Image denoising

Some approaches

Forthem Alliance / Team Teaching - April 2023

olivier.laligant@u-bourgogne.fr

in collaboration with

liliana.lopresti@unipa.it

Contents

- About noise
- Denoising
  - PDE
  - Bilateral filtering
  - Non local means
  - Wawelet shrinkage

About noise

Noise definition

- Real signal = information & noise
- noise
  - inside the phenomenon : instability (chemical reaction, turbulence, smoke, vibrations, …)
  - introduced by the observation (measure device) : mainly electronics in sensors
- information can be random signal (with correlation)
- noise can be deterministic signal (ex : vibration)

needs for noise definition before addressing the denoising problem

Noise definition

- Natural noise (generally not wished)
- Synthetic noise
  - to simulate real/natural noise in synthetic (or low noise) data
  - to test and measure performances of the algorithm
  - to put in light lacks of stability of the algorithm
  - to analyze the performances in respect of different noises

Noise definition

- Model
  - defining : real signal = information signal & noise signal
  - noise signal : generally random variable, characterized by distribution, moments, etc.
  - the way this noise signal disturbs the information signal : additive noise, multiplicative noise, salt and pepper noise, etc.
  - whatever the way the noise signal disturbs the information signal, we can write :

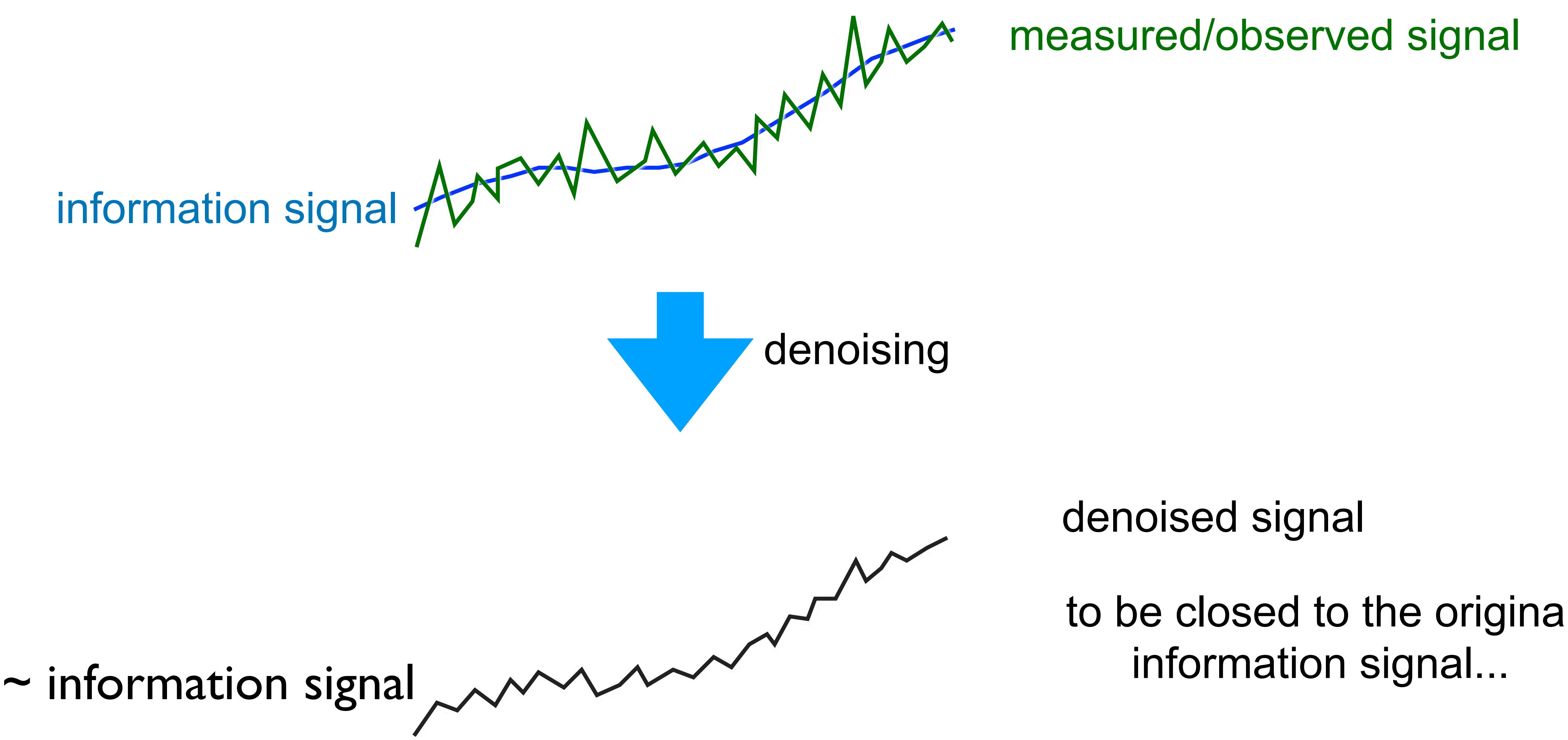
$$y = x + n$$

y : real signal
x : information signal
n : noise signal (= noise signal in the additive case)

real signal = information signal + noise signal

Denoising

real signal = information signal + noise signal

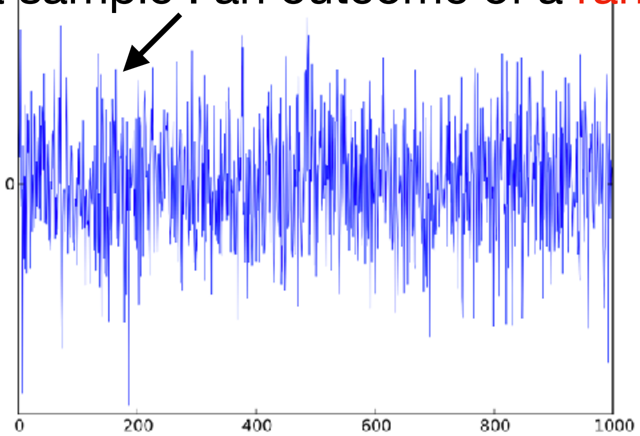


final result : a trade-off between a priori assumptions and a posteriori results

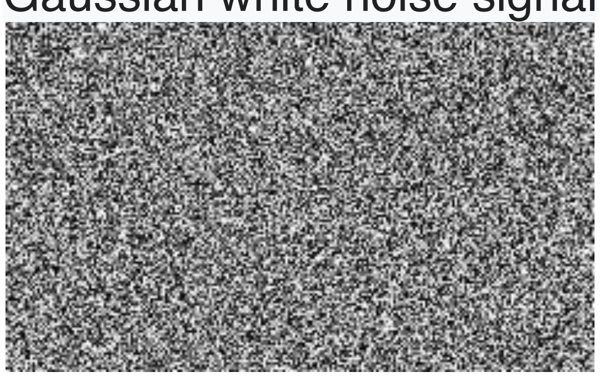
Chosen synthetic noise

- Chosen model : gaussian white noise - additive « contribution » $y = x + n$
  - Realization of a Random Process
  - noise : a discrete signal whose samples are regarded as a sequence of serially uncorrelated random variables (random process) with zero mean and finite variance
  - Gaussian distribution
  - White Power Spectral Density (PSD : Fourier Transform of the Autocorrelation)

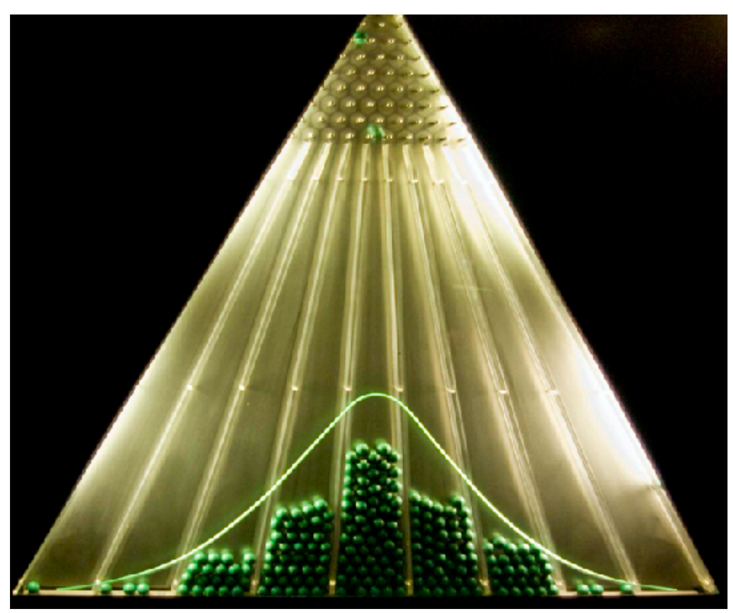
a sample : an outcome of a random variable



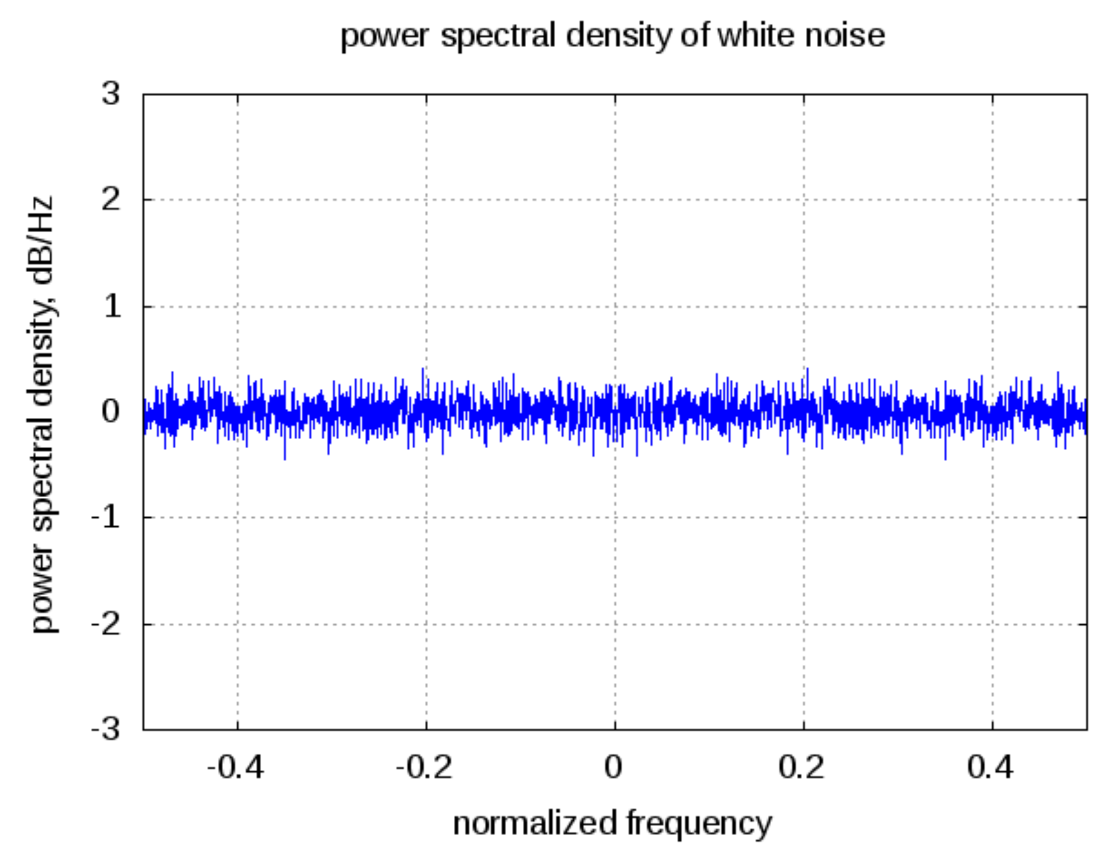
Gaussian white noise signal



white noise image



The bean machine : the first generator of Gaussian (normal) random variables



source : wikipedia

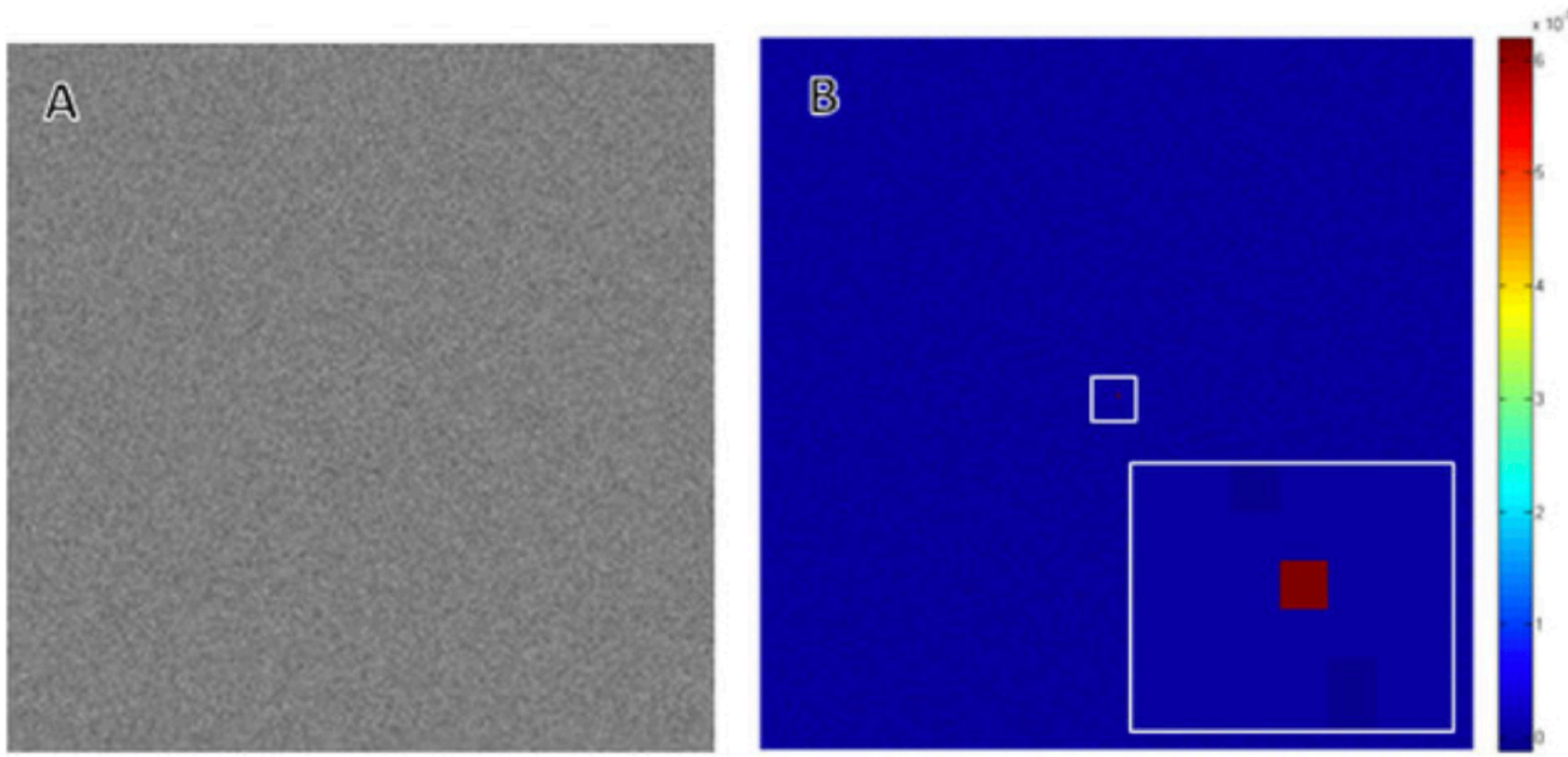
Distribution

- Uniform distribution
  - you roll the dice 100 times
  - you observe approximately a flat distribution of the different faces of the dice

- Gaussian distribution
  - you roll the dice 100 times
  - you count the number of times the dice shows for example 6
  - you repeat this experiment a lot of times
  - you plot the number of times the 6 is shown 1time, 2 times, etc.. The max will be around 50 times. The plots follow a gaussian function.

About random/stochastic

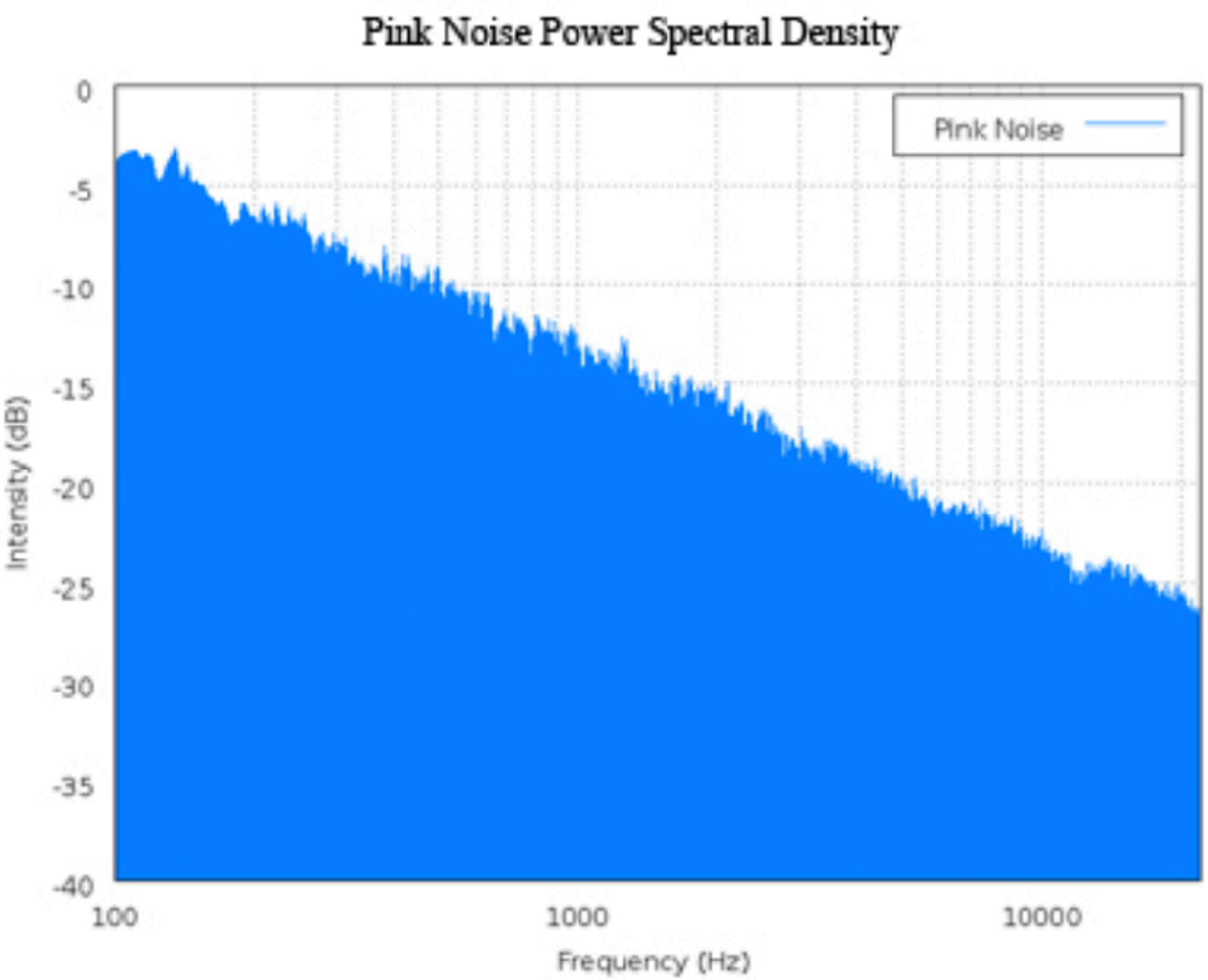
- Random = stochastic.
  uses : stochastic process, random variable, random signal
- White means no correlation : each sample is independent on the others

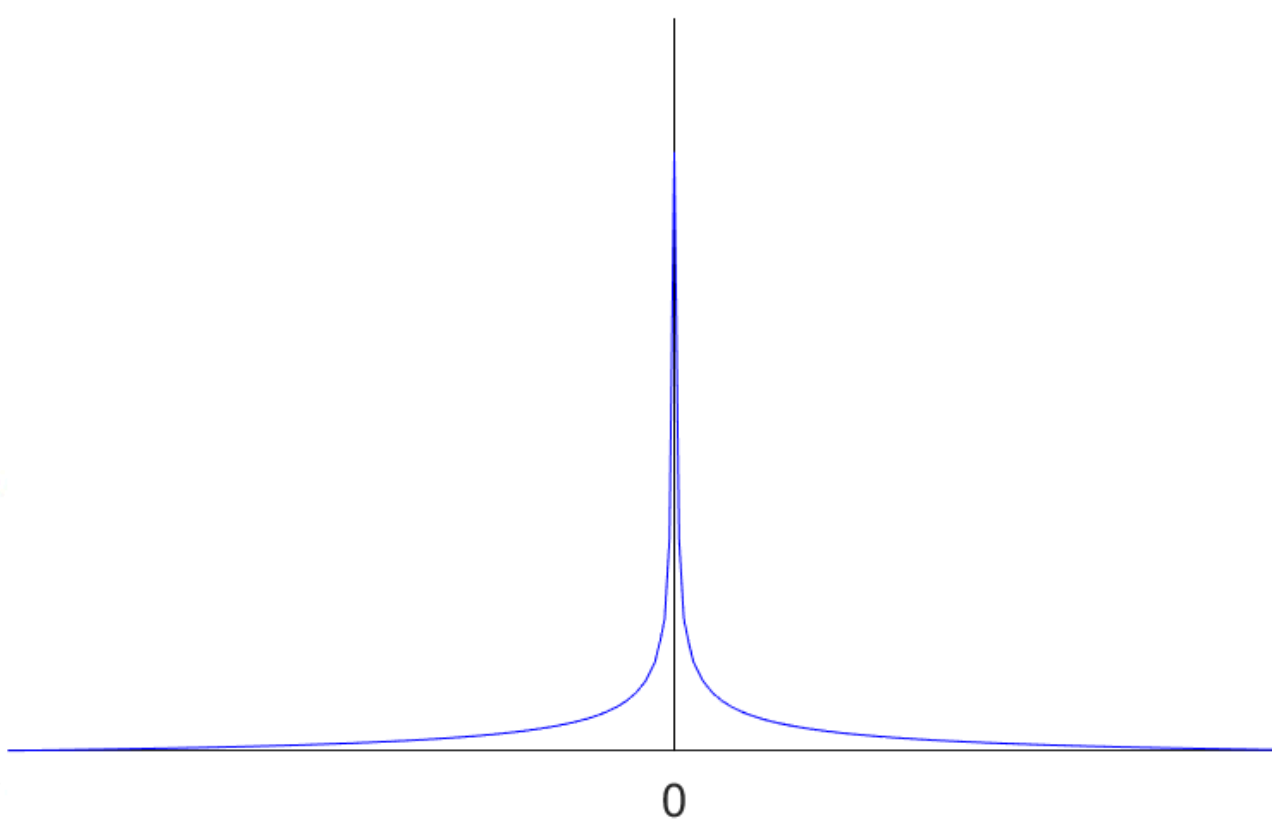


white noise image　　autocorrelation image

About random/stochastic

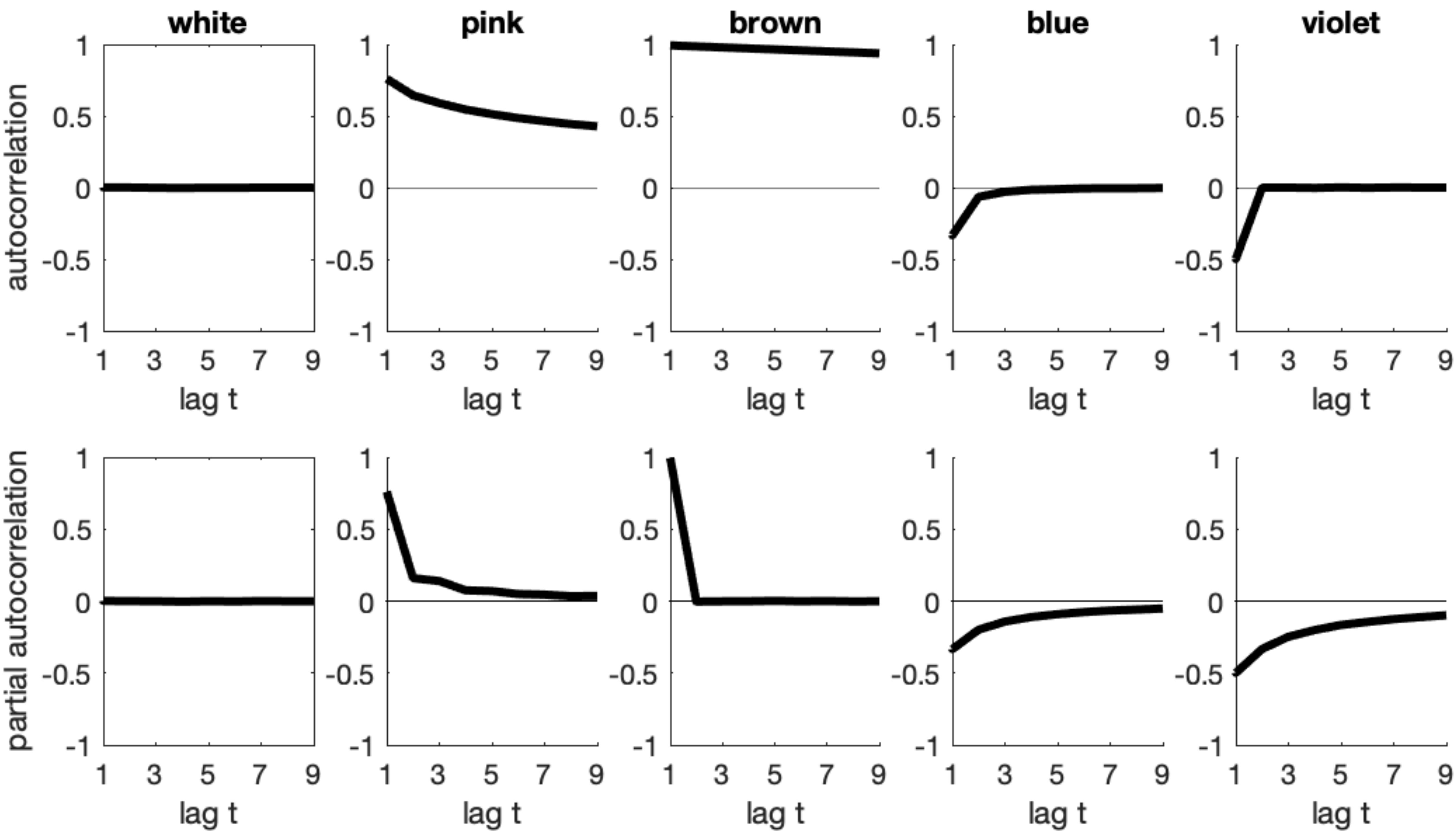
- Pink, red, blue, … noises contain correlation and then are not totally random





About random/stochastic

- Pink, red, blue, … noises contain correlation and then are not totally random



https://stats.stackexchange.com/questions/485449/are-colored-noises-correlated-uncorrelated

Denoising

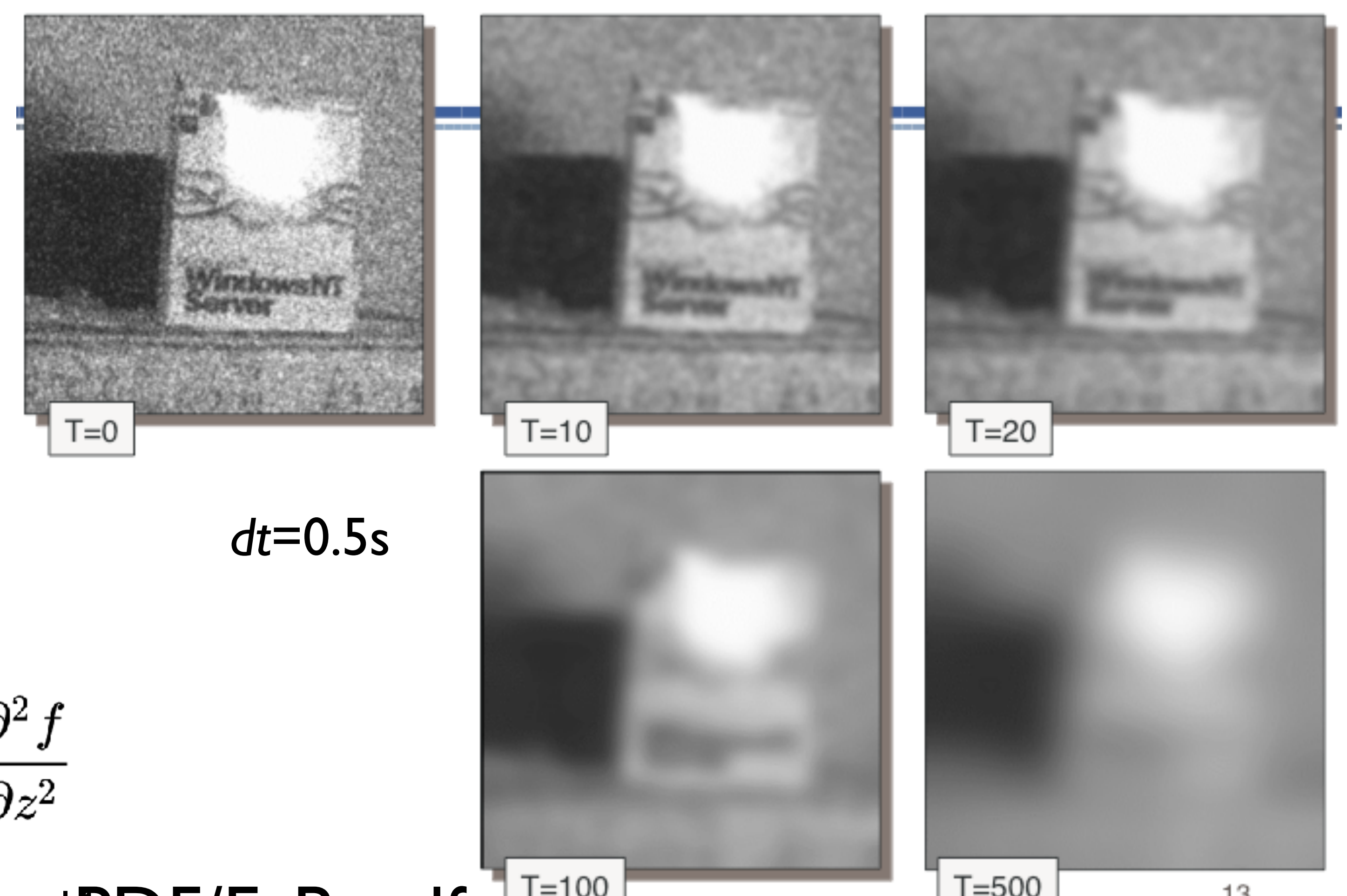
PDE (Partial Differential Equations)

- Principle : heat equation

$$\begin{cases} \frac{\partial I}{\partial t} & = & \Delta I \\ I(x,y,0) & = & I_0(x,y) \end{cases}$$

- Solution : convolution of the image with the Gaussian function
  - discrete derivative(s) gives the edge points

the worst equation ! (isotropic smoothing)



dt=0.5s

$$\Delta f = \overrightarrow{\nabla}^2 f = \overrightarrow{\nabla} \cdot (\overrightarrow{\nabla} f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

http://www.lgi2p.ema.fr/~montesin/Cours/PDF/EcPr.pdf

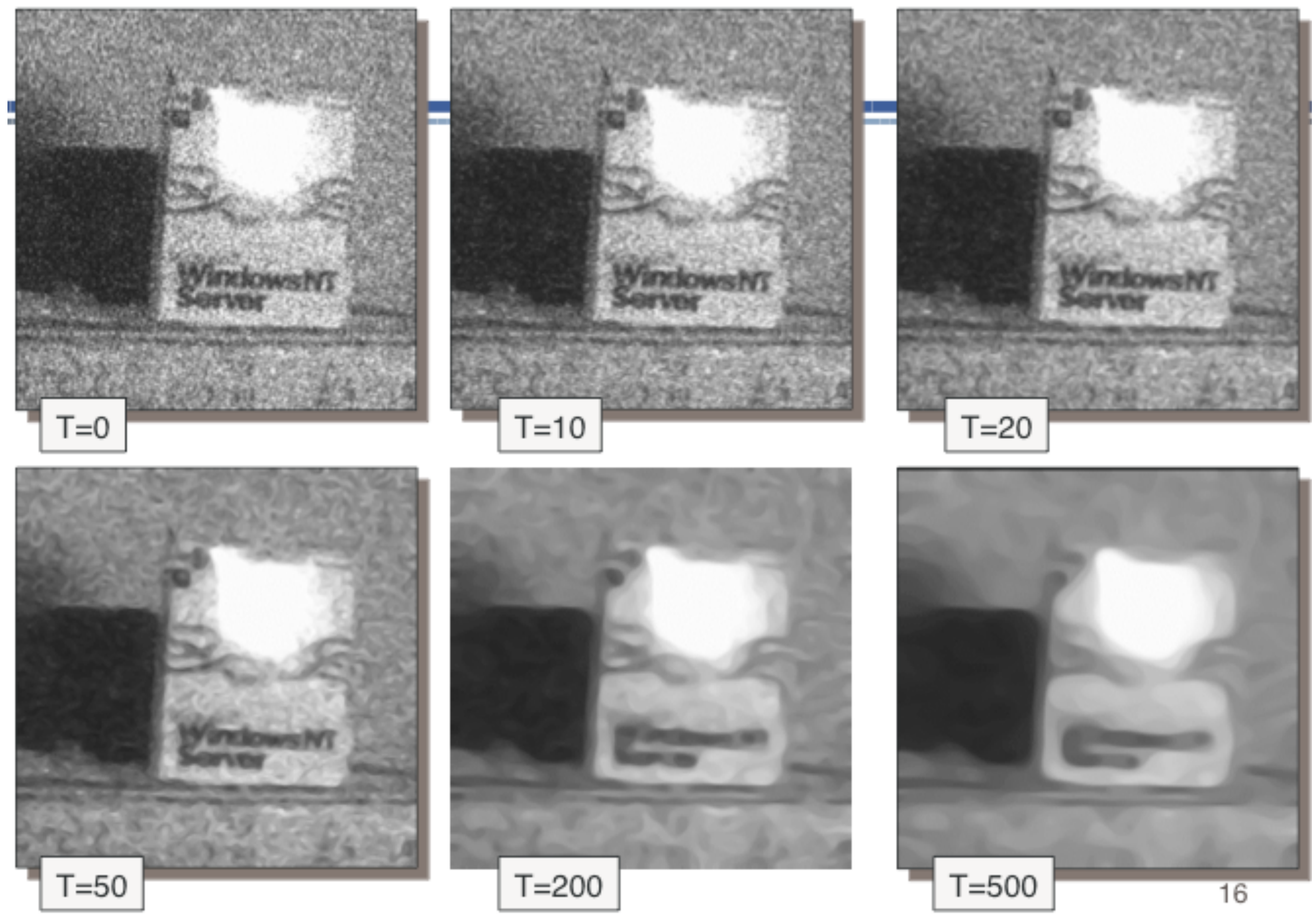
PDE

- Anisotropic smoothing

$$\begin{cases} \frac{\partial I}{\partial t} & = & t|\nabla I| div\left(\frac{\nabla I}{|\nabla I|}\right) \\ I(x,y,0) & = & I_0(x,y) \end{cases}$$

$$div\left(\frac{\nabla I}{|\nabla I|}\right) = \frac{I_y^2 I_{xx} - 2I_x I_y I_{xy} + I_x^2 I_{yy}}{|\nabla I|^3}$$

isotropic smoothing on isophotes*



$$\vec{\nabla} = \frac{\partial}{\partial x}\vec{\mathbf{i}} + \frac{\partial}{\partial y}\vec{\mathbf{j}} + \frac{\partial}{\partial z}\vec{\mathbf{k}}$$

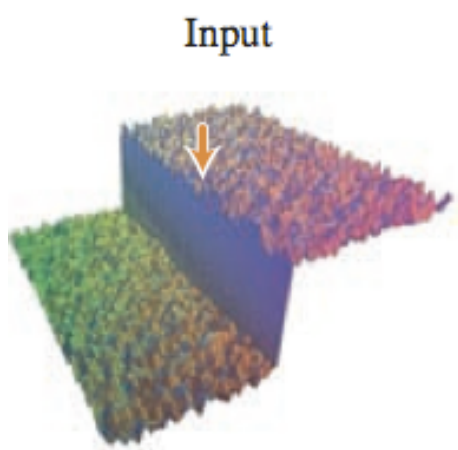
$$\operatorname{div}\mathbf{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}$$

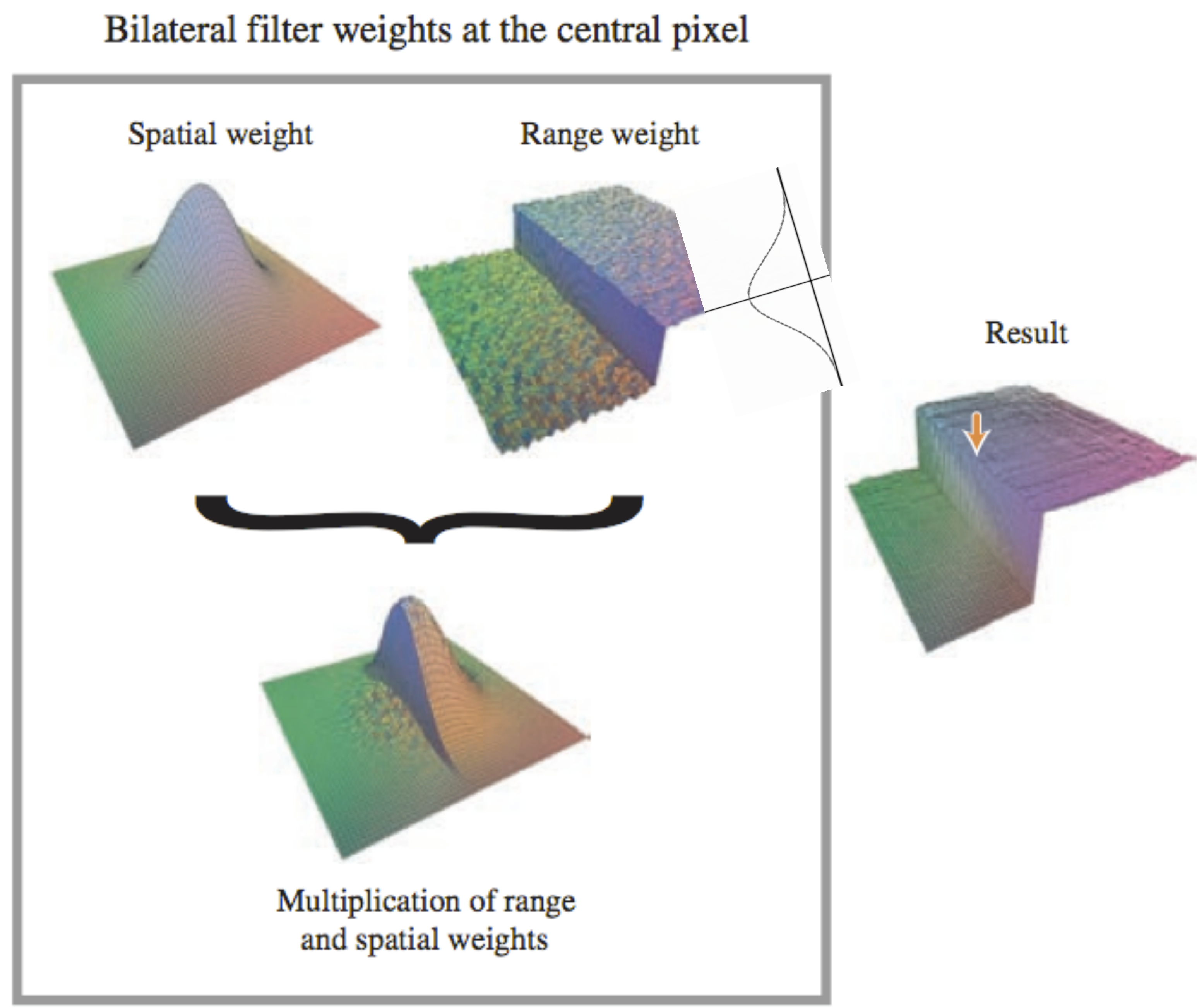
* isophote : a curve on a chart joining points of equal light intensity from a given source

http://www.lgi2p.ema.fr/~montesin/CoursPDF/EcPr.pdf

Bilateral filtering

- Principle :
  - bilateral filtering preserves a given range of edges





usually, g_l and h_n are the gaussian function

Bilateral filtering

- Definition
  - Regularized image :

$$I_r(\mathbf{x}) = \mathcal{N}^{-1}(\mathbf{x}) \sum_{\mathbf{x}_\omega \in \omega} I(\mathbf{x}_\omega) h_n(||\mathbf{x} - \mathbf{x}_\omega||) g_l(I(\mathbf{x}) - I(\mathbf{x}_\omega))$$

where :

$$\mathcal{N}(\mathbf{x}) = \sum_{\mathbf{x}_\omega \in \omega} h_n(||\mathbf{x} - \mathbf{x}_\omega||) g_l(I(\mathbf{x}) - I(\mathbf{x}_\omega))$$

  - This corresponds to a convolution of the image I by a low pass-filter (h_n) associated to a limited range of luminance (g_l)
  - $\mathcal{N}(x)$ is a normalization factor depending on the current point $\mathbf{x}$
  - ω is the neighorhood required for the application of h_n

C. Tomasi and R. Manduchi, "Bilateral Filtering for Gray and Color Images", *Proceedings of the 1998 IEEE International Conference on Computer Vision*, Bombay, India.

Bilateral filtering

- Example

(a) Input image

(b) Output of the bilateral filter

Bilateral filtering

original

result

Non local means (NL-means)

- Principle
  - consider a point p (to denoise) in a given neighborhhod
  - we search for similar neighborhoods in the whole image
  - the similarity (w) in based on the luminance
    - neighborhoods of $q1$ and $q2$ are similar (strong w) to the one of p
  - the new value (denoised) of p is the mean of the points having similarity

Antoni Buades, Bartomeu Coll, and Jean-Michel Morel, *Non-Local Means Denoising*, Image Processing On Line, vol. 2011.

Non local means (NL-means)

- Definition
  - the new $v_{NL}(\mathbf{x})$ value of the image I at the position $\mathbf{x}$ is:

$$v_{NL}(\mathbf{x}) = \sum_{\mathbf{y} \in I} w(\mathbf{x}, \mathbf{y}) v(\mathbf{y})$$ contribution of the neighborhood

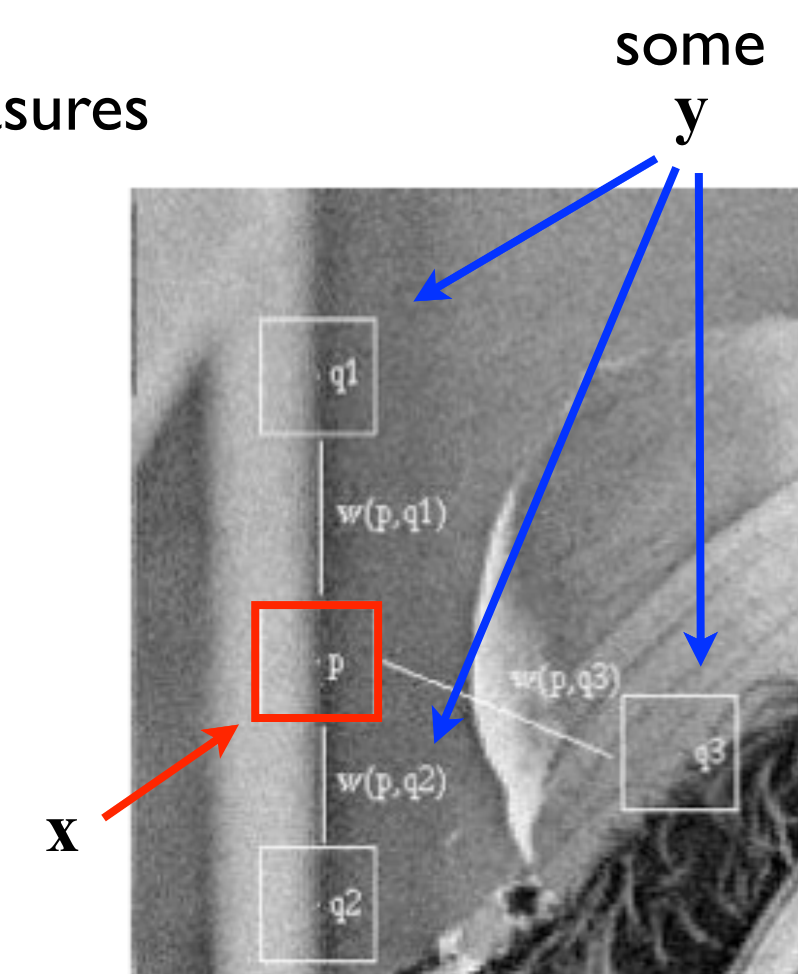
with :

$$w(\mathbf{x}, \mathbf{y}) = Z(\mathbf{x})^{-1} e^{-\frac{1}{h^2}||v(N_{\mathbf{x}}) - v(N_{\mathbf{y}})||^2}$$ similarity measure

where $$Z(\mathbf{x}) = \sum_{\mathbf{y}} e^{-\frac{1}{h^2}||v(N_{\mathbf{x}}) - v(N_{\mathbf{y}})||^2}$$ all measures

N_{x} and N_{y} are the neighborhoods around x and y respectively

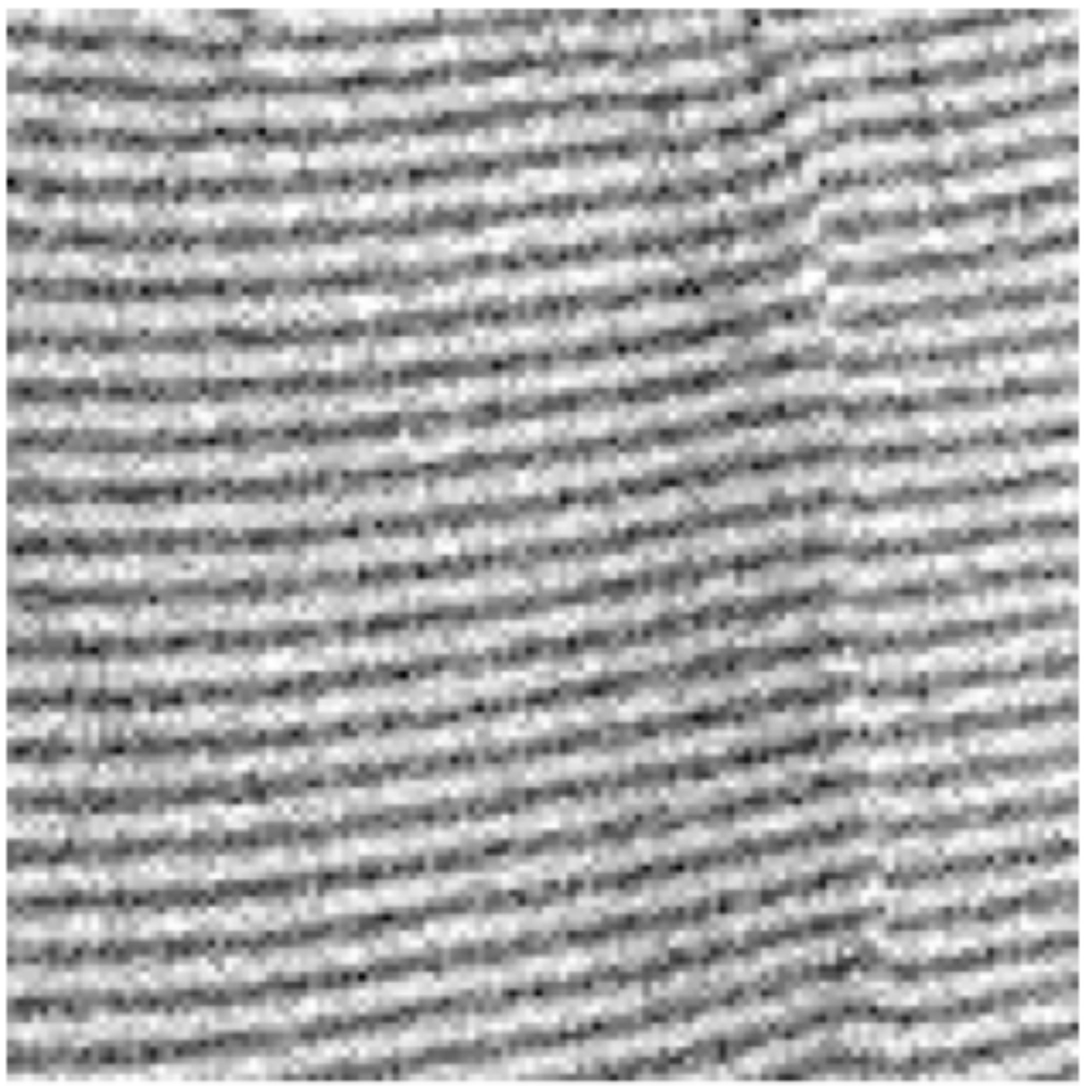
- h is a parameter, its optimal value correspond to the noise level
- the minimum size of the neighborhood is 7x7 (little noise)



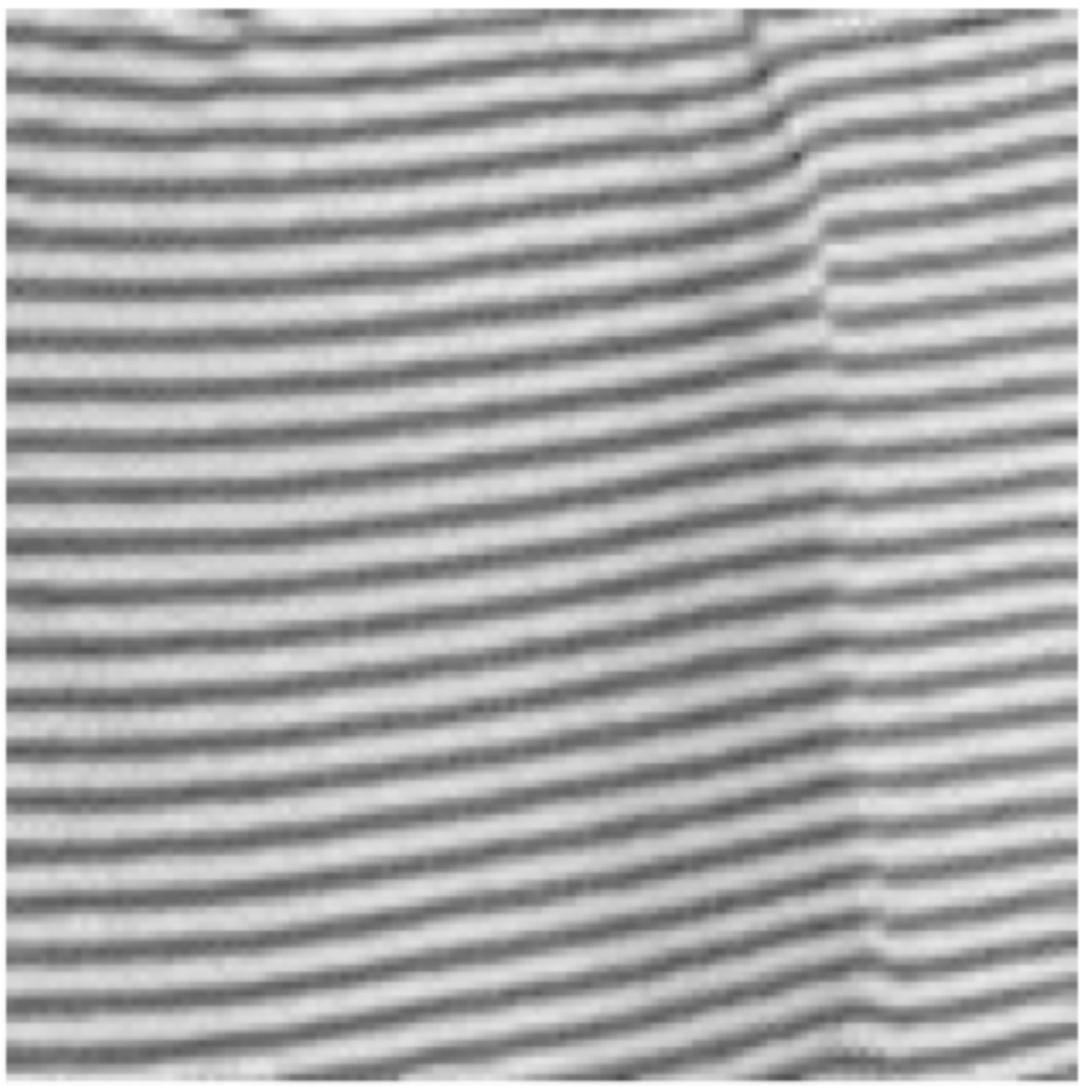
Non local means (NL-means)

- Particularly suitable for :
  - periodic information
  - textures

Noisy image with standard deviation 30.



NL-means restored image



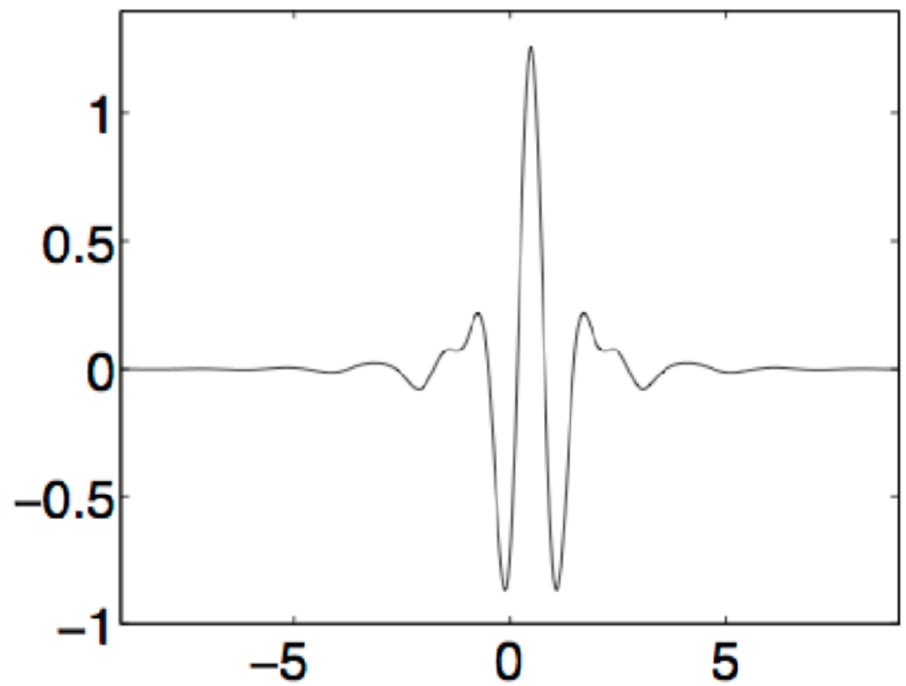
NL-means denoising experiment with a nearly periodic image.

Non local means (NL-means)

- Example of results

Wavelet transform : shrinkage

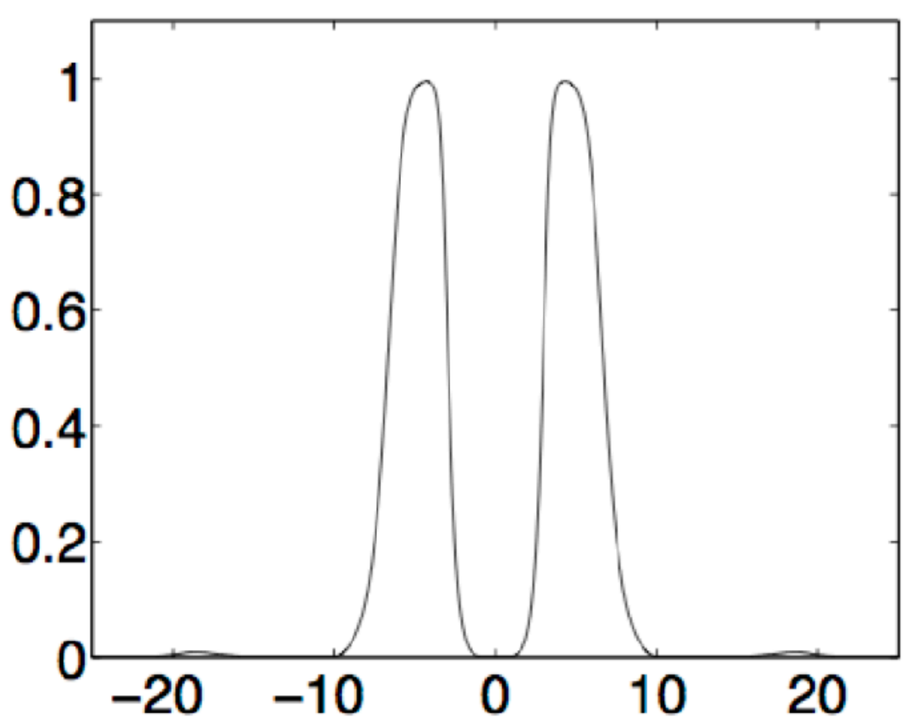
- Wavelet



Wavalet in signal (time or space) domain

$$[W_\psi f](a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} \overline{\psi\left(\frac{x-b}{a}\right)} f(x)dx$$

a : scale factor
b : translation

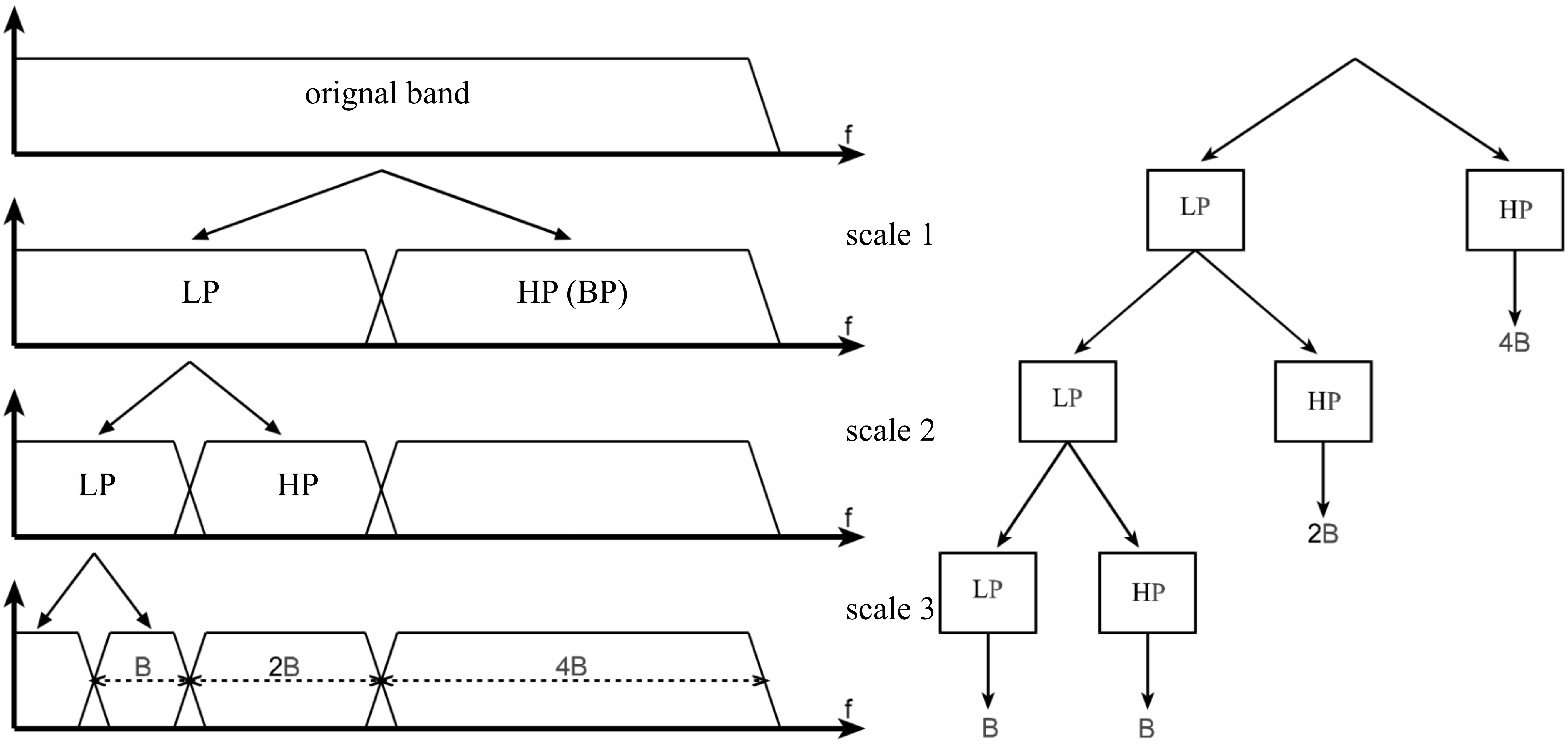


Wavalet in frequency (Fourier) domain

equivalent to a band-pass filter

Wavelet transform : shrinkage

- Pyramidal decomposition / analysis



Inverse scheme : exact reconstruction

Wavelet transform : shrinkage

- Example

N samples

N/32 samples

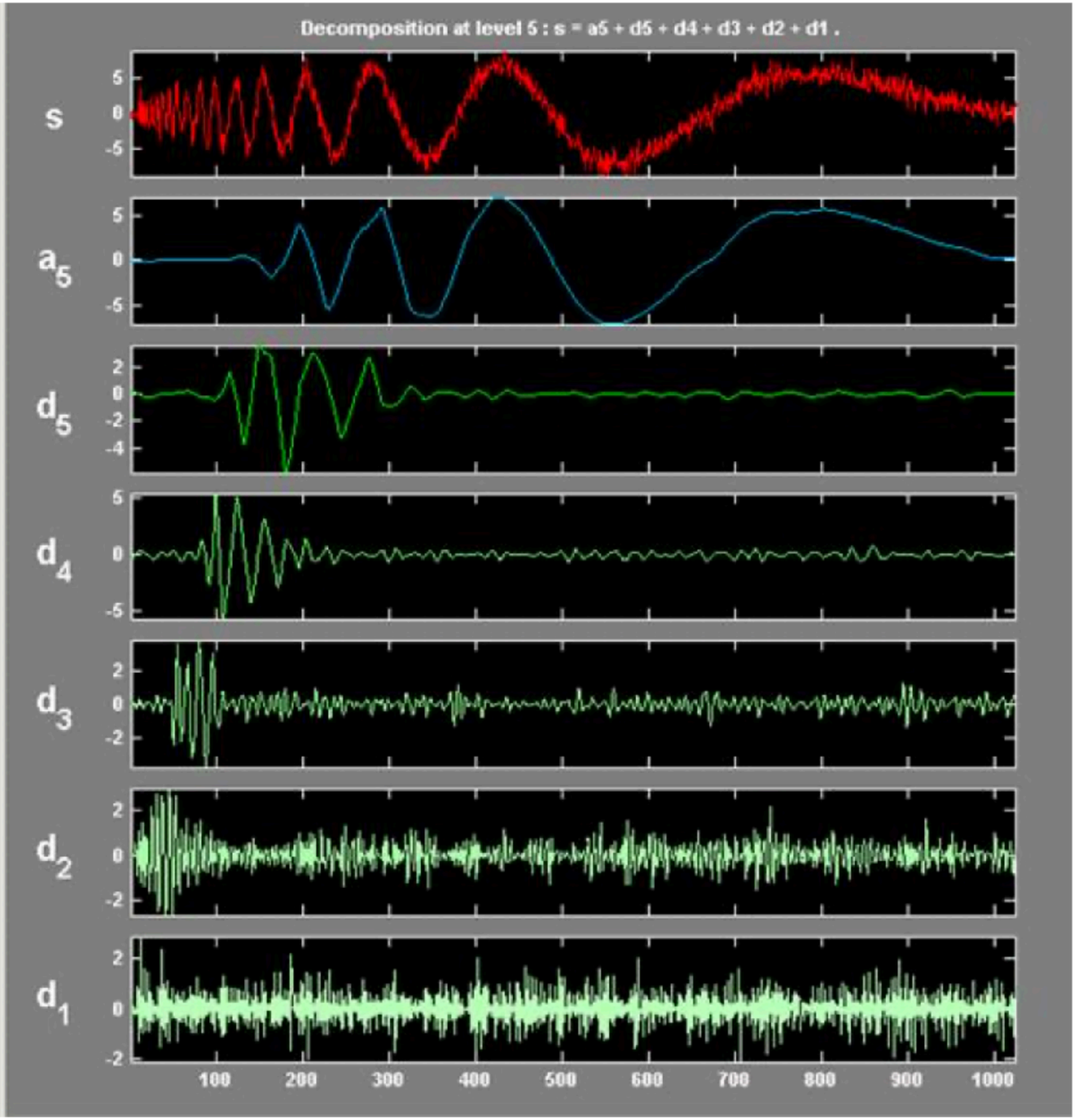
N/32 samples

N/16 samples

N/8 samples

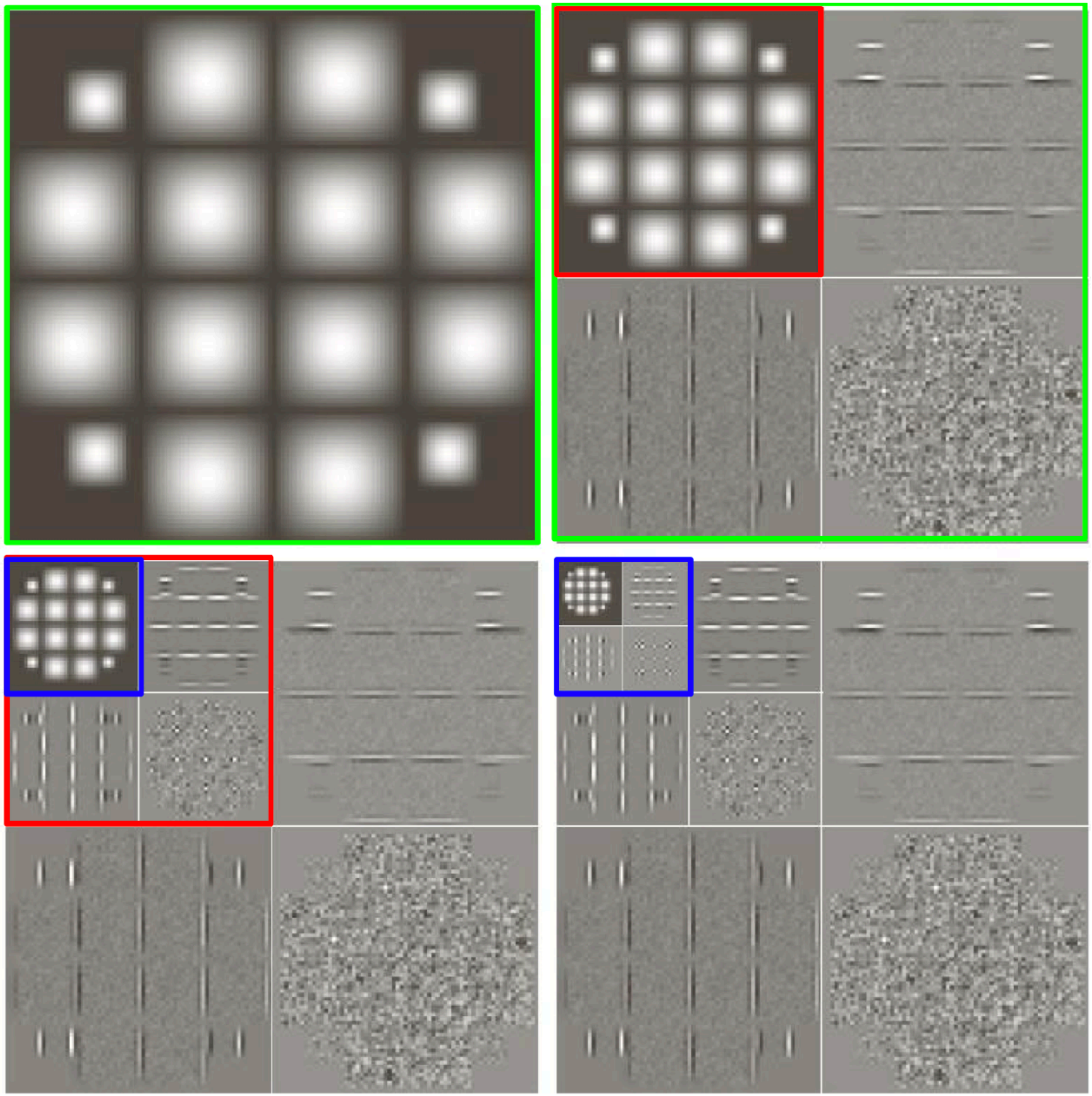
N/4 samples

N/2 samples



Wavelet transform : shrinkage

- Examples in 2D
  - 1 level
  - 2 levels
  - 3 levels



Wavelet transform : shrinkage

- Assumption : correlation in signal, no correlation in noise
- Property of the Wavelet transform / noise
  - Sparsity : a space transform which contains only few coefficients to represent efficiently information
    - ✓ Fourier transform : sparsity only for signals close to *sin* function
    - ✓ Wavelet transform : more general and more efficient for various signals. The wavelet can be chosen according to the signal.
  - Singularity localization
    - ✓ noise peak is a singularity
    - ✓ noise peak is <u>locally</u> detected whatever the level / scale (*persistence*)

Wavelet transform : shrinkage

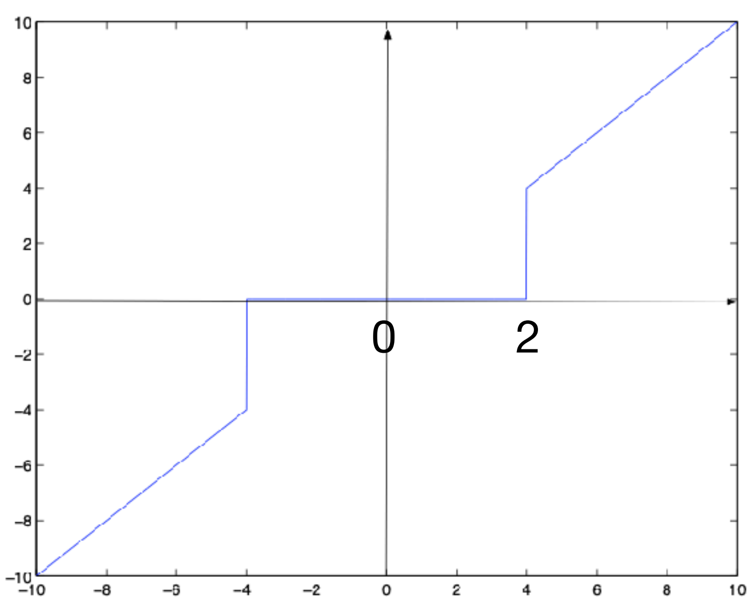
- Algorithm

Image > WT > Soft thresholding T_u > WT^{-1} > denoised image

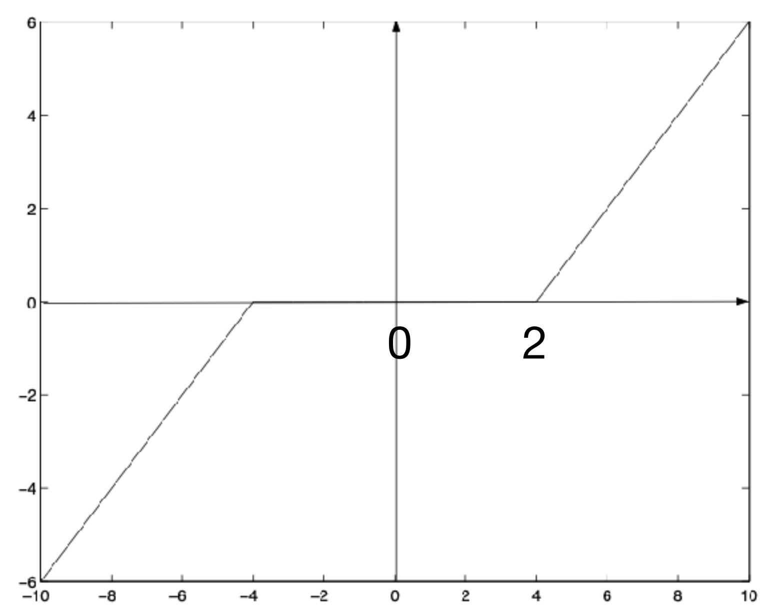
$$T_u = \sigma\sqrt{2\log N}$$

N size of the current signal (level)

σ is generally unknown …



Hard thresholding



Soft thresholding

[… 0 1 2 3 4 …]
=> [… 0 0 0 3 4 …]

[… 0 1 2 3 4 …]
=> [… 0 0 0 1 2 …]

(additive noise)

Wavelet transform : shrinkage

- Examples

(c) Denoised using Hard Thresholding

(d) Denoised using Soft Thresholding

Wavelet transform : shrinkage

- Improvements
  - threshold (noise estimation)
  - wavelet
  - …

Other approaches

- A starting reference

  Linwei Fan , Fan Zhang , Hui Fan and Caiming Zhang, « Brief review of image denoising techniques », Visual Computing for Industry, Biomedicine, and Art, 2019

- Other noises (presence of noise is not equivalent to GWN)
  - example : salt & pepper noise (typically median filters)