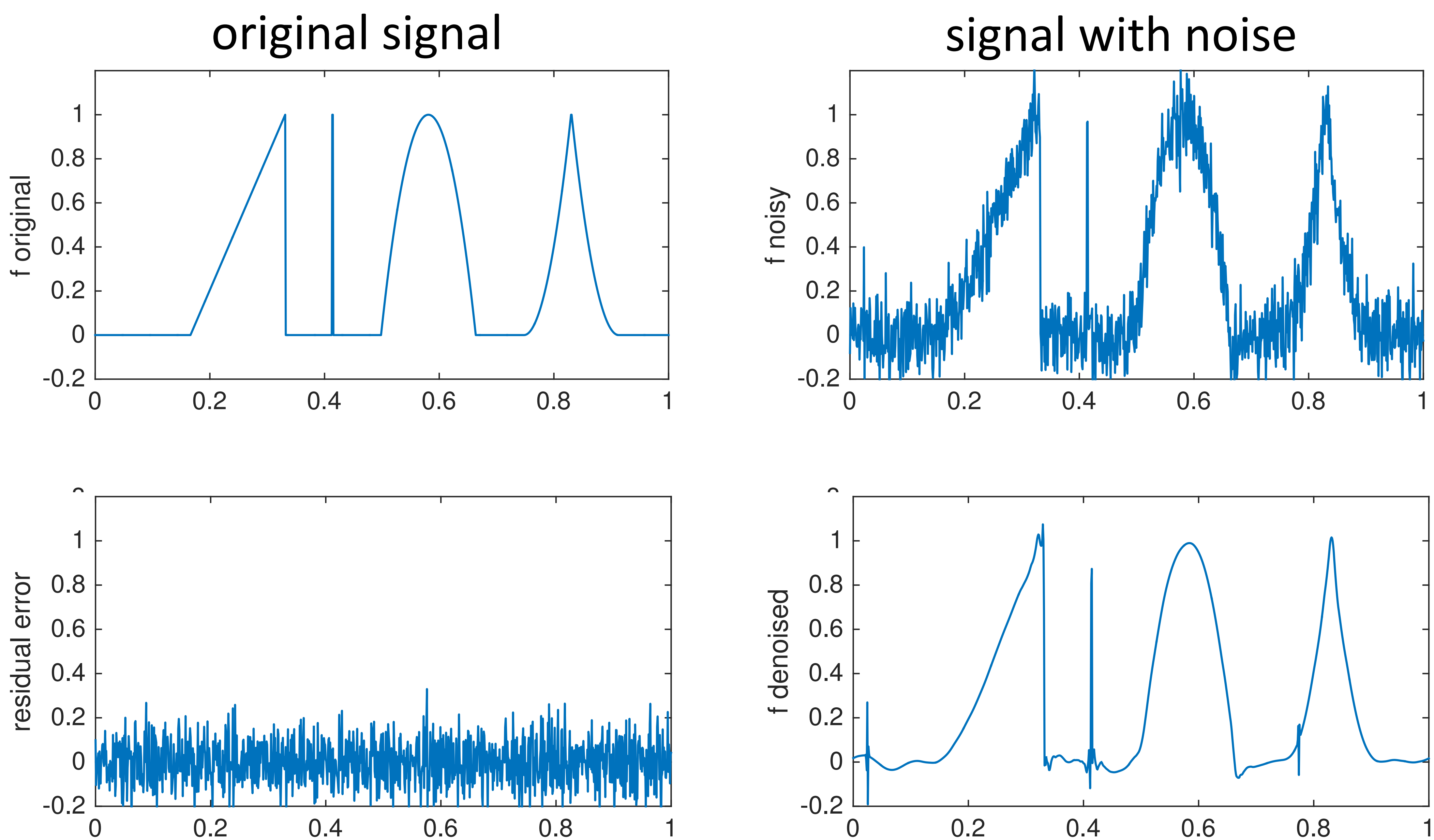
g h j d s e 2 2 z c l b g r o n k h f
r v de 4 g 7 n 0 n v e 3 c q z d g
b m l p 0 7 c s g y 6 g 🐔 t d 4 f h
i 4 s 3 h 8 6 h h f 4 g 7 n 0 n v d
5 d d s w t u k i b y f 4 g 7 n 0 n
v e 3 c q z d g b m h i g 5 v k 9 j
v 5 c 5 h u g t y l p j n q v f c

How wavelet denoising works



Noise is spread out evenly, whereas the coherent structures are concentrated in a few coefficients

How wavelet denoising works



residual error (= noise)

denoised signal

Denoising

- These methods are well documented
  - there exists a rigorous framework for determining optimal thresholds, etc. [Donoho, Mallat, ...]

Denoising

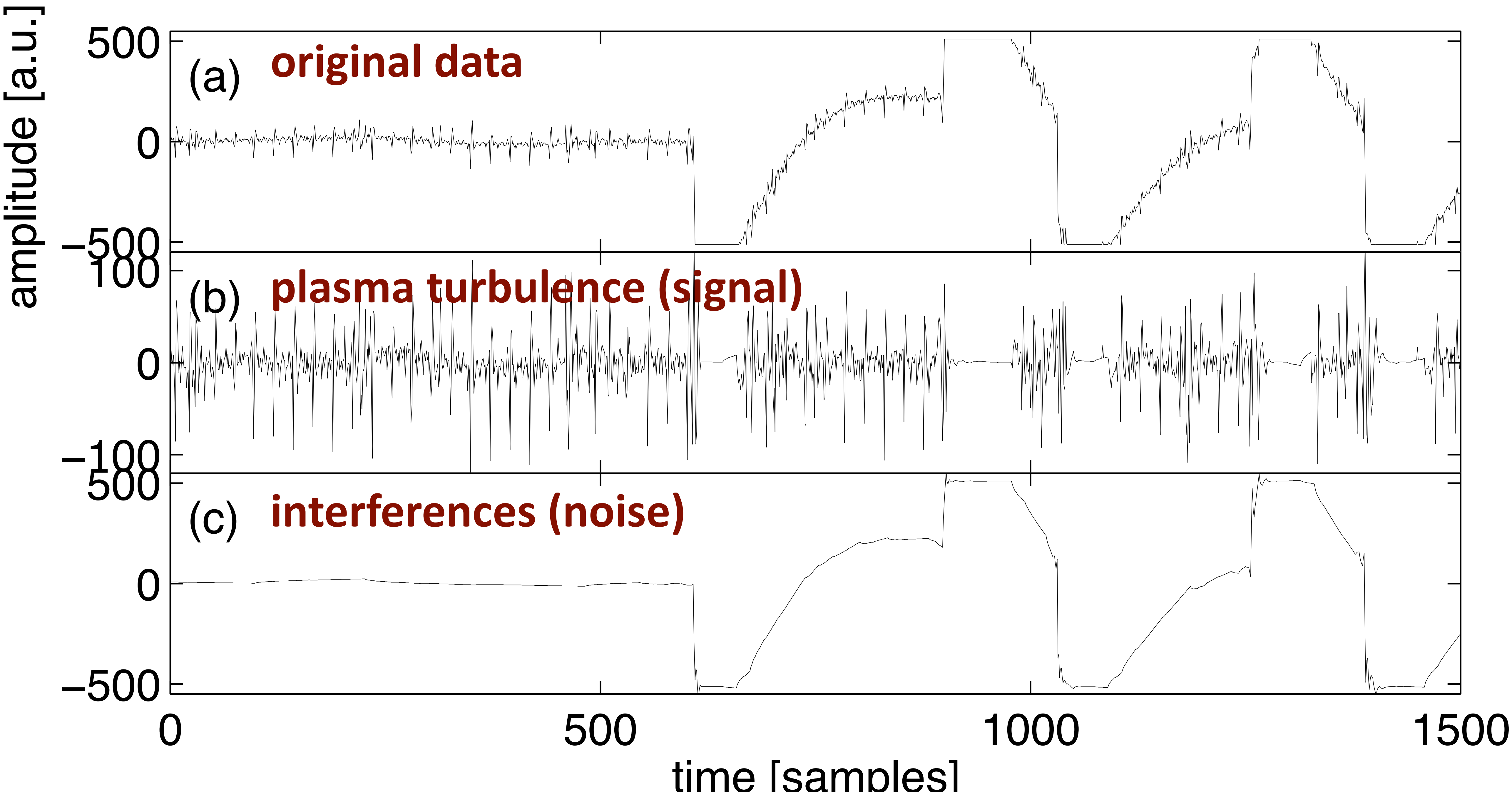
One man’s noise is another man’s signal

Example : noisy magnetic field measurements in the ionosphere

M. C. Escher

Denoising

Interference in AC magnetic field data (CUSP2000 sounding rocket)

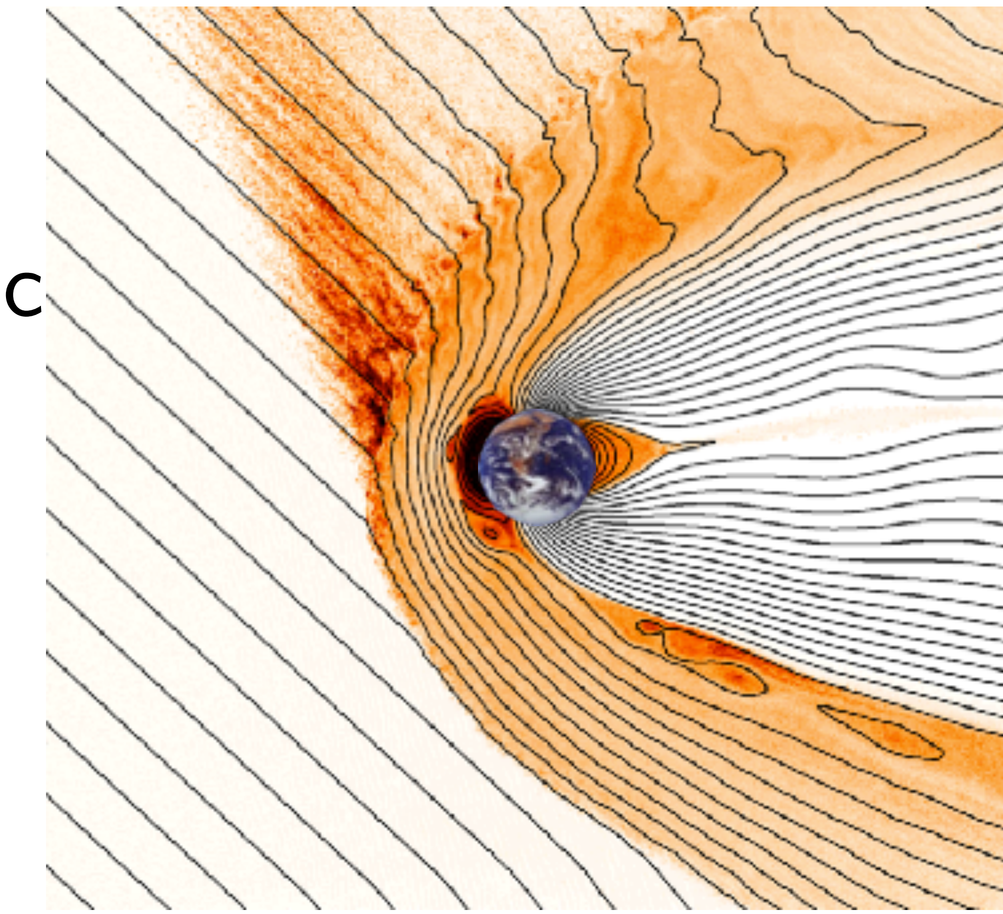


M. C. Escher

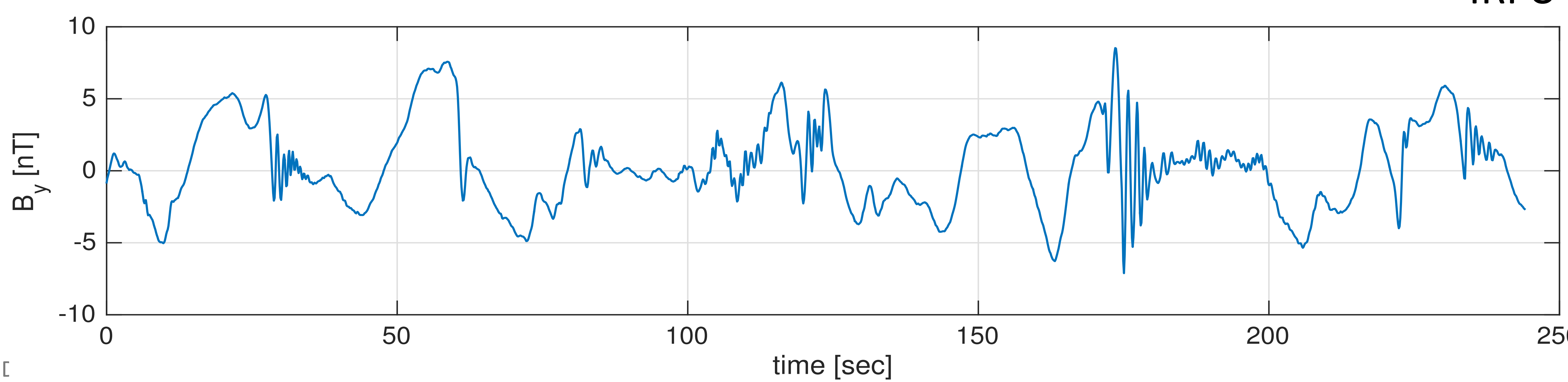
Example : magnetic structures in the magnetosphere

Example

- Short-Large Amplitude Magnetic Structures (SLAMS) upstream of the Earth's quasi parallel bow shoc

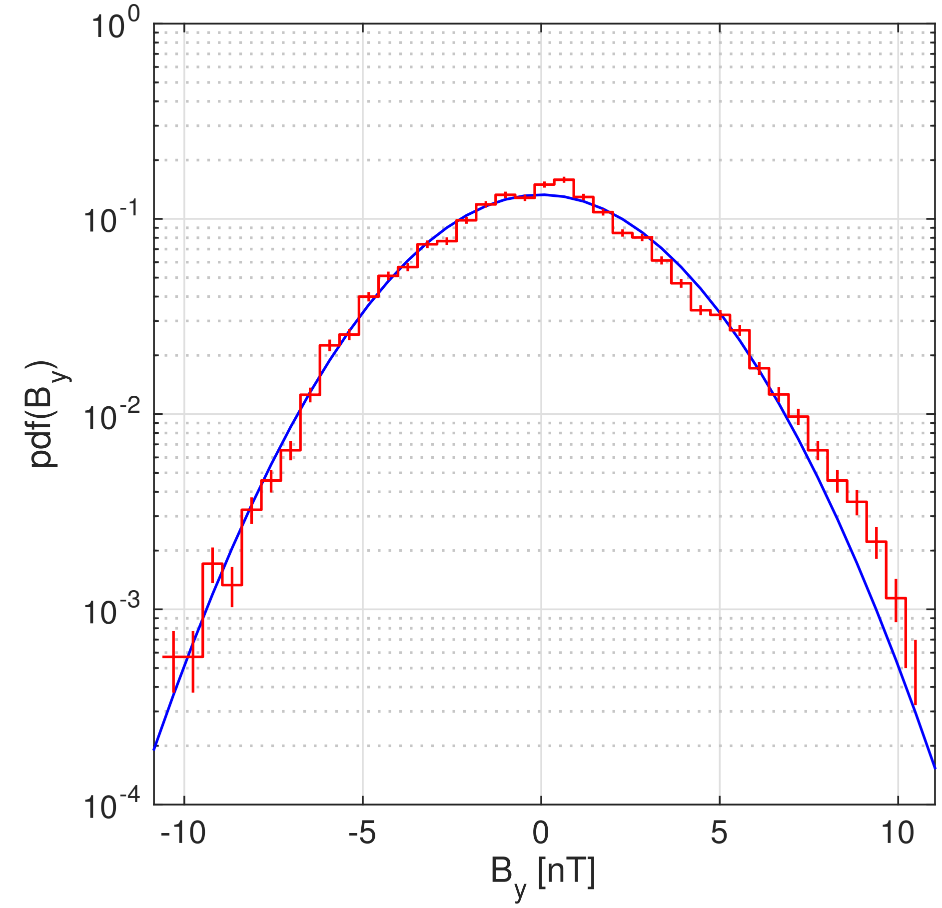


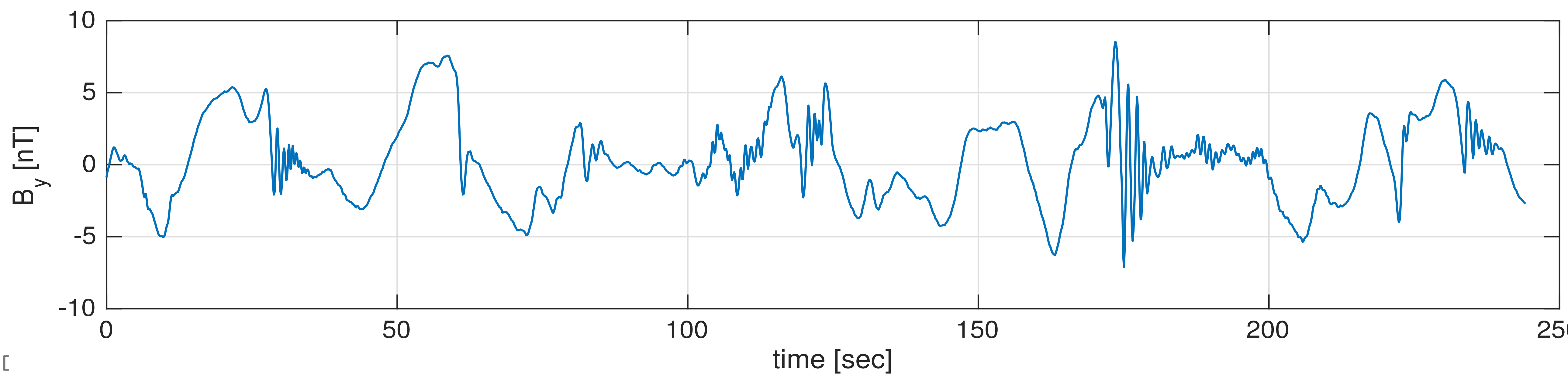
IRFU



Example

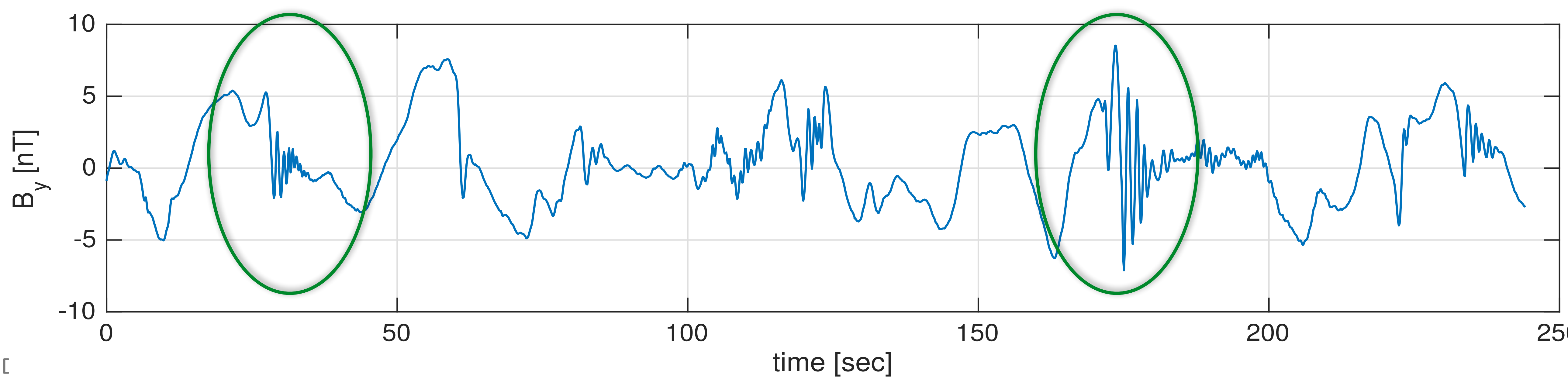
- The probability density function is close to Gaussian...





Example

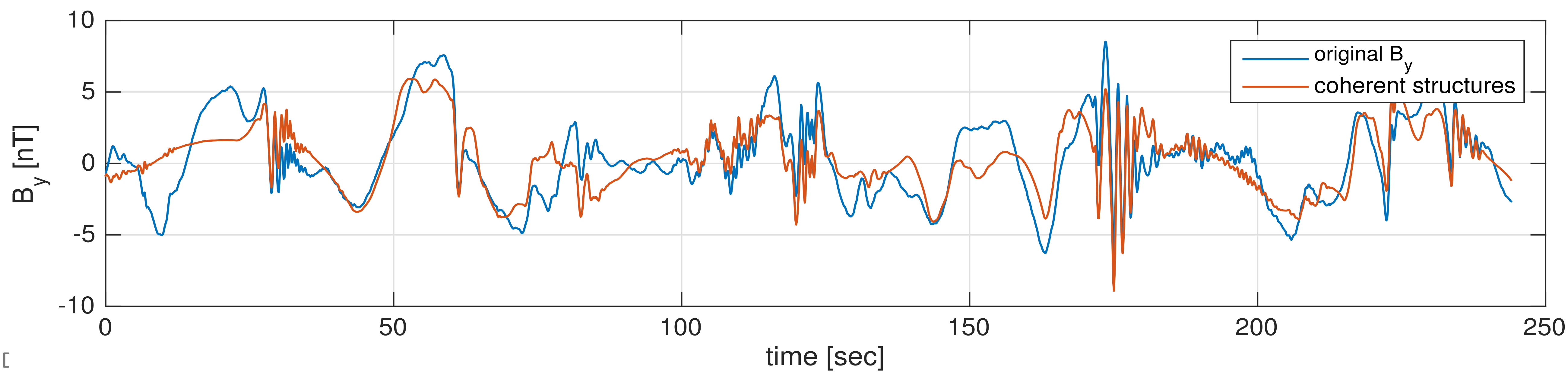
- But the variations are not that random : dispersive wave packets



Example

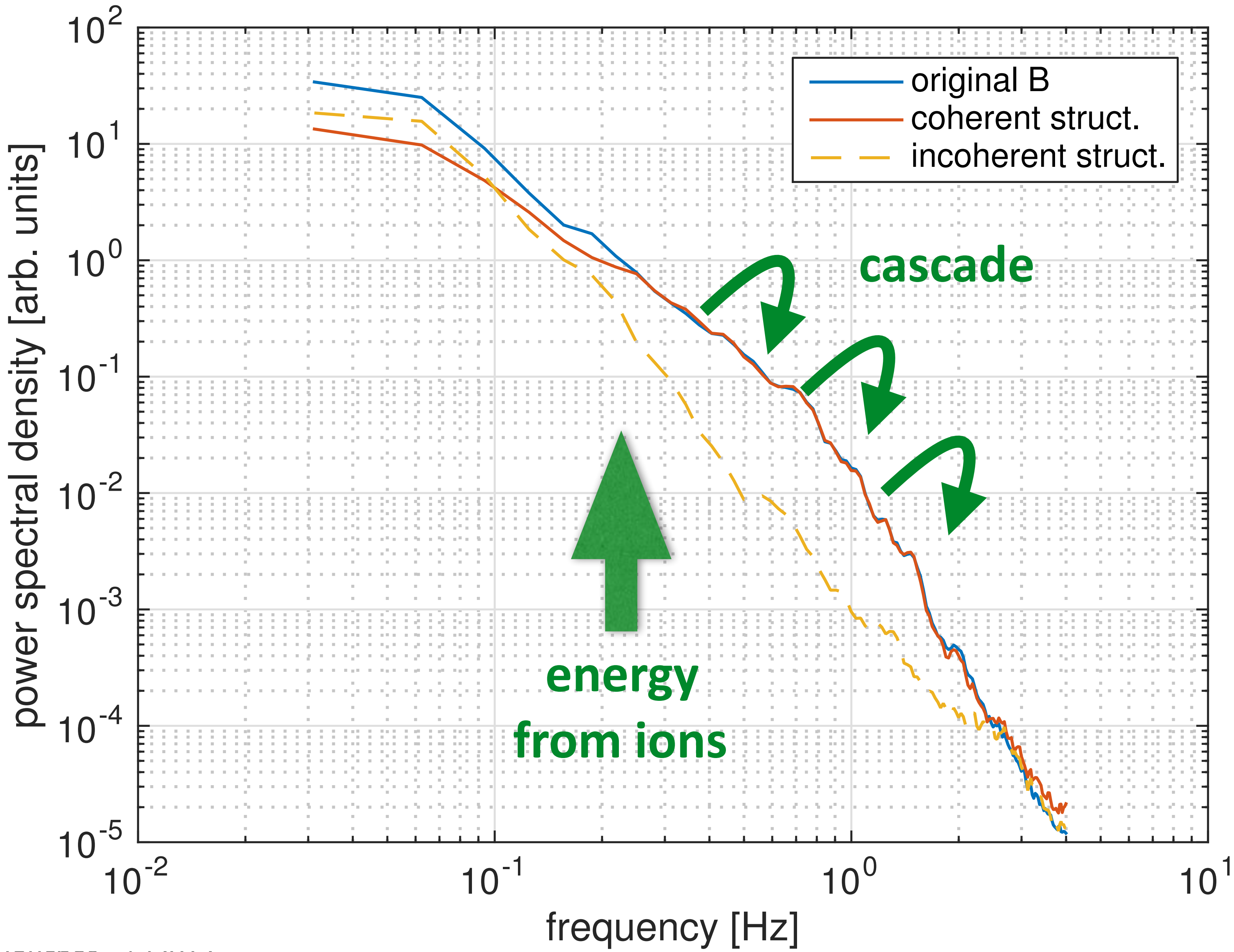
- Using the discrete wavelet transform, decompose

$$B(t) = B_{\text{coh. struct.}}(t) + B_{\text{incoherent struct.}}(t)$$



Example

- The power spectral density shows that these structures are concentrated in the spectral domain

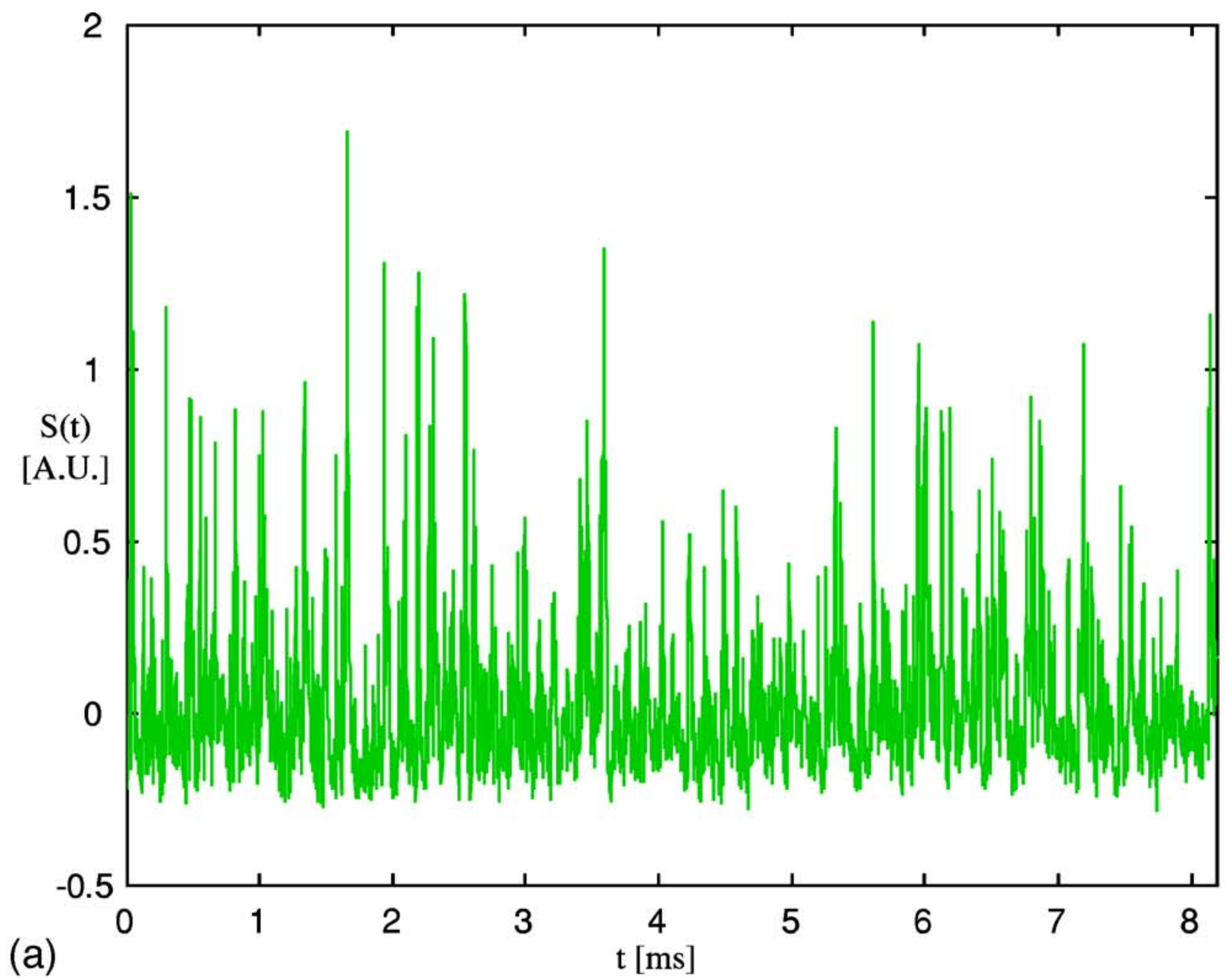


M. C. Escher

Example : intermittency in tokamak edge turbulence

Example

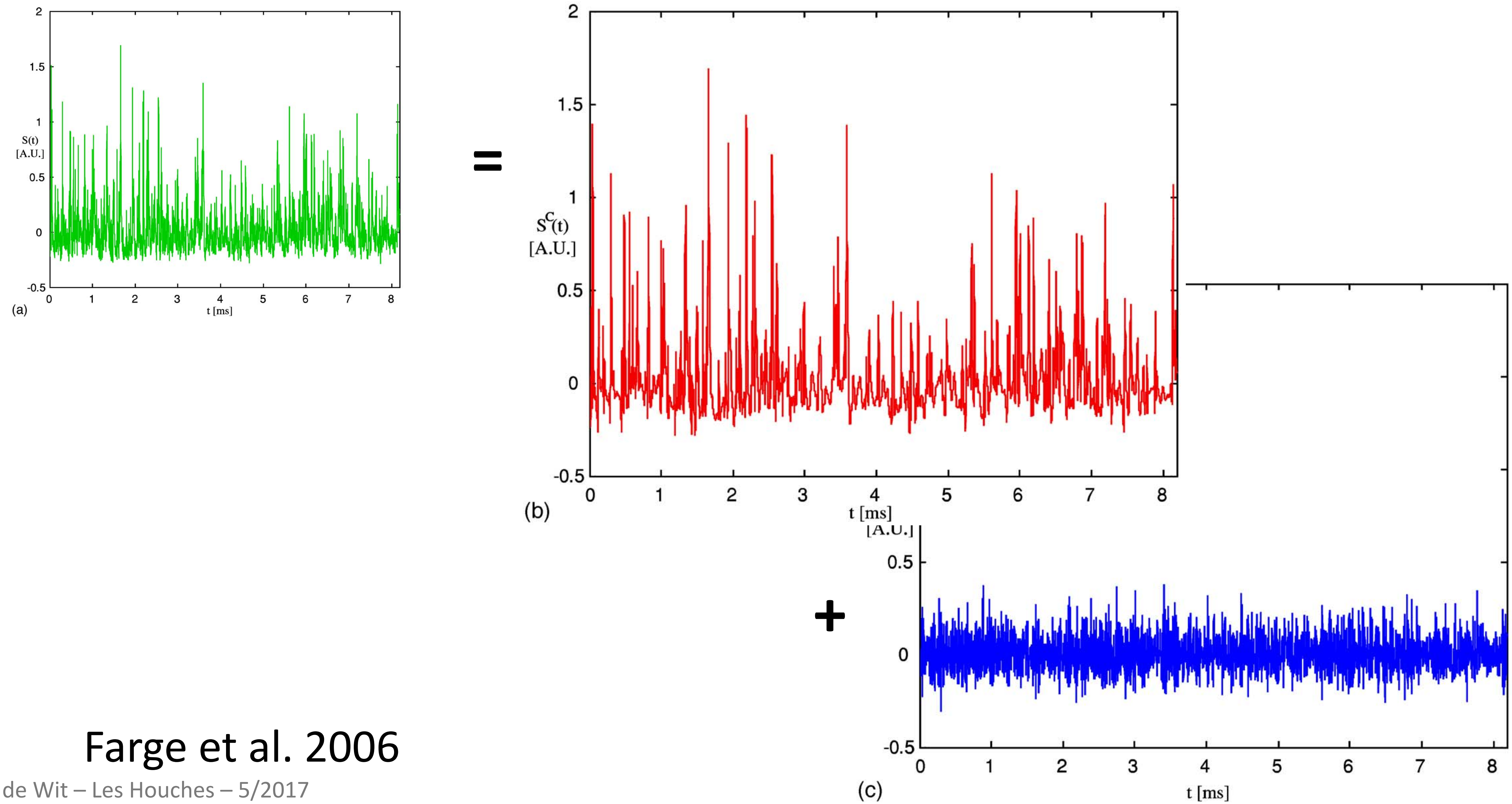
- Intermittent density bursts in a tokamak edge plasma (Tore Supra)



Farge et al. 2006

Example

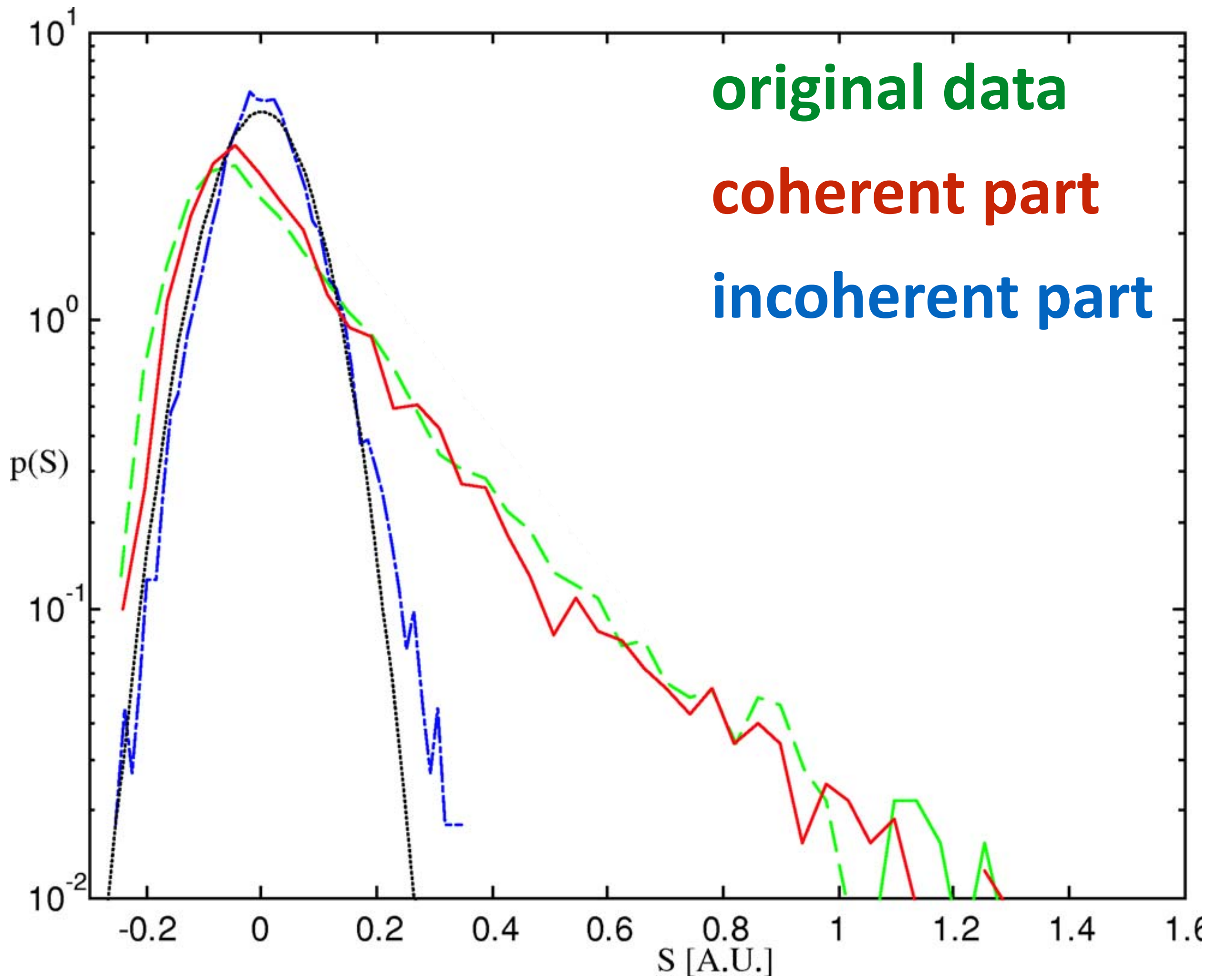
- Density fluctuations = bursts + incoherent fluctuations



Farge et al. 2006

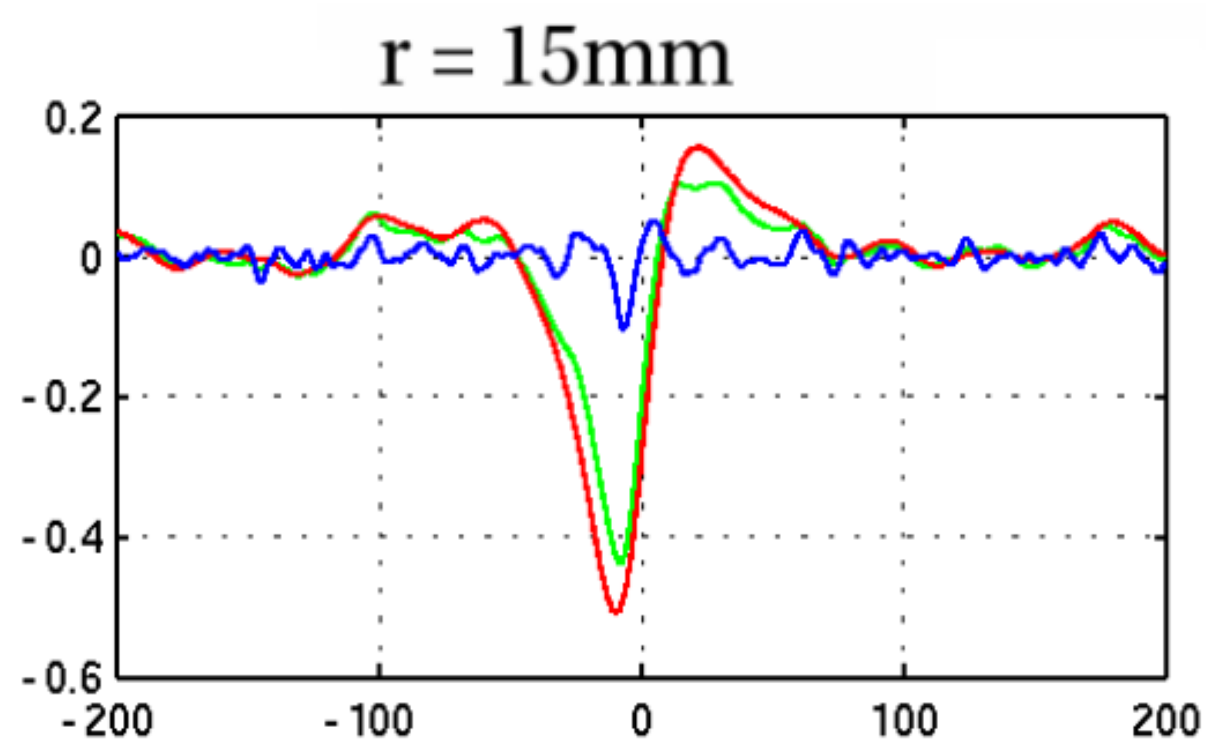
Example

- pdf of density fluctuations = exponential + Gaussian



(Farge et al. 2006)

Energy flux vs time
(Schneider et al., 2015)



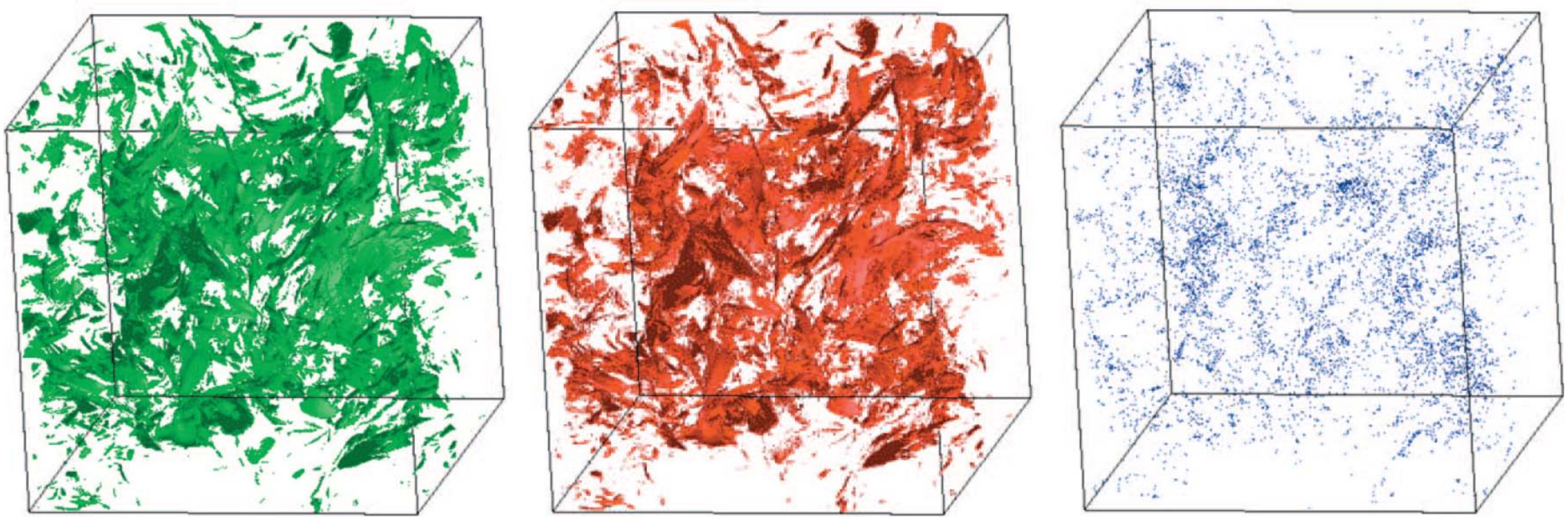
M. C. Escher

Example :
MHD simulations

Example

- Application to 3D resistive MHD turbulence simulations (Yoshimatsu et al., 2009)

(Schneider et al., 2015)



Vorticity = current sheets + dissipative incoherent structures

Take home message

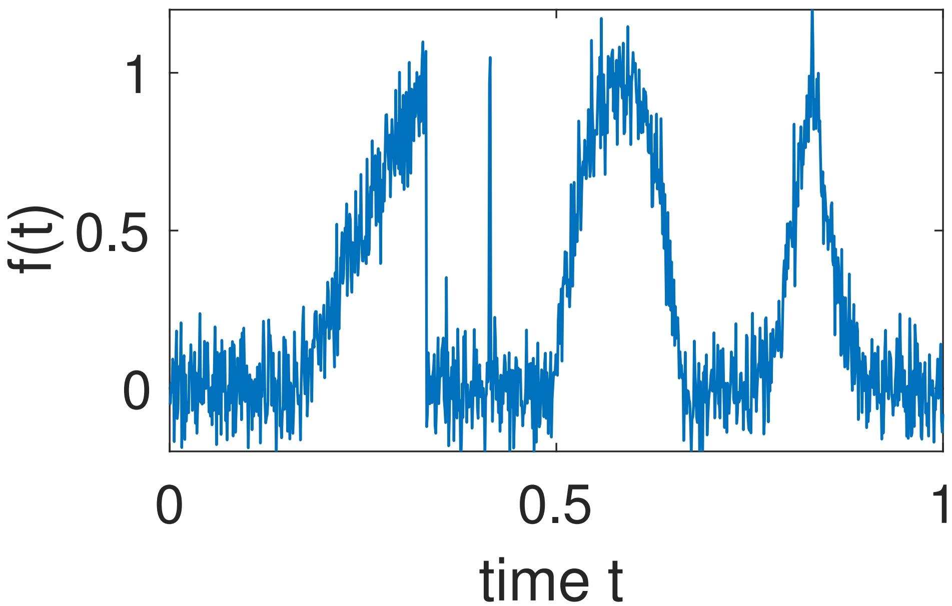
- Multiscale decompositions offer a **natural description** of plasma phenomena
- The discrete wavelet transform is more relevant than the continuous one for extracting structures.
- Coherent structures $\neq$ noise

M. C. Escher

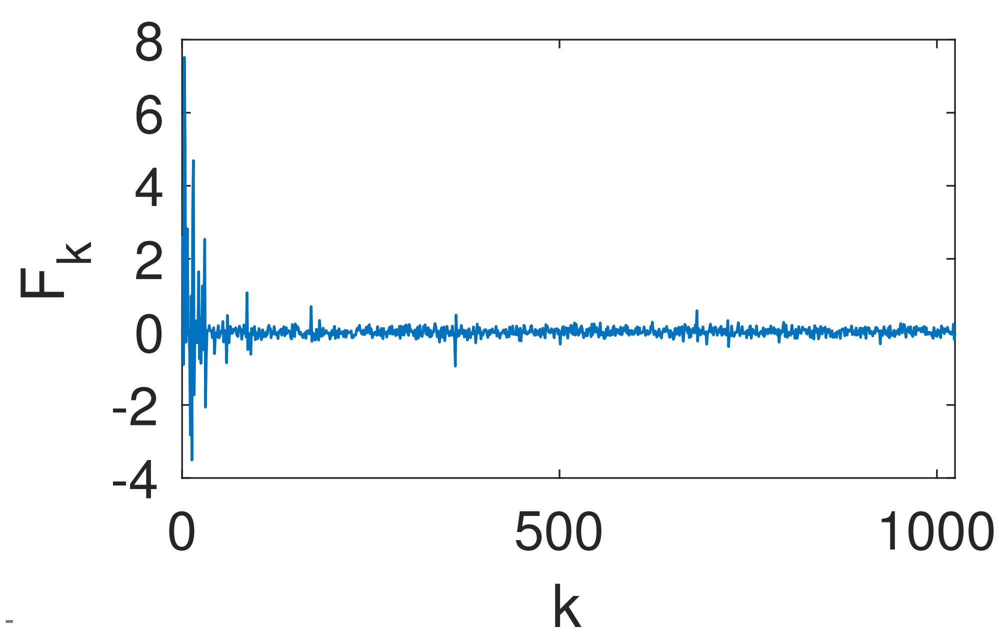
Sparsity

Sparsity

- We saw that the discrete wavelet transform of a time series generally has few outstanding coefficients



time series



its wavelet coefficients :
less than 5% are significant

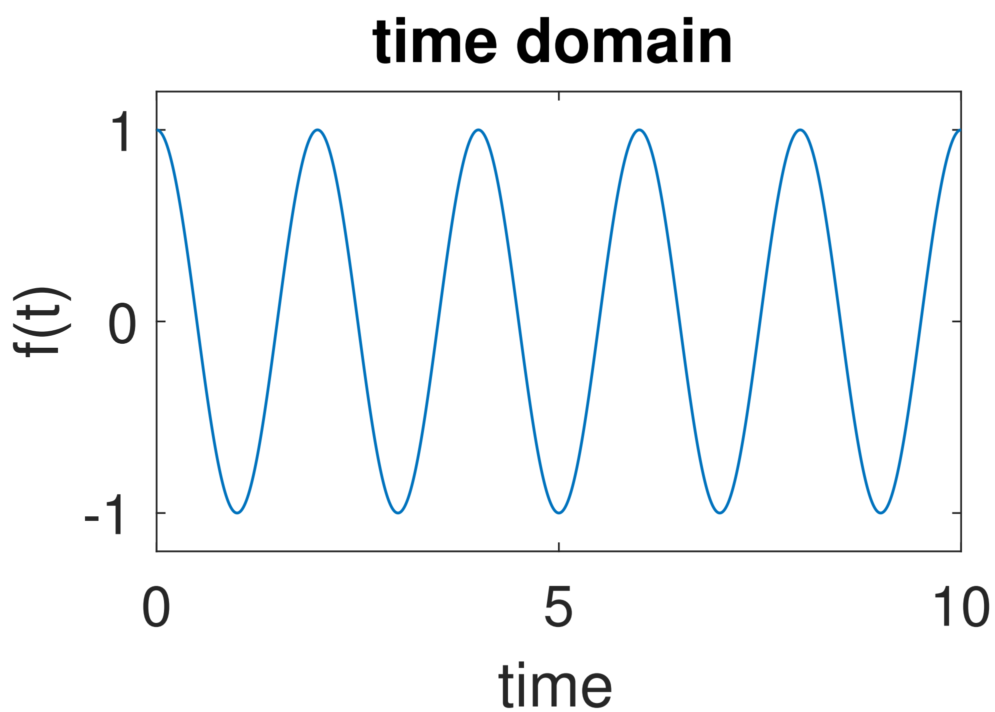
Sparsity and compression

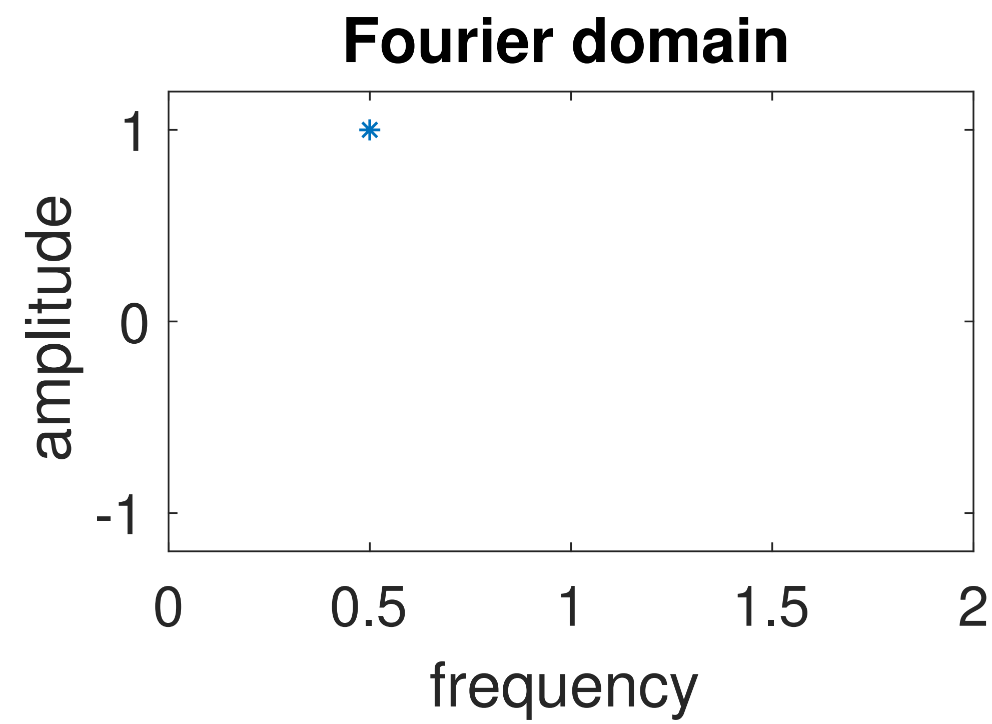
- By keeping a small fraction of the largest coefficients we can achieve high compression rates (JPEG2000 standard)

- The time series is **sparse** in the wavelet domain

Sparsity : example

- **Example** : sine wave



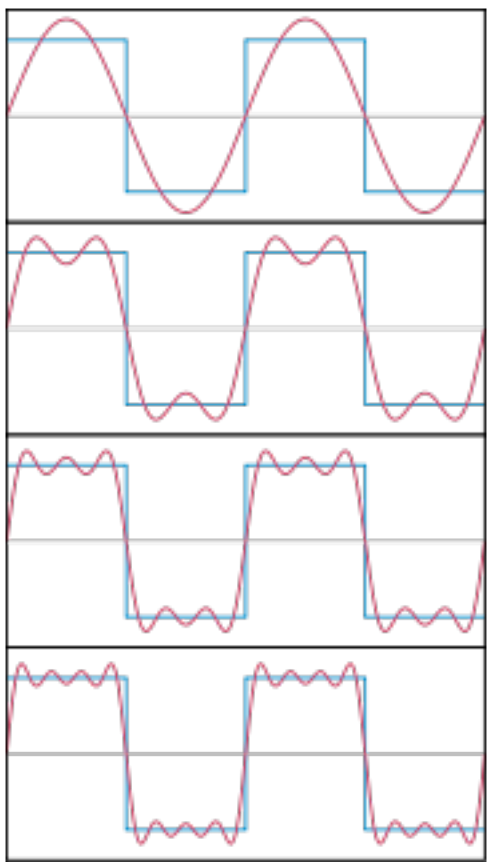


requires many data points	requires only 1 value
= NOT sparse	= VERY sparse

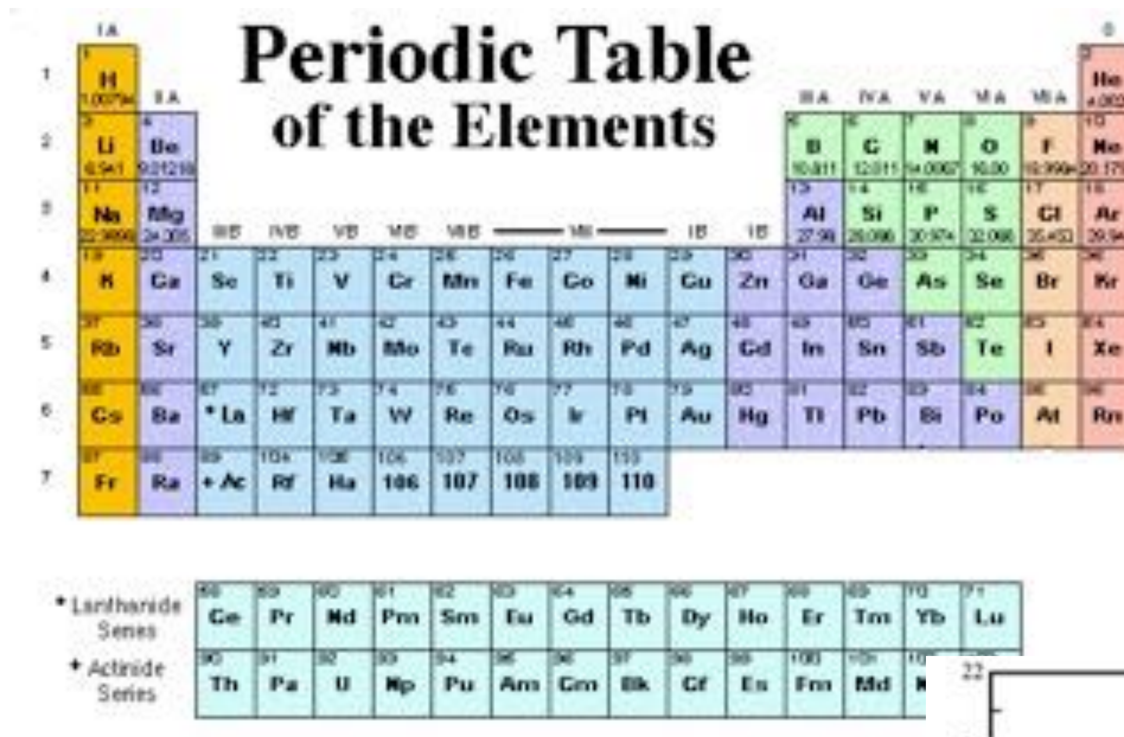
- Fourier representation is better suited here (does it provide physical invariants?)

Sparsity

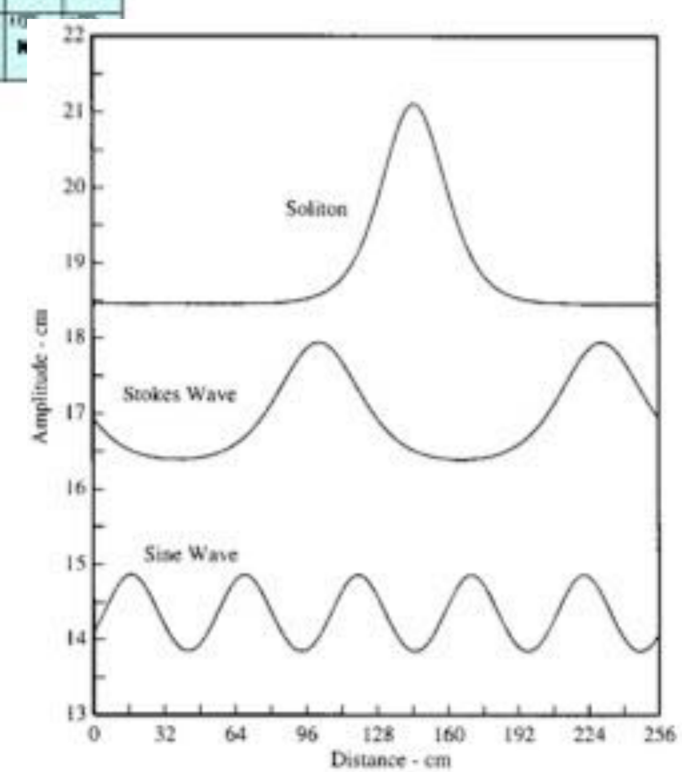
- **Sparsity** = find the representation (alphabet & grammar) that offers the most compact representation of the data



Fourier modes for linear wave phenomena



Nonlinear Fourier modes for the KdV equation



For the physicist to become

Sparsity

- What is the best alphabet / grammar ?
- Two possibilities

The alphabet/grammar are **UN**known
→ infer them from the data (empirical)

The alphabet/grammar are known
→ model them (Fourier modes, …)

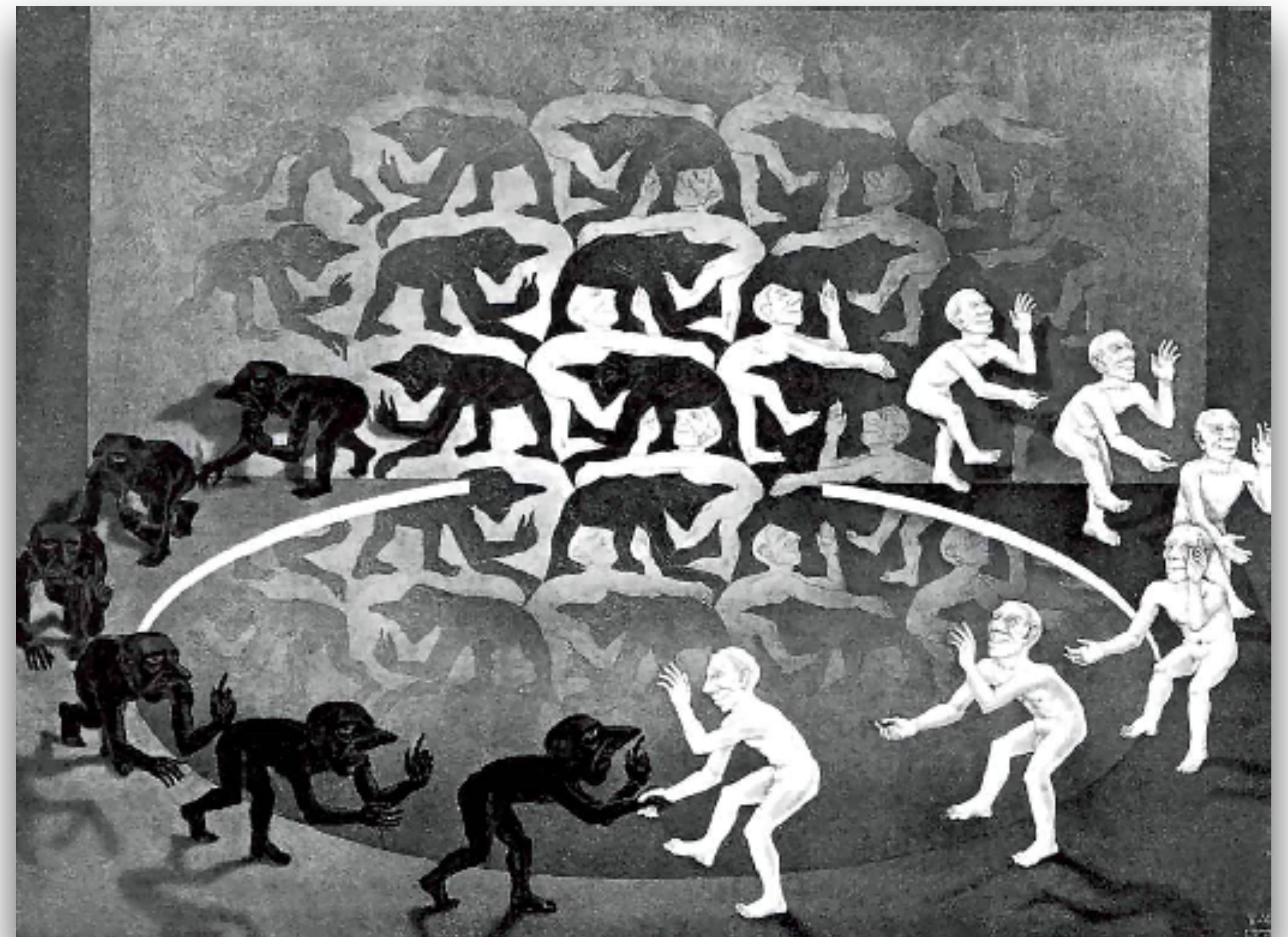
Blind source separation

Find the alphabet (sources) from a mixture with the least a priori information

“Cocktail party problem”

Assumptions

- Frequent assumptions
  - the mixture is *linear* $\rightarrow$ otherwise intractable
  - the mixture is *instantaneous* (non convolutive)$\rightarrow$ not always realistic, but eases the solution. Can be alleviated.
  - the sources are *sparse* $\rightarrow$ they are coherent in time, in space, or both

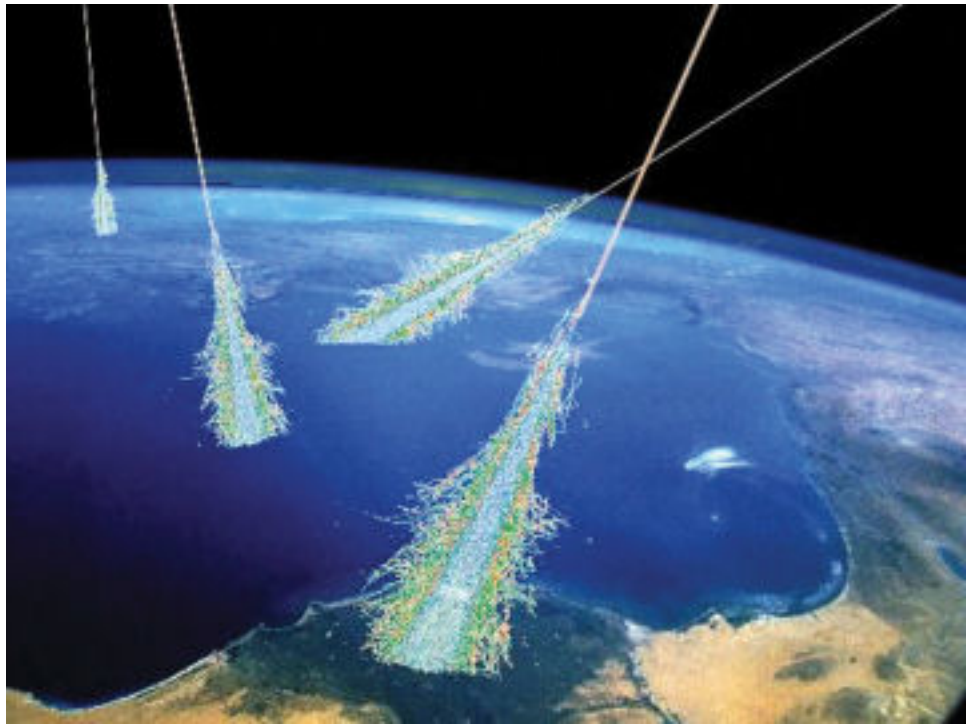


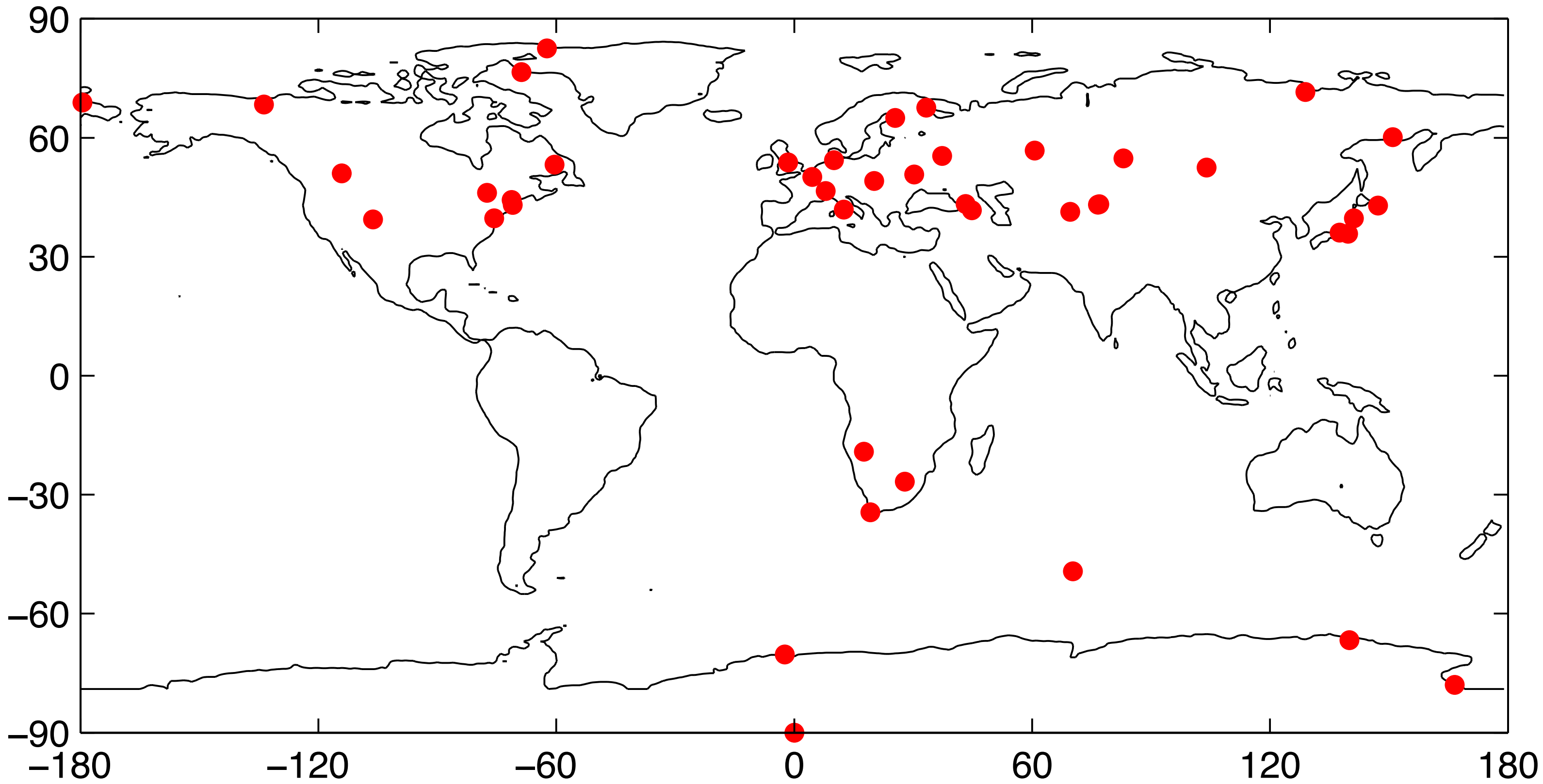
M. C. Escher

Example : cosmic radiation from neutron monitors

Neutron monitor data

- Neutrons are produced in the atmosphere by cosmic ray impacts
- They are a proxy for the cosmic ray flux & solar proton events

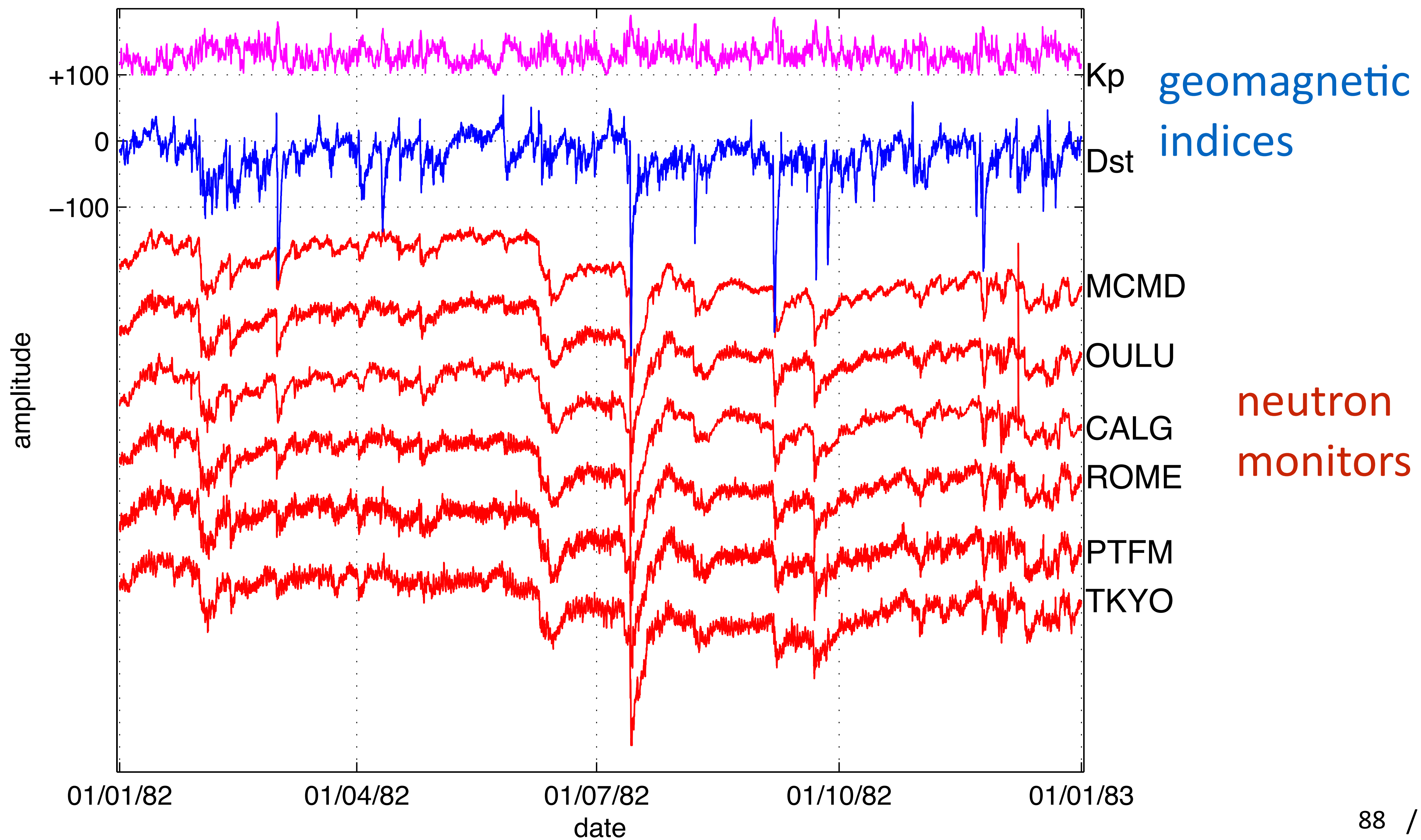




Network of neutron monitor stations

Neutron monitor data

All monitors see the same variations, with some subtle differences



Neutron monitor data

What the physics is telling us

The observed flux is a **linear superposition** of different contributions that are ~ independent

- anisotropy of the cosmic flux
- modulation by the solar magnetic field (Forbush decreases)
- modulation by the geomagnetic field (hardness)

Assumptions

- Consider a separable solution is separable

$$\Phi(x,t) = \sum_k V_k(t) S_k(x)$$

neutron flux

temporal modulation

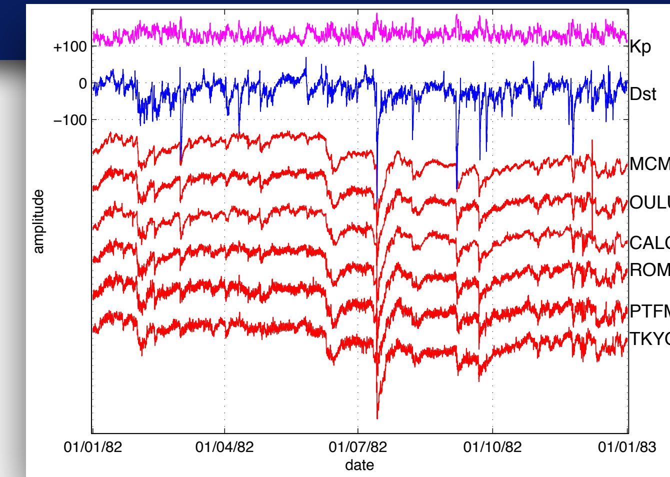
spatial signature

- The different modulation amplitudes $V_k(t)$ are independent $\Longrightarrow$ use Independent component analysis (ICA)

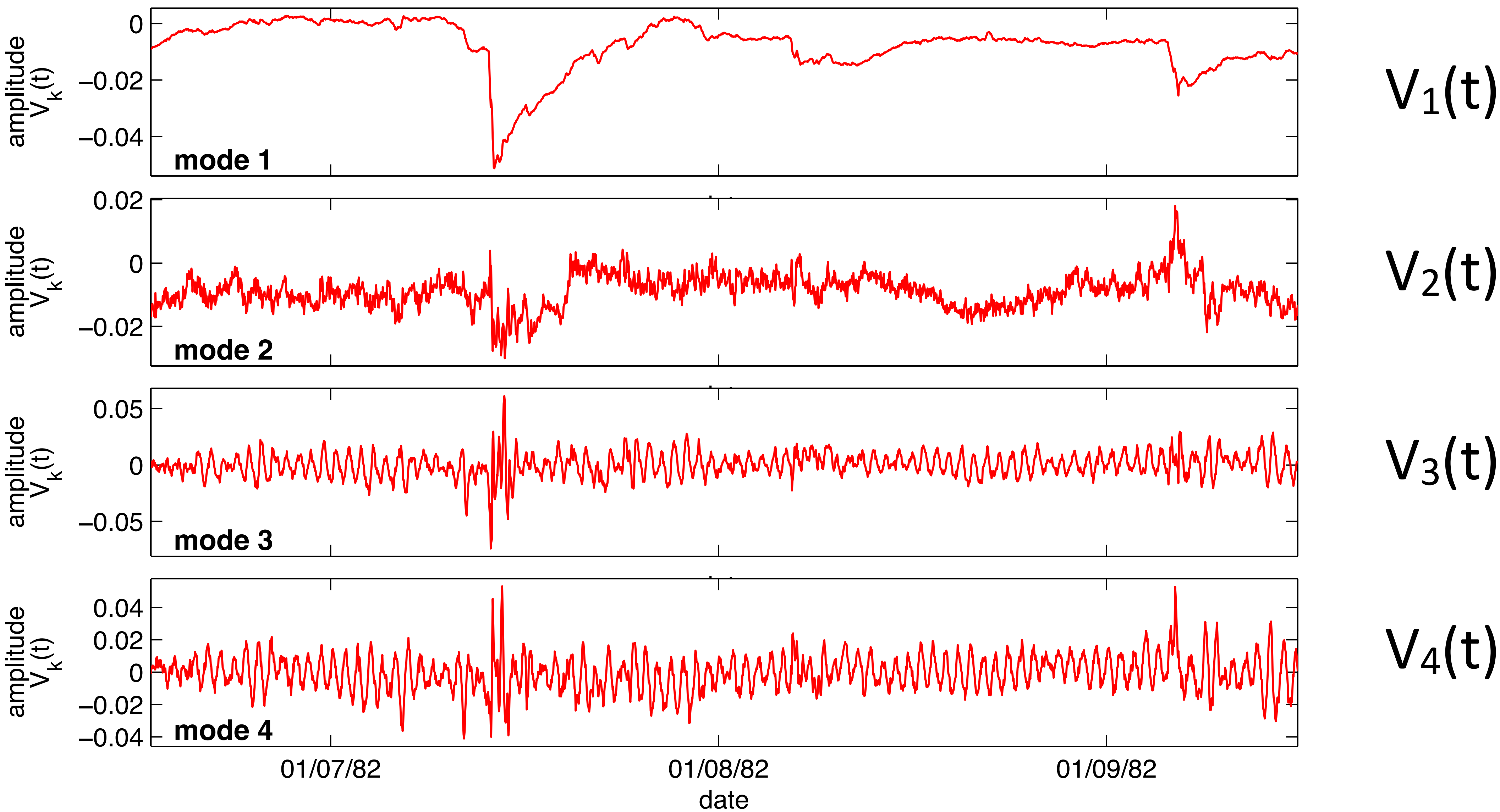
$$\mathcal{P}\left(V_k(t), V_l(t)\right) = \mathcal{P}\left(V_k(t)\right) \cdot \mathcal{P}\left(V_l(t)\right)$$

probability distribution

Results



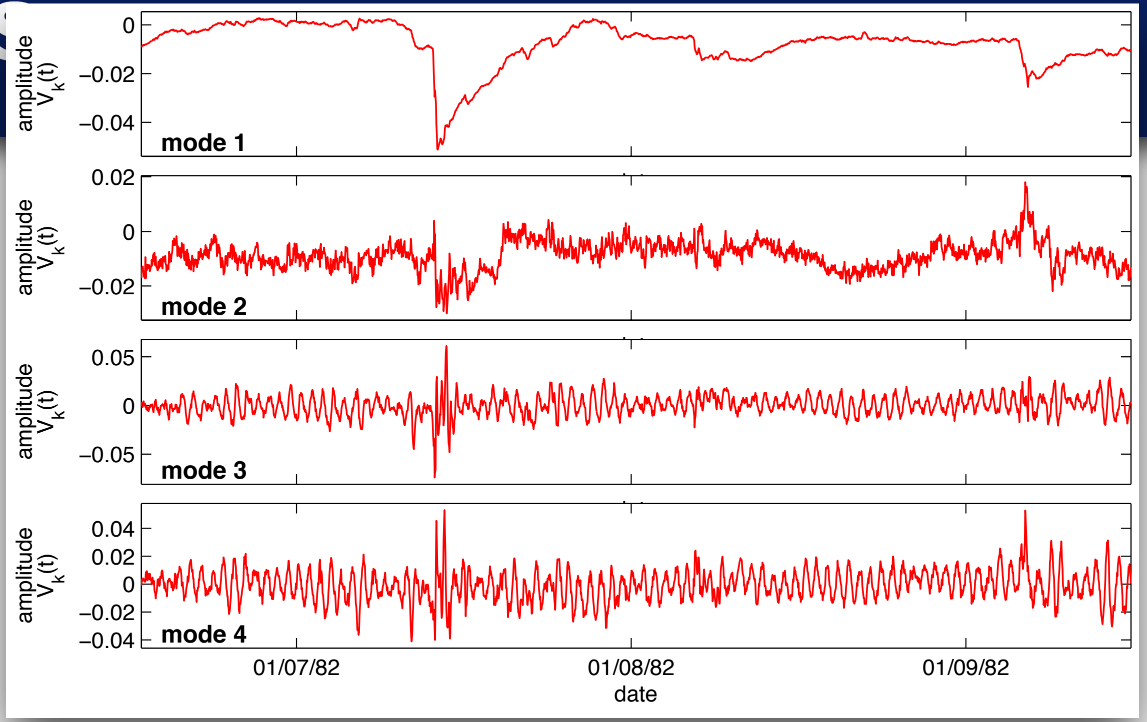
- 98.4 % of the variance can be reconstructed with just 4 modes out of 43



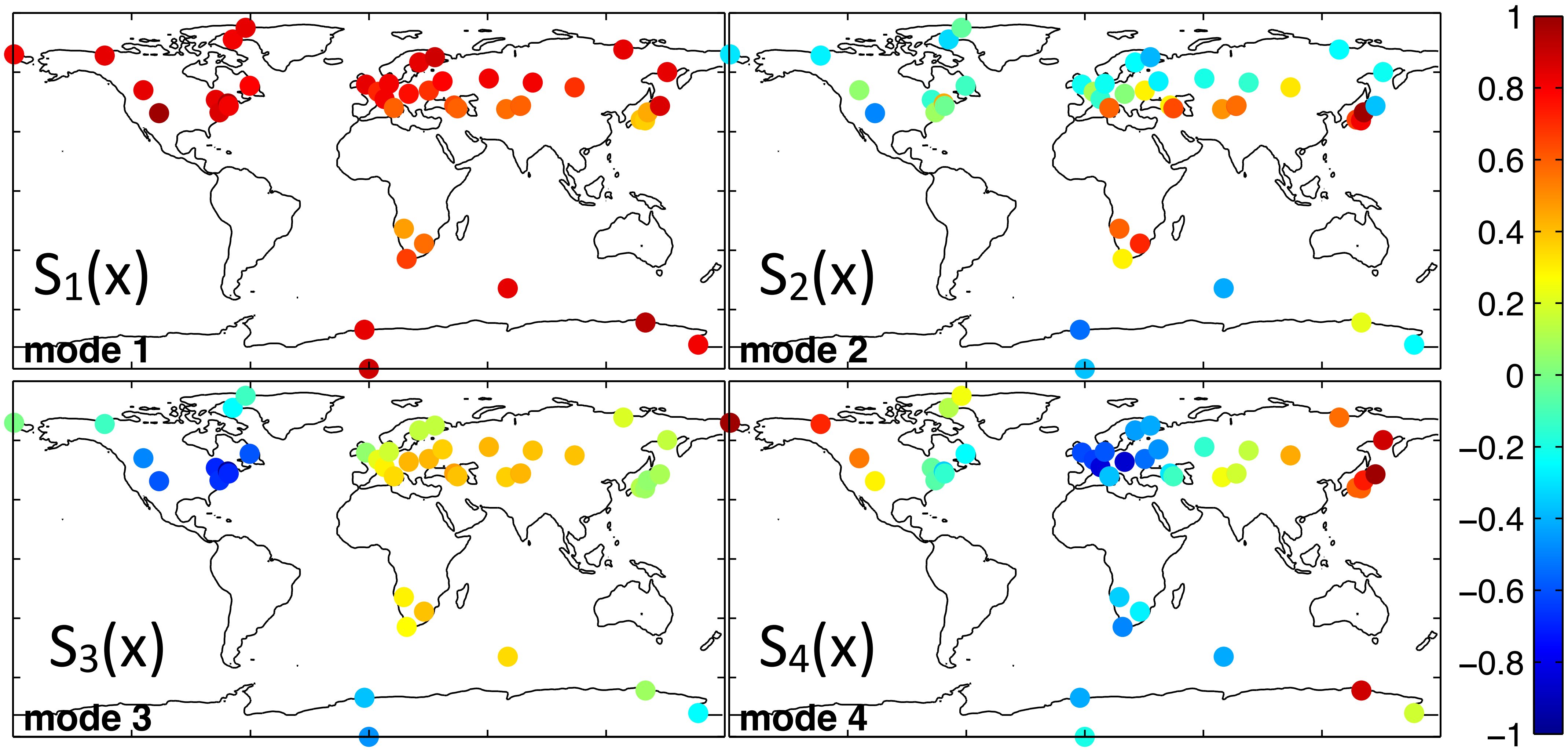
Interpretation

This is purely statistical (empirical) but it suggests that specific physical processes are at play

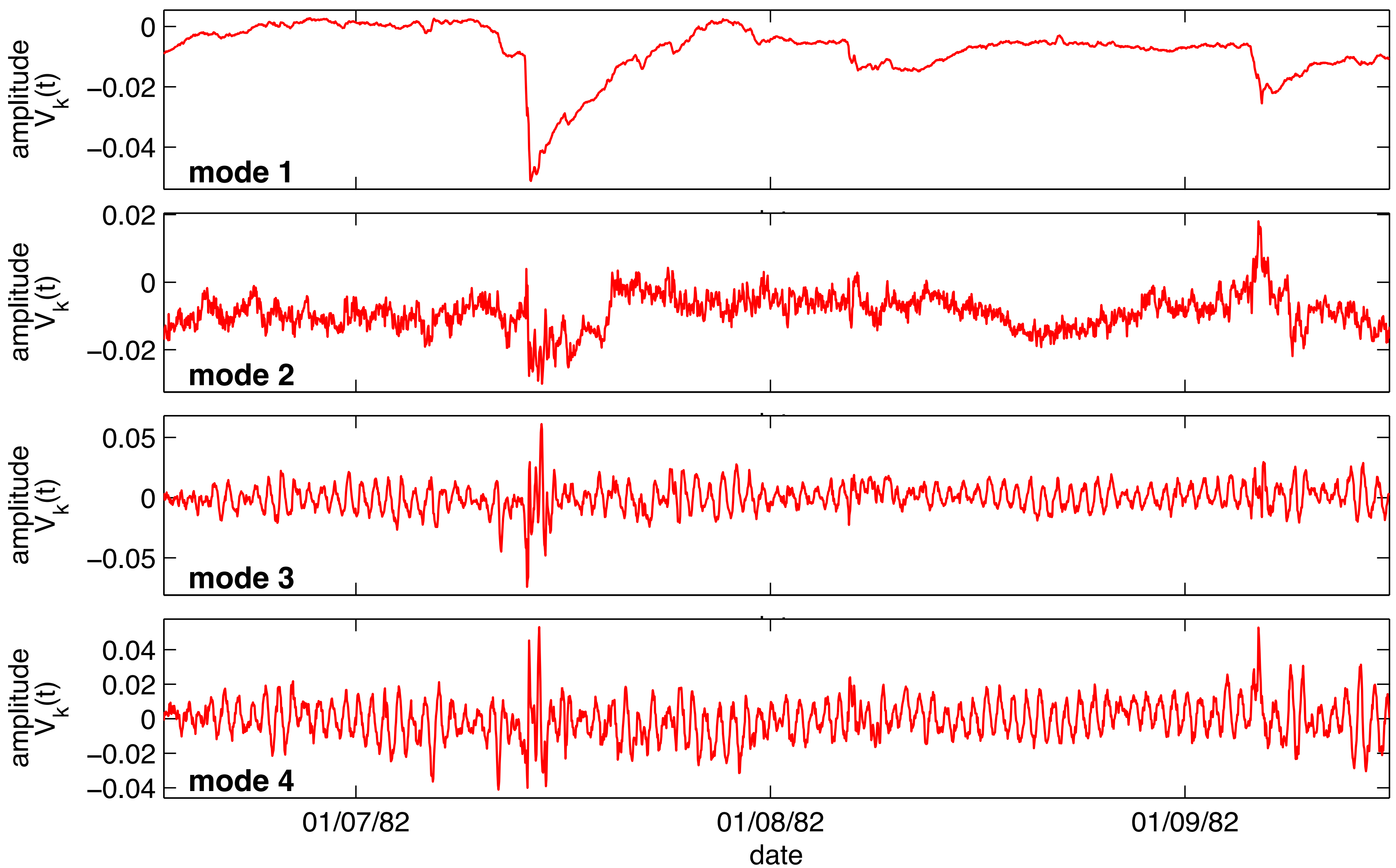
Results



- Spatial footprint of the modes



Physical interpretation



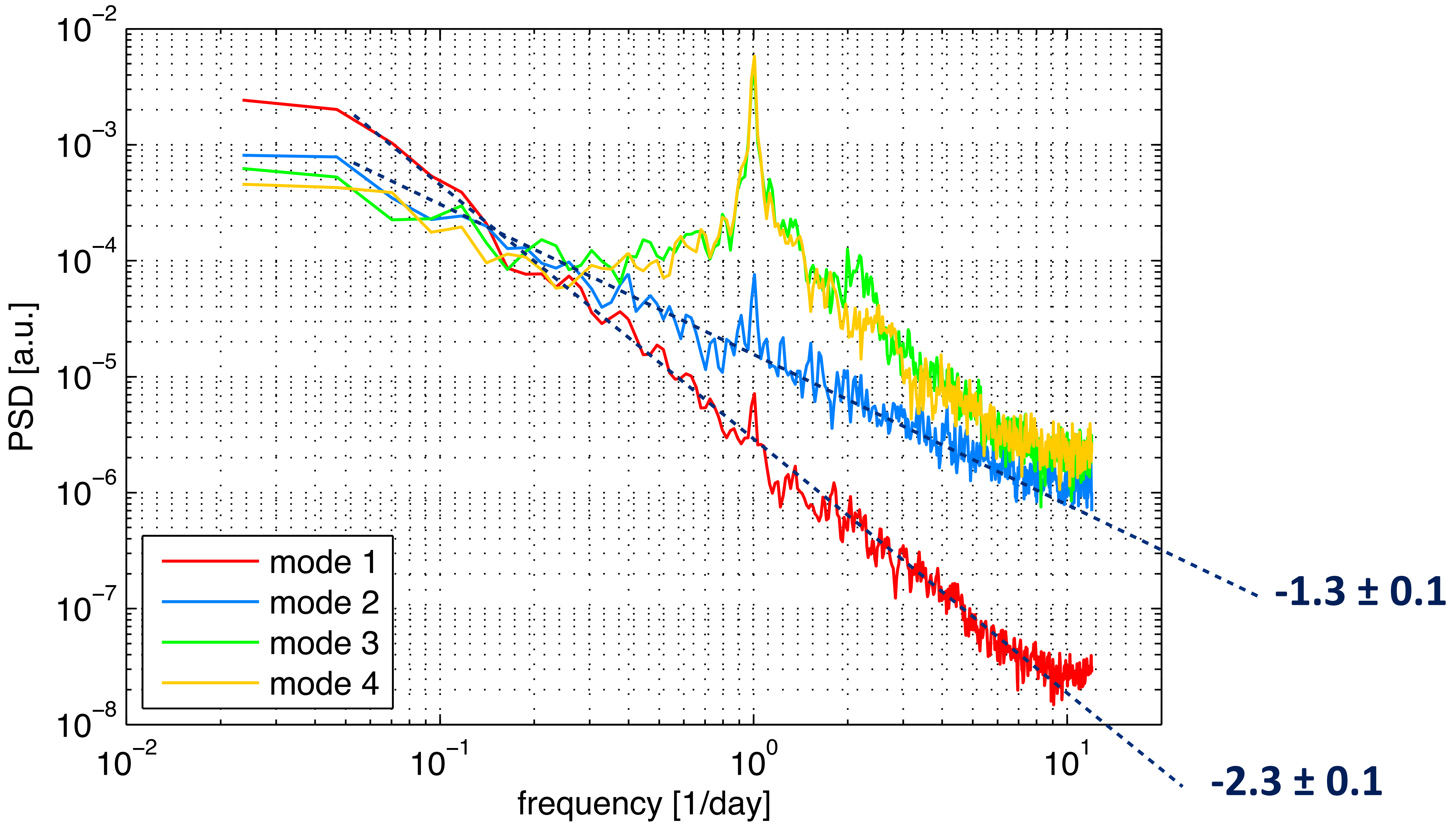
Isotropic cosmic ray flux (+ Forbush decreases)

Rigidity changes due to geomagnetic field

Anisotropic contribution

Physical interpretation

- Power spectral density reveals differing origins



M. C. Escher

Inpainting

Inpainting

- **Inpainting** = reconstruct lost information in images or time series

- Use sparsity & wavelets to recover that information from local context (intensity, texture, patterns, ...)

Inpainting : example

Inpainting is an ancient art

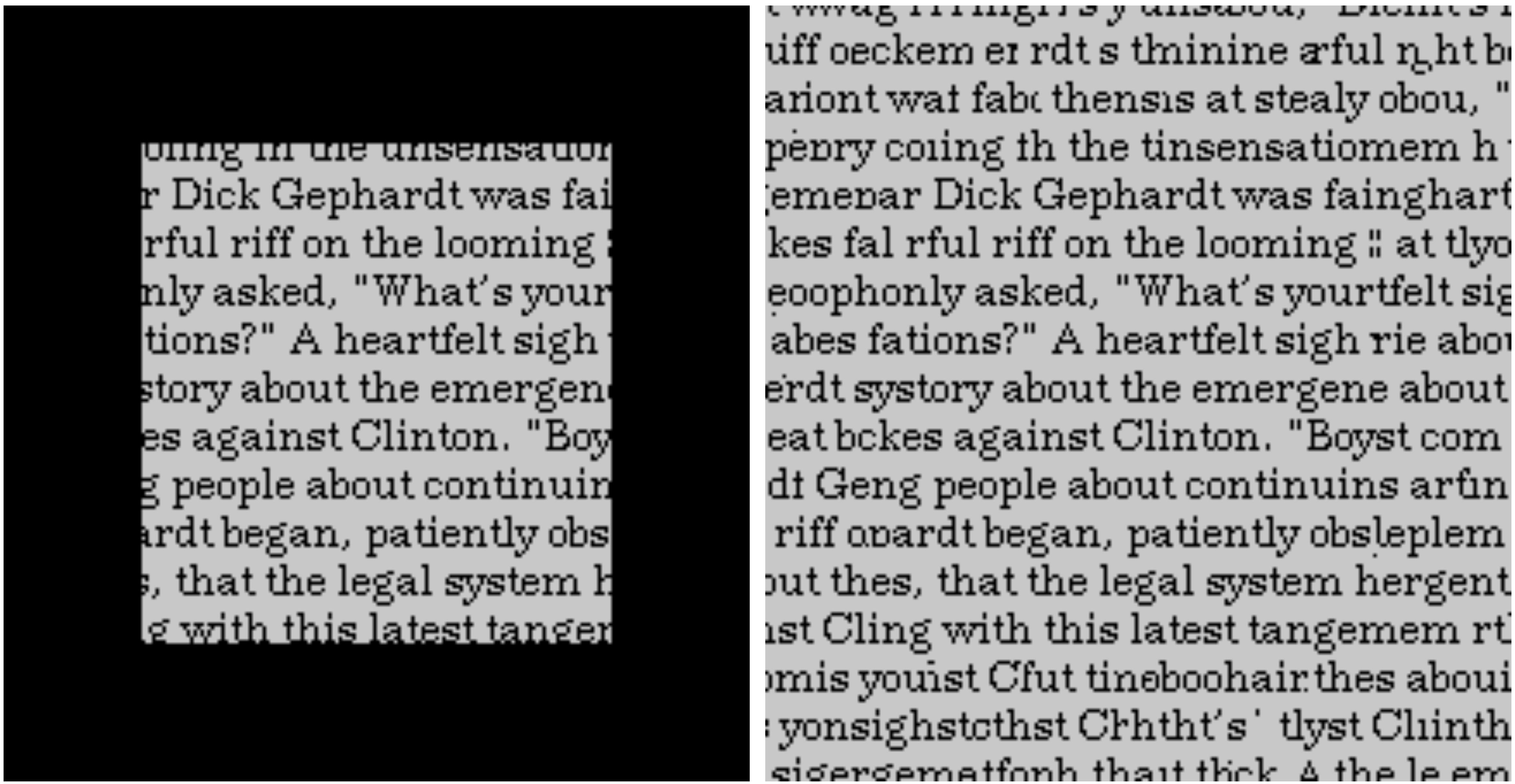
Staline with/without Nicolaï Yezhov

Inpainting : example

Criminisi et al., 2004

Inpainting : example

- Another example with texture preservation

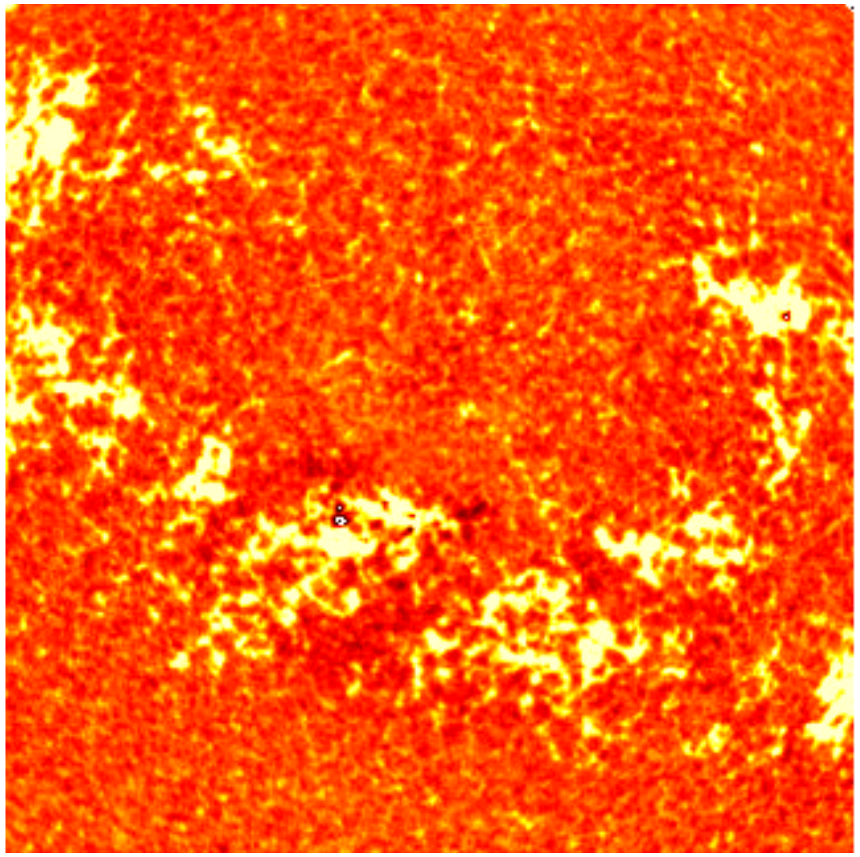


Efros & Leung, 1999

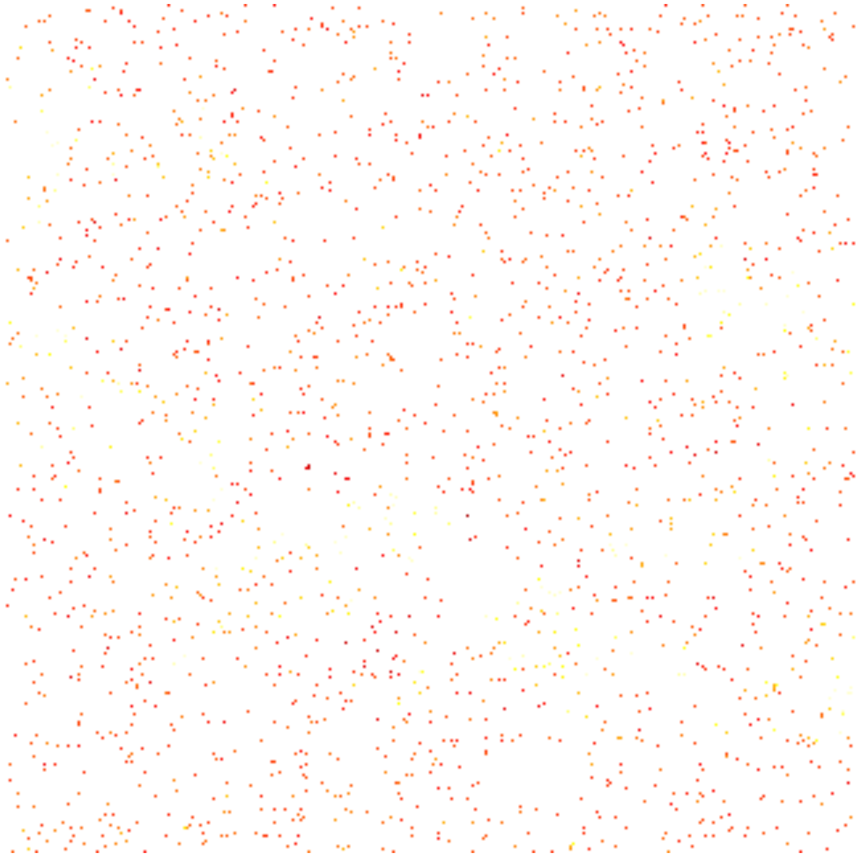
Inpainting : example

- Solar image taken in the Ca II K line (293 nm) : 98% of pixels were removed and then filled by inpainting

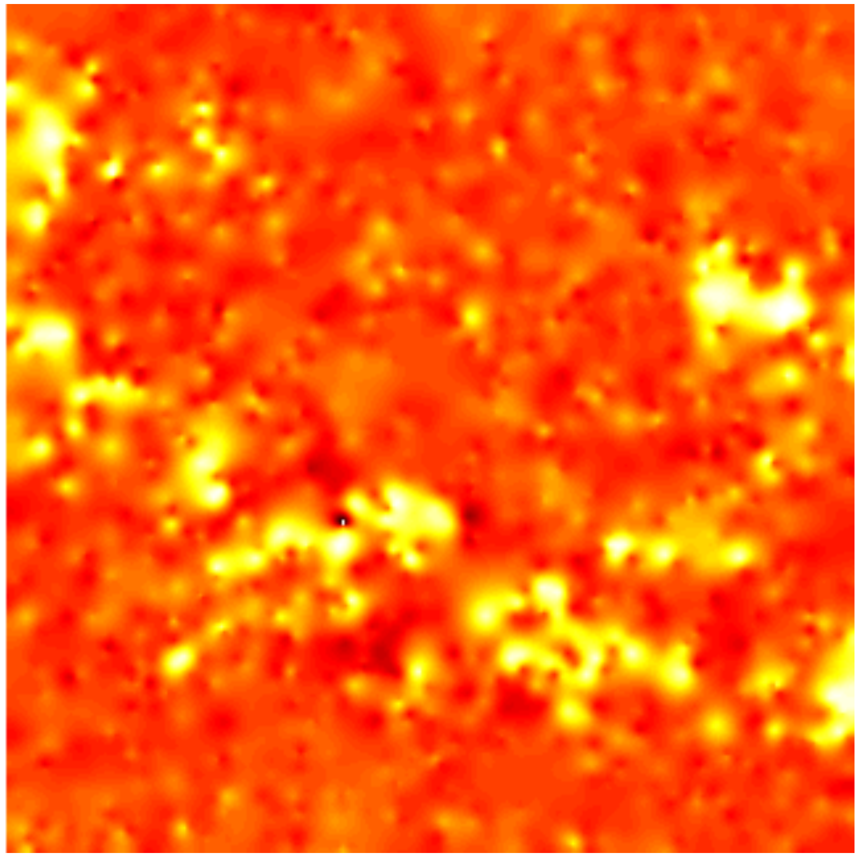
original image



only 2% of pixels left



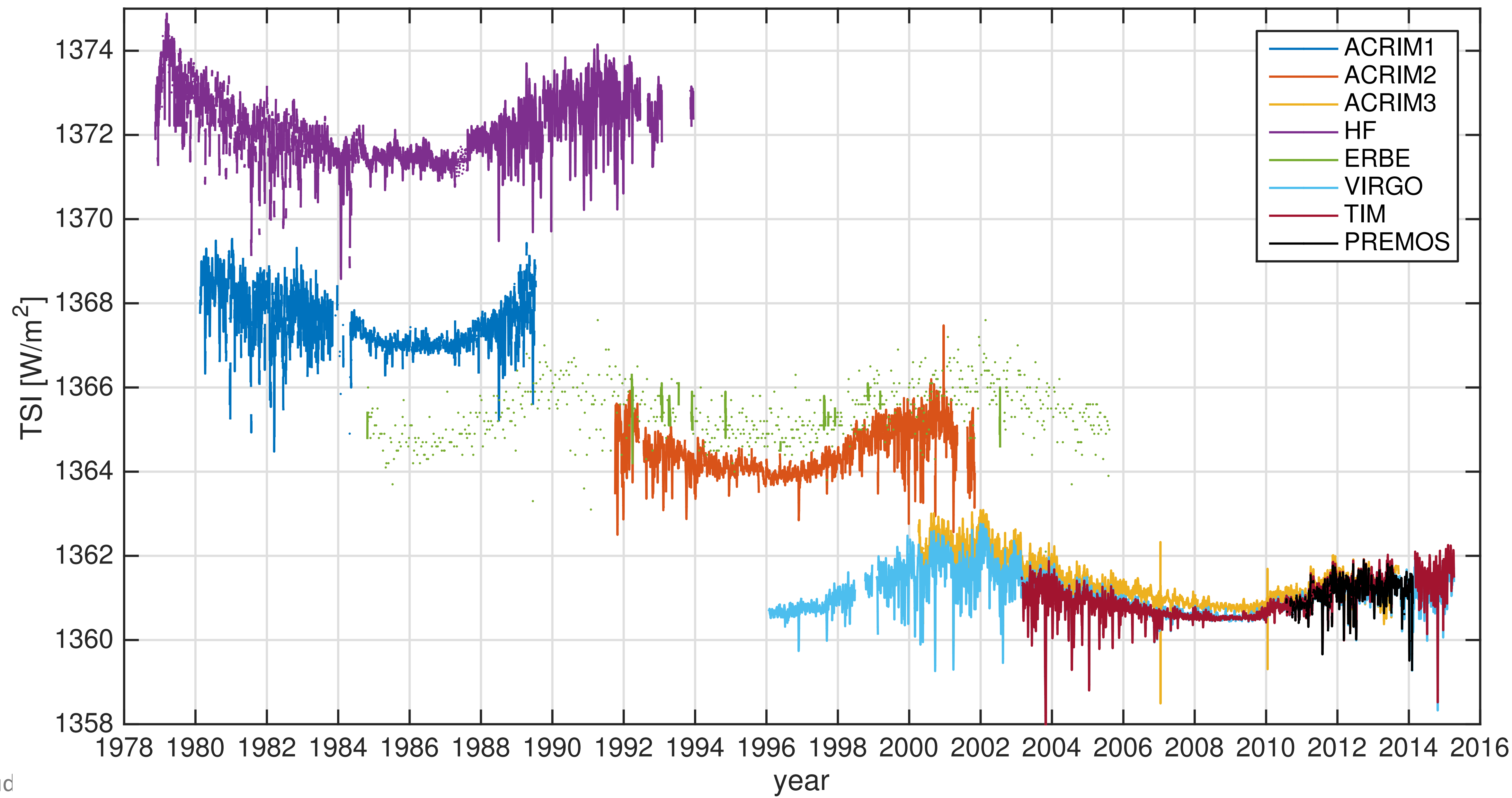
reconstructed image



Inpainting : example

- Inpainting can also be applied to time series, to fill in data gaps

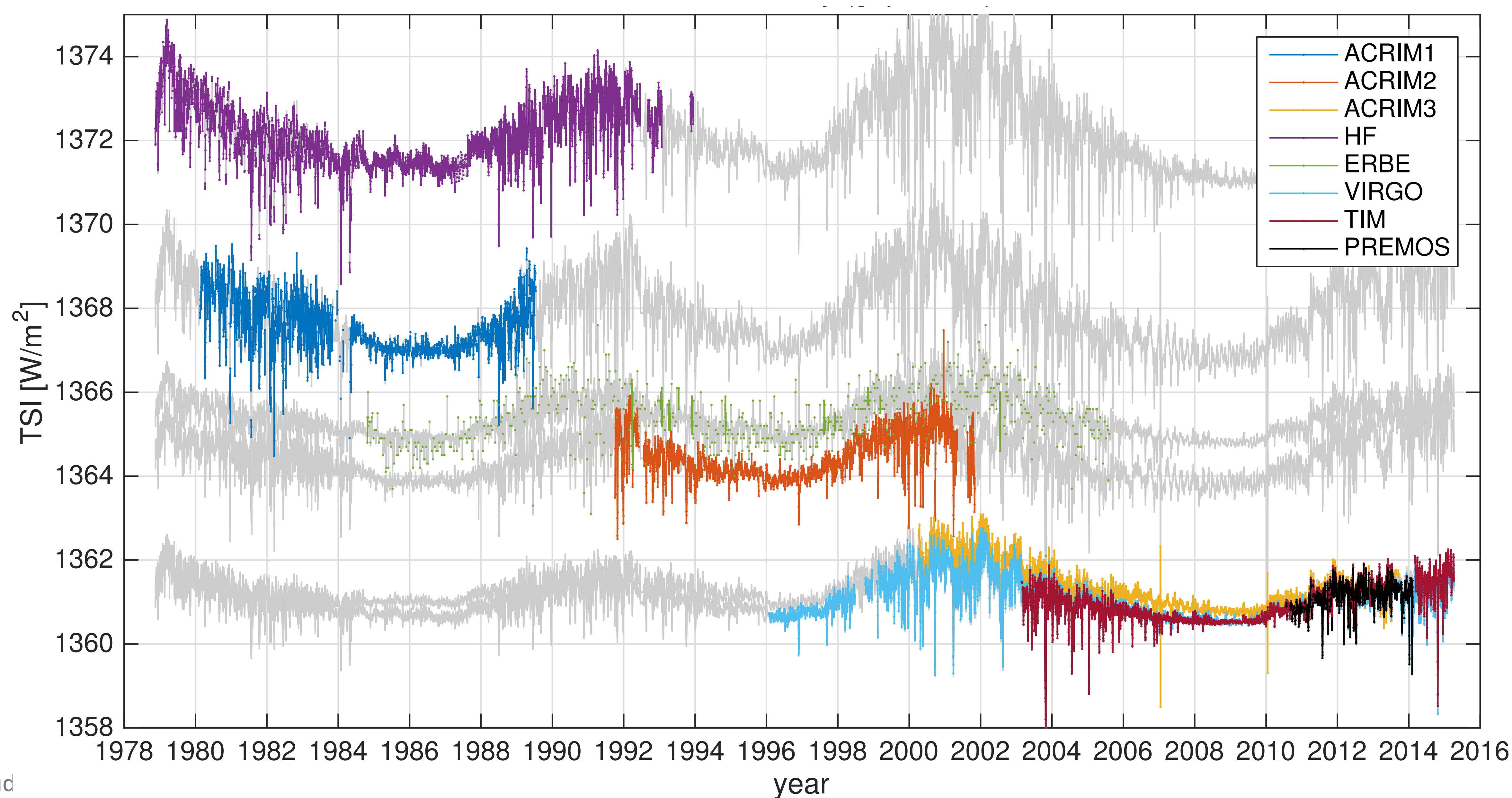
Example : total solar irradiance observations (from space)



Inpainting : example

- Inpainting can also be applied to time series, to fill in data gaps

Example : total solar irradiance observations (from space)



Conclusion

M. C. Escher

Take home message

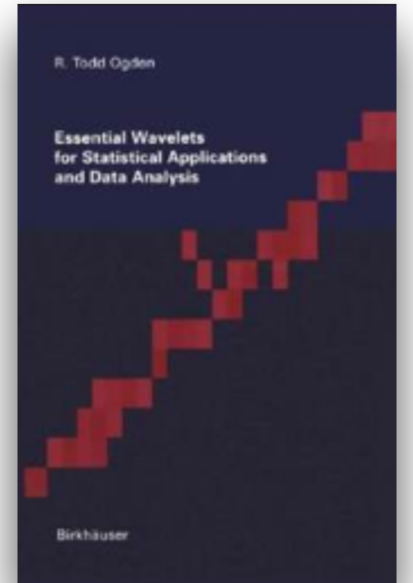
- Think multiscale (because plasmas are instrinsically nonstationary)

- Don't use a method just because everyone has always done so before.

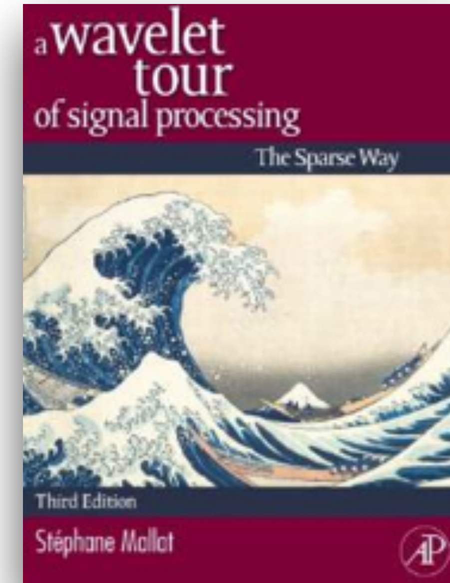
"We use the best physical models we can, the best computers for processing data (...) But there is one weak link. We interpret the data using mathematics that's 100 years old."

(Dana McKenzie, New Scientist, 2004)

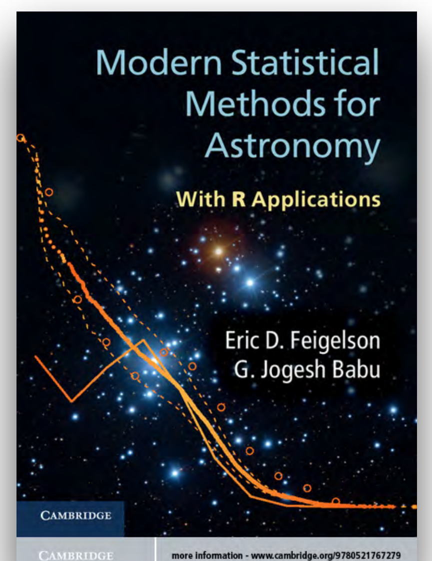
Further reading



T. Ogden
Essential wavelets for statistical applications and data processing
(Academic Press, 1996)

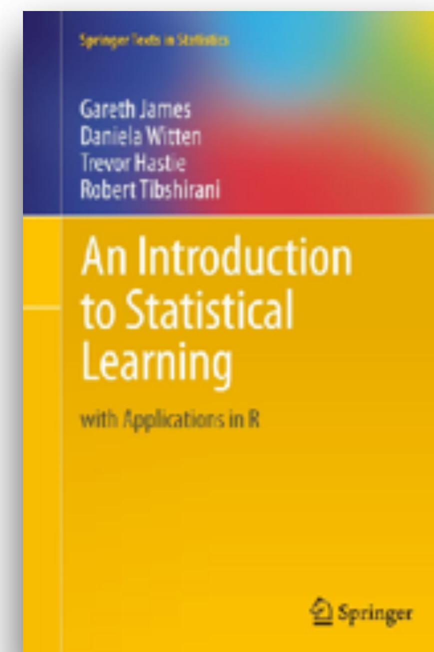


S. Mallat
A wavelet tour of signal processing
(Academic Press, 2008)



E. Feigelson & G. Babu

Modern statistical methods for astronomy
(Cambridge, 2012)



G. James, et al., *An Introduction to Statistical Learning: with Applications in R*
(Springer, 2013).

Collection of articles at: https://tinyurl.com/les-houches-2017