# Wavelet in Astronomy: From the Isotropic Undecimated WT to Compressed Sensing

J.A. Starck CEA, IRFU, Service d'Astrophysique, France jstarck@cea, fr https://jstarck.free.fr

Collaborators: J. Bobin, CEA.

- Wavelet in Astronomy
- The Isotropic Undecimated Wavelet Transform
- Compressed Sensing in Astronomy
- CS and the Herschel Satellite

## **Huge Impact of Wavelets in Astronomy**

3000 astronomical papers with the word "Wavelet" in the abstract, and in ALL domains of astrophysics.

76 papers already published in 2008 with the word "Wavelet" in the abstract, and 15 in the title!

- Sun: active region oscillations (Ireland et al., 1999; Blanco et al., 1999), determination of solar cycle length variations (Fligge et al., 1999), feature extraction from solar images (Irbah et al., 1999), velocity fluctuations (Lawrence et al., 1999).
- Solar system: asteroidal resonant motion (Michtchenko and Nesvorny, 1996), classification of asteroids (Bendjoya, 1993), Saturn and Uranus ring analysis (Bendjoya et al., 1993; Petit and Bendjoya, 1996).
- Star studies: Ca II feature detection in magnetically active stars (Soon et al., 1999), variable star research (Szatmary et al., 1996).
- Interstellar medium: large-scale extinction maps of giant molecular clouds using optical star counts (Cambrésy, 1999), fractal structure analysis in molecular clouds (Andersson and Andersson, 1993).
- Planetary nebula detection: confirmation of the detection of a faint planetary nebula around IN Com (Brosch and Hoffman, 1999), evidence for extended high energy gamma-ray emission from the Rosette/Monoceros Region (Jaffe et al., 1997).
- Galaxy: evidence for a Galactic gamma-ray halo (Dixon et al., 1998).
- QSO: QSO brightness fluctuations (Schild, 1999), detecting the non-Gaussian spectrum of QSO  $Ly_{\alpha}$  absorption line distribution (Pando and Fang, 1998).
- Gamma-ray burst: GRB detection (Kolaczyk, 1997; Norris et al., 1994) and GRB analysis (Greene et al., 1997; Walker et al., 2000).
- Black hole: periodic oscillation detection (Steiman-Cameron et al., 1997; Scargle, 1997)
- Galaxies: starburst detection (Hecquet et al., 1995), galaxy counts (Aussel et al., 1999; Damiani et al., 1998), morphology of galaxies (Weistrop et al., 1996; Kriessler et al., 1998), multifractal character of the galaxy distribution (Martínez et al., 1993a).
- Galaxy cluster: sub-structure detection (Pierre and Starck, 1998; Krywult et al., 1999; Arnaud et al., 2000), hierarchical clustering (Pando et al., 1998a), distribution of superclusters of galaxies (Kalinkov et al., 1998).
- Cosmic Microwave Background: evidence for scale-scale correlations in the Cosmic Microwave Background radiation in COBE data (Pando et al., 1998b), large-scale CMB non-Gaussian statistics (Popa, 1998; Aghanim et al., 2001), massive CMB data set analysis (Gorski, 1998).
- Cosmology: comparing simulated cosmological scenarios with observations (Lega et al., 1996), cosmic velocity field analysis (Rauzy et al., 1993).

The Astrophysical Journal, 686:1–12, 2008 October 10 © 2008. The American Astronomical Society. All rights reserved. Printed in U.S.A. DOI: 10.1086/591236

#### Discreteness Effects in ACDM Simulations: A Wavelet-Statistical View

Alessandro B. Romeo, Oscar Agertz, Ben Moore, and Joachim Stadel Onsala Space Observatory, Chalmers University of Technology, SE-43992 Onsala, Sweden and Institute for Theoretical Physics, University of Zurich, CH-8057 Zurich, Switzerland

#### ABSTRACT

The effects of particle discreteness in N-body ACDM simulations are still an intensively debated issue. In this paper we explore such effects, taking into account the scatter caused by the randomness of the initial conditions and focusing on the statistical properties of the cosmological density field. For this purpose, we run large sets of ACDM simulations and analyze them using a wide variety of diagnostics, including new and powerful wavelet statistics. Among other facts, we point out (1) that dynamical evolution does not propagate discreteness noise up from the small scales at which it is introduced and

(2) that one should aim to satisfy the condition  $\varepsilon$  2d, where  $\varepsilon$  is the force resolution and d is the interparticle distance. We clarify what such a condition means and how to implement it in modern cosmological codes.



# **COSMOS data :**

Maps of the Universe's Dark matter scaffolding, Massey et al, Nature,

Vol. 445, pp. 286-290, 2007

# Baryonic and non-baryonic matter comparison at large scale

The total projected mass map from WL (dominated by dark matter) is shown as contours. It is compared to 3 independent baryonic tracers : stellar mass (in blue), galaxy number density seen in optical and near-IR light (in green) and the hot gas seen in **x-rays** (in red).



This broad success of the wavelet transform is due to the fact that astronomical data generally gives rise to complex hierarchical structures, often described as fractals.

Using multiscale approaches such as the wavelet transform, an image can be decomposed into components at different scales, and the wavelet transform is therefore well-adapted to the study of astronomical data.

## The Filter Bank

In order to get an exact reconstruction, two conditions are required for the filters:

- Dealiasing condition:  $\hat{h}(\nu+\frac{1}{2})\hat{\tilde{h}}(\nu)+\hat{g}(\nu+\frac{1}{2})\hat{\tilde{g}}(\nu)=0$
- Exact restoration:  $\hat{h}(
  u)\hat{ ilde{h}}(
  u)+\hat{g}(
  u)\hat{ ilde{g}}(
  u)=1$

#### The Isotropic Undecimated Wavelet Transform

 Filters do not need to verify the dealiasing condition. We need only the exact restoration condition:

$$\hat{h}(
u)\hat{ ilde{h}}(
u)+\hat{g}(
u)\hat{ ilde{g}}(
u)=1$$

- Filters do not need to be (bi) orthogonal.
- Filters must be symmetric.
- In 2D, we want h(x, y) = h(x)h(y) for fast calculation and more important, h(x, y) must nearly isotropic.

*h* is derived from a  $B_3$  spline:  $h_{1D}(k) = [1, 4, 6, 4, 1]/16$ , and in 2D  $h_{2D} = h_{1D}h_{1D} =$ 

 $\left(\begin{array}{cccc} \frac{1}{16} & \frac{1}{4} & \frac{3}{8} & \frac{1}{4} & \frac{1}{16} \end{array}\right) \otimes \left(\begin{array}{cccc} \frac{1/16}{1/4} \\ \frac{3/8}{1/4} \\ \frac{1}{1/16} \end{array}\right) = \left(\begin{array}{ccccc} \frac{\frac{1}{256}}{164} & \frac{1}{128} & \frac{1}{64} & \frac{1}{256} \\ \frac{1}{64} & \frac{1}{16} & \frac{3}{32} & \frac{1}{16} & \frac{1}{64} \\ \frac{3}{128} & \frac{3}{32} & \frac{9}{64} & \frac{3}{32} & \frac{3}{128} \\ \frac{1}{164} & \frac{1}{16} & \frac{3}{32} & \frac{1}{16} & \frac{1}{64} \\ \frac{1}{256} & \frac{1}{64} & \frac{1}{128} & \frac{1}{64} & \frac{1}{256} \end{array}\right)$ 







## **Isotropic Undecimated Wavelet Transform**

$$\varphi = B_3 - \text{spline}, \quad \frac{1}{2}\psi(\frac{x}{2}) = \frac{1}{2}\varphi(\frac{x}{2}) - \varphi(x) \qquad \qquad I(k,l) = c_{J,k,l} + \sum_{j=1}^J w_{j,k,l} \\ h = [1,4,6,4,1]/16, \quad g = \delta - h, \quad \tilde{h} = \tilde{g} = \delta$$



## **Dynamic Range Compression**

$$I_{k,l} = \log(c_{J,k,l}) + \sum_{j=1}^J \operatorname{sgn}(w_{j,k,l}) \log(\mid w_{j,k,l} \mid +\epsilon)$$



Left - Hale-Bopp Comet image. Middle - histogram equalization results, Right - wavelet-log representations.







Data restoration: ringing artefacts may appear around strongest sources, which is due to the negative part of the wavelet function.

Large scale structure thresholding may create artefacts.

$$\hat{h}^{*}\left(\nu + \frac{1}{2}\right)\hat{\tilde{h}}(\nu) + \hat{g}^{*}\left(\nu + \frac{1}{2}\right)\hat{\tilde{g}}(\nu) = 0$$
$$\hat{h}^{*}(\nu)\hat{\tilde{h}}(\nu) + \hat{g}^{*}(\nu)\hat{\tilde{g}}(\nu) = 1$$

## **MODIFIED Isotropic Undecimated Wavelet Transform**

J.-L. Starck, J. Fadili and F. Murtagh, "The Undecimated Wavelet Decomposition and its Reconstruction", IEEE Trans. on Image Processing, 16, 2, pp 297--309, 2007.

$$\hat{h}(\nu)\hat{\tilde{h}}(\nu) + \hat{g}(\nu)\hat{\tilde{g}}(\nu) = 1$$

$$\varphi = B_3 - \text{spline}, \quad \frac{1}{2}\psi(\frac{x}{2}) = \frac{1}{2}\varphi(\frac{x}{2}) - \varphi(x)$$

$$h = [1,4,6,4,1]/16, \quad g = Id - h, \quad \tilde{h} = \tilde{g} = Id$$

$$I(k,l) = c_{J,k,l} + \sum_{j=1}^{J} w_{j,k}$$

$$\varphi = B_3 - \text{spline}, \quad \frac{1}{2}\psi(\frac{x}{2}) = \frac{1}{2}\varphi(\frac{x}{2}) - \varphi(x)$$

$$h = [1,4,6,4,1]/16, \quad g = Id - h \quad \tilde{h} = h$$

$$ig = Id + h$$

$$\frac{1}{2}\tilde{\psi}(\frac{x}{2}) = \phi(x) + \frac{1}{2}\phi(\frac{x}{2})$$

$$I(k,l) = c_{f,k,l} + \sum_{j=1}^{l} w_{j,k,l}$$

$$\varphi = B_3 - \text{spline}, \quad \frac{1}{2}\psi(\frac{x}{2}) = \frac{1}{2}\varphi(\frac{x}{2}) - \varphi(x) \qquad \longrightarrow \qquad \tilde{g} = Id$$

$$h = [1,4,6,4,1]/16 \quad \tilde{h} = h$$

$$g = Id - h * h$$

$$\tilde{g} = Id - h * h$$

## **Reconstruction using Scaling Functions**

$$s_l \hspace{0.1 cm} = \hspace{0.1 cm} \sum_k c_{J,k} \phi_{J,l}(k) + \sum_k \sum_{j=1}^J ilde{\phi}_{j,l}(k) w_{j,k} \hspace{0.1 cm}$$





### Haar Transform with Smooth Reconstruction Filters





## **Compressed Sensing**



 \* E. Candès and T. Tao, "Near Optimal Signal Recovery From Random Projections: Universal Encoding Strategies? ", IEEE Trans. on Information Theory, 52, pp 5406-5425, 2006.
 \* D. Donoho, "Compressed Sensing", IEEE Trans. on Information Theory, 52(4), pp. 1289-1306, April 2006.
 \* E. Candès, J. Romberg and T. Tao, "Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information", IEEE Trans. on Information Theory, 52(2) pp. 489 – 509, Feb. 2006.
 A non linear sampling theorem

"Signals with exactly K components different from zero can be recovered perfectly from ~ K log N incoherent measurements"

Replace samples with *few linear projections*  $y=\Theta x$ 



⇒Application: Compression, tomography, ill posed inverse problem.

## **Compressed Sensing Reconstruction**

Measurements:  $y_k = \langle x, \theta_k \rangle$ Reconstruction via non linear processing:  $\min \|x\|_1$  s.t.  $y = \Theta_{\Lambda} x$ In practice, x is sparse in a given **dictionary**:  $x = \Phi lpha$ and we need to solve:  $\min \|lpha\|_1$  s.t.  $y = \Theta_\Lambda \Phi lpha$  $\mu_{\boldsymbol{\Theta}, \boldsymbol{\Phi}} = \max_{i, j} \left| \left\langle \theta_i, \phi_j \right\rangle \right|$ the number of required measurements is :  $m \geq C.\mu_{\Theta.\Phi}^2.S.\log n$ 

## **Compressed Sensing For Data Compression**

Compressed Sensing presents several interesting properties for data compress:

- •Compression is **very fast** ==> good for on-board applications.
- •Very **robust** to bit loss during the transfer.
- •Decoupling between compression/decompression.
- •Data protection.
- •Linear Compression.



But clearly not as competitive to JPEG or JPEG2000 to compress an image.

## **Compressed Sensing Impact in astronomy**

$$y = \Theta x$$

Typical Astronomical Data related to CS

- (radio-) Interferometry:  $\Theta$  = Fourier transform  $\Phi$  = Id (or Wavelet transform)
- Period detection in temporal series  $\Theta = Id$ 
  - $\Phi$  = Fourier transform
- Gamma Ray Instruments (Integral) Acquisition with coded masks

CS gives another point of view on some existing methods

- Inpainting: 
$$\Theta$$
 = Ic

$$\min_{\alpha} \| \alpha \|_0 \quad \text{s.t.} \quad y = M \Phi \alpha = M x$$

New problems that can be addressed by CS

==> Data compression: the case of Herschel satellite

## (Radio-) Interferometry



J.L. Starck, A. Bijaoui, B. Lopez, and C. Perrier, "Image Reconstruction by the Wavelet Transform Applied to Aperture Synthesis", Astronomy and Astrophysics, 283, 349--360, 1994.

## Wavelet - CLEAN minimizes well the $l_0$ norm

But recent  $l_0$ - $l_1$  minimization algorithms would be clearly much faster.





## **Interpolation of Missing Data: Inpainting**

•M. Elad, J.-L. Starck, D.L. Donoho, P. Querre, "Simultaneous Cartoon and Texture Image Inpainting using Morphological Component Analysis (MCA)", ACHA, Vol. 19, pp. 340-358, 2005.
•M.J. Fadili, J.-L. Starck and F. Murtagh, "Inpainting and Zooming using Sparse Representations", in press.

$$\Theta_{\Lambda} = \operatorname{Id}_{\Lambda} \quad \min_{\alpha} \parallel \alpha \parallel_{0} \qquad y = M \Phi \alpha = M x$$



Where M is the mask:  $M(i,j) = 0 \implies$  missing data  $M(i,j) = 1 \implies$  good data



## **Interpolation of Missing Data: Inpainting**

•M. Elad, J.-L. Starck, D.L. Donoho, P. Querre, "Simultaneous Cartoon and Texture Image Inpainting using Morphological Component Analysis (MCA)", ACHA, Vol. 19, pp. 340-358, 2005.
•M.J. Fadili, J.-L. Starck and F. Murtagh, "Inpainting and Zooming using Sparse Representations", in press.

$$\Theta_{\Lambda} = \mathbf{Id}_{\Lambda}$$
  $\min_{\alpha} \|\alpha\|_{\ell_0}$  s.t.  $y = Mx$ 

Where M is the mask:  $M(i,j) = 0 \implies$  missing data  $M(i,j) = 1 \implies$  good data

$$x^{(n+1)} = \mathcal{S}_{\Phi,\lambda^{(n)}} \left\{ x^{(n)} + M\left(y - x^{(n)}\right) \right\}$$











Inpainted with the curvelet dictionary (80% data missing)

Jalal Fadili's web page (<u>http://www.greyc.ensicaen.fr/~jfadili</u>).



## HERSCHEL



This space telescope has been designed to observe in the far-infrared and sub-millimeter wavelength range.

Its launch is scheduled for the beginning of 2009. The shortest wavelength band, 57-210 microns, is covered by PACS (Photodetector Array Camera and Spectrometer).

Herschel data transfer problem: -no time to do sophisticated data compression on board.

-a compression ratio of 6 must be achieved.

==> solution: averaging of six successive images on board

CS may offer another alternative.

## The proposed Herschel compression scheme





Good measurements must be incoherent with the basis in which the data are assumed to be sparse.

Noiselets (Coifman, Geshwind and Meyer, 2001) are an orthogonal basis that is shown to be highly incoherent with a wide range of practical sparse representations (wavelets, Fourier, etc).

Advantages:

Low computational cost (O(n))

Most astronomical data are sparsely represented in a wide range of wavelet bases

## The decoding scheme



Physical priors

$$\alpha^{(h)} = \mathcal{S}_{\gamma} \left\{ \alpha^{(h-1)} + \boldsymbol{\Phi}^{T} \boldsymbol{\Theta}^{T} \left( y^{\sharp} - \mathbf{I}_{\Lambda} \boldsymbol{\Theta} \boldsymbol{\Phi} \alpha^{(h-1)} \right) \right\}$$

1. Set the number of iterations  $I_{\max}$  and threshold  $\gamma^{(0)} = \| \Phi^T \Theta^T y^{\sharp} \|_{\infty}$ .  $x^{(0)}$  is set to zero. 2. While  $\gamma^{(h)}$  is higher than a given lower bound  $\gamma_{\min}$ 

- Compute the measurement projection of  $x^{(h-1)}$ :  $y^{(h)} = \mathbf{I}_{\Lambda} \Theta x^{(h-1)}$ .
- Estimate the current coefficients  $\alpha^{(h)}$ :  $\alpha^{(h)} = S_{\gamma^{(h)}} \left\{ \alpha^{(h-1)} + \Phi^{T} \Theta^{T} \left[ y^{\sharp} - y^{(h)} \right] \right\}.$
- Get the new estimate of x by reconstructing from the selected coefficients  $\alpha^{(h)}$ :  $x^{(h)} = \Phi \alpha^{(h)}$ .
- 3. Decrease the threshold  $\gamma^{(h)}$  following a given strategy.

$$y \longrightarrow \frac{\min_{\alpha} \|\alpha\|_{\ell_1} \text{ s.t. } \|y - \Theta_{\Lambda} \Phi \alpha\|_{\ell_2} \le \epsilon}{x = \Phi \alpha} \longrightarrow \mathcal{X}$$
  
Physical priors

Six consecutive observations of the same field can be decompressed together:

$$x^{(n+1)} = \frac{1}{6} \mathcal{S}_{\Phi,\lambda^{(n)}} \left\{ x^{(n)} + \sum_{i=1,\dots6} \Theta^T \left( y - \Theta x^{(n)} \right) \right\}$$

## Sensitivity: CS versus mean of 6 images



## Resolution: CS versus Mean

Simulated image



Mean of six images

Simulated noisy image with flat and dark





Compressed sensing reconstructed images

SNR	-17.3	-9.35	-3.3	0.21	2.7	4.7	6.2	7.6	8.7
Intensity	900	2250	4500	6750	9000	11250	13500	15750	18000
MO6	3	3	3	3	3	3	3	3	3
CS	2.33	2.33	2	2	2	2	2	2	2
The CS-based compression entails a resolution gain equal to a $30\%$ of the spatial									
RESOLUTION PROVIDED BY MO6.									

#### Resolution limit versus SNR

# **CS and Herschel Status**

- CS compression is implemented in the Herschel on-board software (as an option).
- CS Tests in flight will be done.
- Software developments required for an efficient decompression (taking into account dark, flatp-field, PSF, etc).
- The CS decompression is fully integrated in the data processing pipeline.

## **Data Fusion: JPEG versus Compressed Sensing**



Simulated source



One of the 10 observations



Averaged of the 10 JPEG compressed images (CR=4)



Reconstruction from the 10 compressed sensing images (CR=4)

## JPEG2000 Versus Compressed Sensing

Compression Rate: 25





## **Conclusions on CS**

Compressed Sensing gives us a clear direction for:

- (radio-) interferometric data reconstruction
- periodic signals with sampled irregularly
- gamma-ray image reconstruction

CS provides an interesting framework and a good theoretical support for our inpainting work.

CS can be a good solution for on board data compression.

CS is a highly competitive solution for compressed data fusion.

**PREPRINT**: