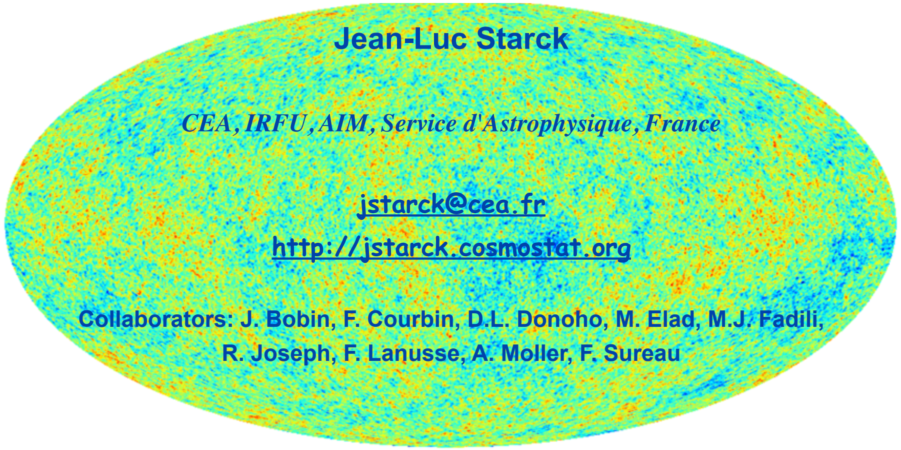
Learning and Blind Source Separation



Jean-Luc Starck

CEA, IRFU, AIM, Service d'Astrophysique, France

jstarck@cea.fr

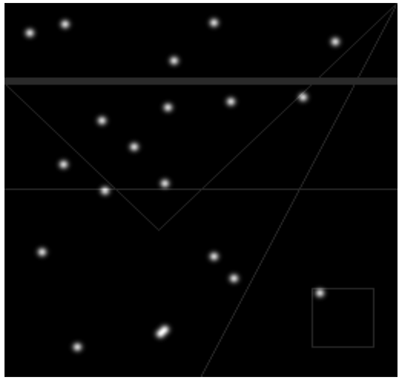
http://jstarck.cosmostat.org

Collaborators: J. Bobin, F. Courbin, D.L. Donoho, M. Elad, M.J. Fadili, R. Joseph, F. Lanusse, A. Moller, F. Sureau

Learning and Blind Source Separation

- **Part 1: Introduction:**
- Part 2: Monochannel Source Separation
- Part 3: Multichannel Blind Source Separation
- Part 4: Dictionary Learning

2D mixture: Lines + Gaussians



Gemini infrared data

Artifact & SNe detection


SNa and its host galaxy

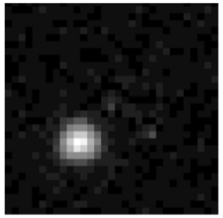


Image after subtraction of the reference sky

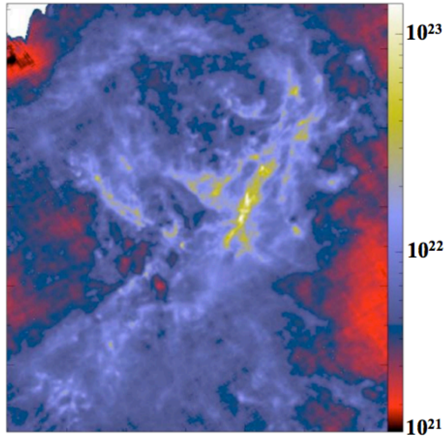
Subtracted image

Detection catalogue

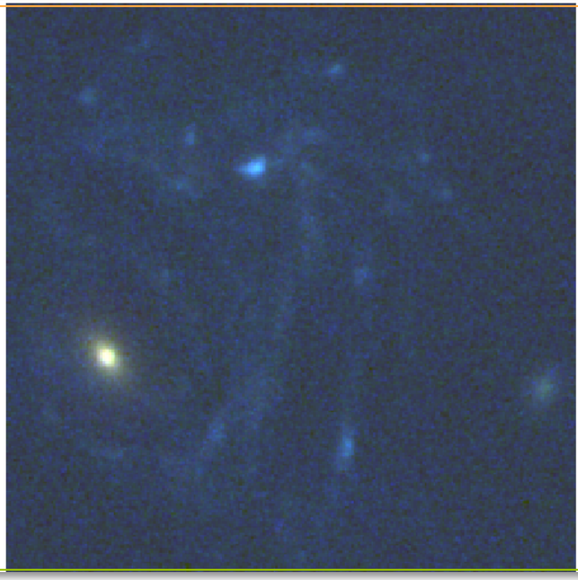
- SNe are detected by subtraction of a reference image.
- In practice, subtracted images are contaminated by artifacts which make the detection difficult

Mixture Problem

Herschel (SPIRE+PACS)
Column density map (H_2/cm^2)



Multichannel data



Mixture modeling

- Any galaxy spectrum

$$Spectrum = Baseline + EmissionLine + AbsorptionLine + Noise$$

Machado, et al, "Darth Fader: Using wavelets to obtain accurate redshifts of spectra at very low signal-to-noise, in press", ArXiv:1309.3579, 560, id.A83, pp 20, 2013.

- Mono-channel mixture:

$$Y = X_1 + X_2 + N$$

- Multichannel mixture:

$$Y_i = H_i * \sum_{s=1}^{S} a_{i,s} X_s + N$$

Learning and Blind Source Separation

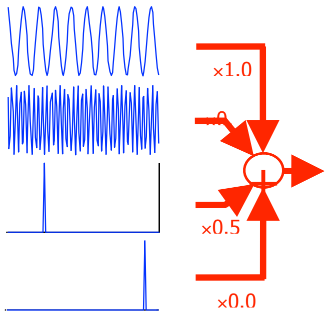
- Part 1: Introduction:
- **Part 2: Monochannel Source Separation**
- Part 3: Multichannel Blind Source Separation
- Part 4: Dictionary Learning

Monochannel Mixture

$$min_{X_1,X_2} \parallel Y - (X_1 + X_2) \parallel^2 + C_1(X_1) + C_2(X_2)$$

C_1: $C_1(X1)$ must be low and $C_1(X2)$ must be high

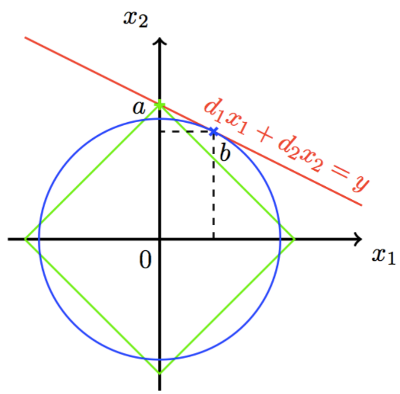
C_2: $C_2(X1)$ must be high and $C_2(X2)$ must be low



$$C_1(X_1) = \parallel \Phi_1^t X_1 \parallel_p$$

$$C_2(X_2) = \parallel \Phi_2^t X_2 \parallel_p$$

L1 Norm & Sparsity



$$C_1(X_1) = \| \Phi_1^t X_1 \|_p$$

$$C_2(X_2) = \| \Phi_2^t X_2 \|_p$$

Weak Sparsity or Compressible Signals

A signal *s* (*n* samples) can be represented as sum of weighted elements of a given dictionary

Dictionary (basis, frame)

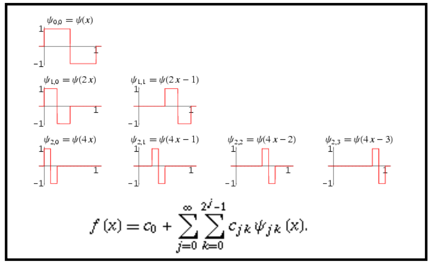
$$\Phi = \{\phi_1, \ldots, \phi_K\}$$

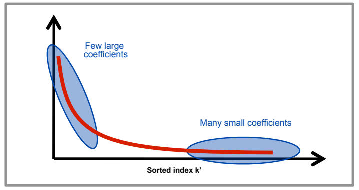
Atoms

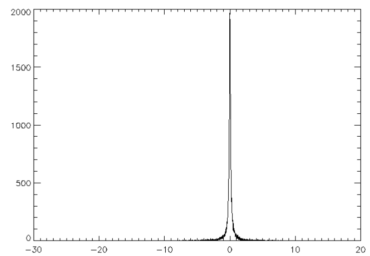
$$s = \sum_{k=1}^{K} \alpha_k \phi_k = \Phi\alpha$$

coefficients

Ex: Haar wavelet

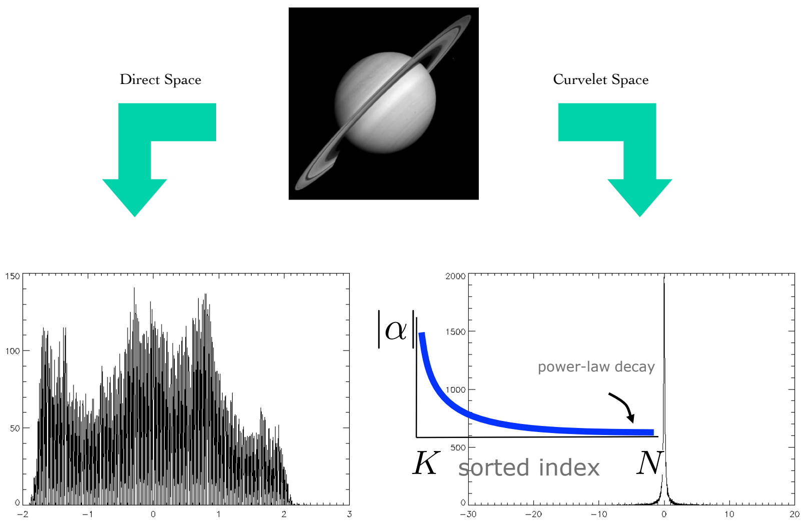






- Fast calculation of the coefficients
- Analyze the signal through the statistical properties of the coefficients
- Approximation theory uses the sparsity of the coefficients

Weak Sparsity or Compressible Signals



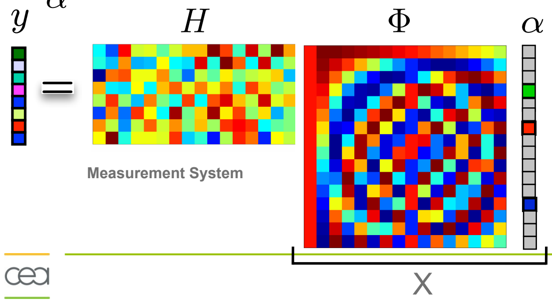
Sparse Recovery & Inverse Problems

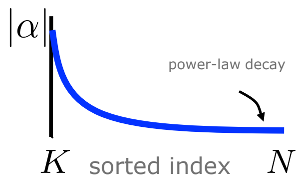
$$Y = HX + N$$

$X = \Phi\alpha$, and α is sparse

- Denoising
- Deconvolution
- Component Separation
- Inpainting
- Blind Source Separation
- Minimization algorithms
- Compressed Sensing

$$\min_{\alpha} \|\alpha\|_p^p \quad \text{subject to} \quad \|Y - H\Phi\alpha\|^2 \leq \epsilon$$





Morphological Diversity

•*J.-L. Starck, M. Elad, and D.L. Donoho, Redundant Multiscale Transforms and their Application for Morphological Component Analysis, Advances in Imaging and Electron Physics, 132, 2004.*

Sparsity Model: we consider a signal as a sum of K components s_k, each of them being sparse in a given dictionary :

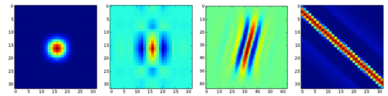
$$Y = X_1 + X_2$$

X_1 can be well approximated with few coefficients in a given domain.

X_2 can be well approximated with few coefficients in **another** domain.

$$min_{X_1, X_2} \parallel Y - (X_1 + X_2) \parallel^2 + C_1(X_1) + C_2(X_2)$$

$$C_1(X_1) = \parallel \Phi_1^t X_1 \parallel_p$$
$$C_2(X_2) = \parallel \Phi_2^t X_2 \parallel_p$$



(A) Starlet (B) Bi-orthogonal (C) Curvelet (D) Ridgelet

Morphological Component Analysis (MCA)

$$min_{X_1,\ldots,X_L} \parallel Y - \sum_{k=1}^{L} X_k \parallel^2 + \lambda \sum_{k=1}^{L} \parallel \Phi_k^t X_k \parallel_p$$

. Initialize all X_k to zero

. Iterate j=1,...,Niter

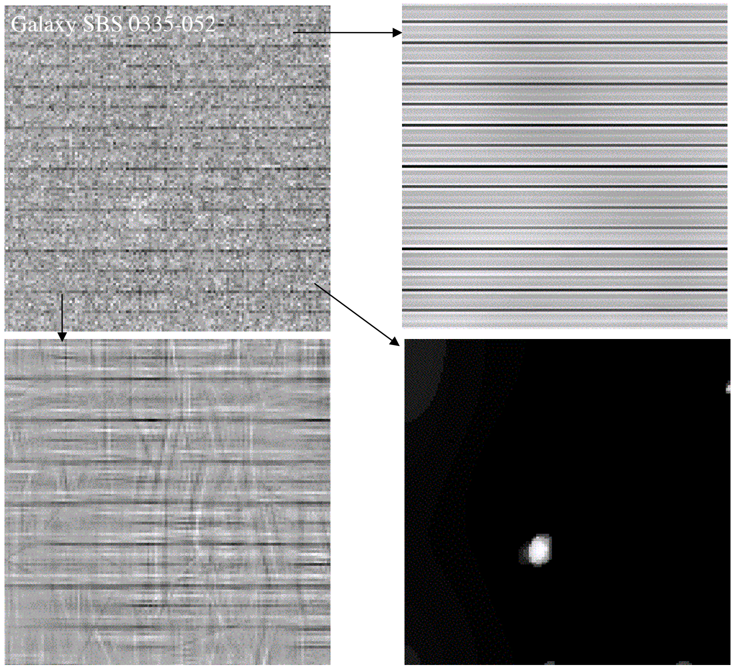
- Iterate k=1,..,L

Update the kth part of the current solution by fixing all other parts and minimizing:

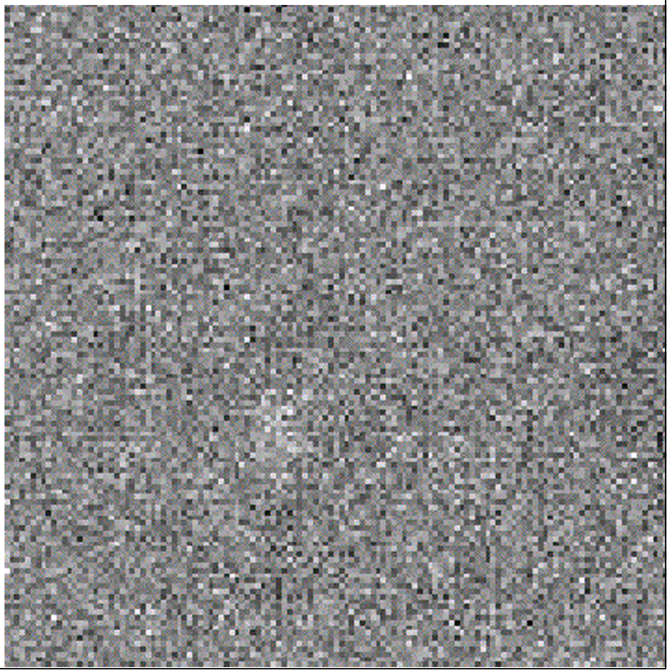
$$min_{X_k} \parallel Y - \sum_{i=1, i \neq k}^{L} X_i - X_k \parallel^2 + \lambda^{(j)} \parallel \Phi_k^t X_k \parallel_p$$

Which is obtained by a simple **hard**/soft thresholding of : $Z = Y - \sum_{i=1, i \neq k}^{L} X_i$

- Decrease the threshold $\lambda^{(j)}$

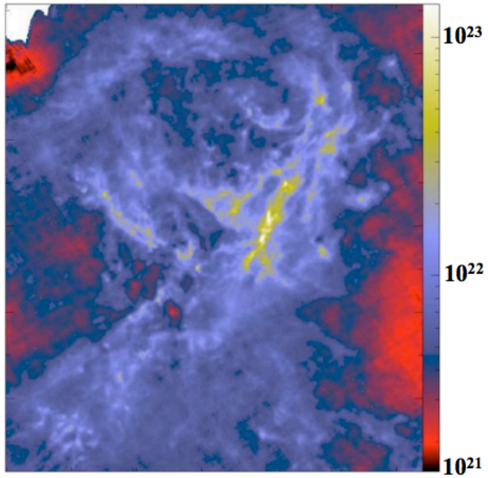

Galaxy SBS 0335-052

Galaxy SBS 0335-052
10 micron
GEMINI-OSCIR



Revealing the structure of one of the nearest infrared dark clouds (Aquila Main: d ~ 260 pc)

Herschel (SPIRE+PACS)
Column density map (H_2/cm^2)



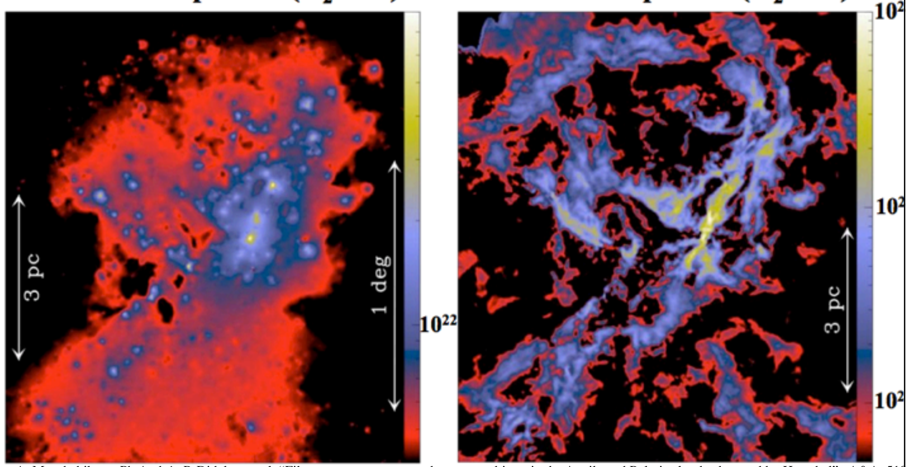
Dense cores form primarily in filaments

Morphological Component Analysis:

***Herschel* Column density map** (P. Didelon based on Starck et al. 2003)

=

Cores — **Wavelet component (H_2/cm^2)** **+** Filaments — **Curvelet component (H_2/cm^2)**



A. Menshchikov, Ph.André, P. Didelon, et al, "Filamentary structures and compact objects in the Aquila and Polaris clouds observed by Herschel", A&A, 518 id.L103, 2010.

MCA based artifact removal for SNe detection

SNa and its host galaxy

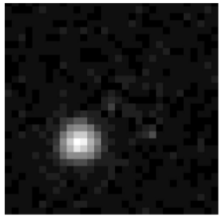


Image after subtraction of the reference sky

Subtracted image

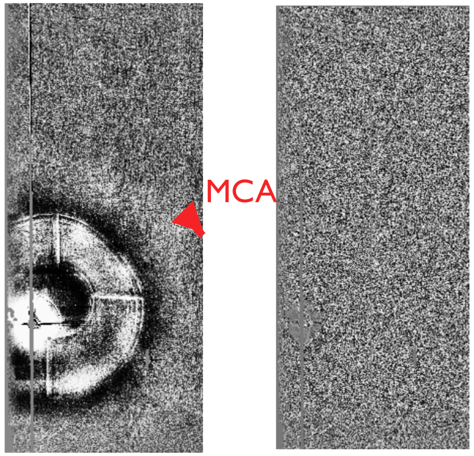
Detection catalogue

- SNe are detected by subtraction of a reference image.

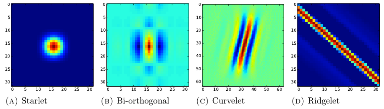
- In practice, subtracted images are contaminated by artifacts which make the detection difficult

Artifact removal for SNe detection

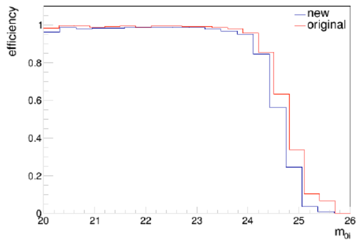
- Möller, et al, 2015, *SNIa detection in the SNLS photometric analysis using Morphological Component Analysis*, *04, Id 041*, JCAP, arxiv:1501.02110.



MCA cleaning of a subtracted image



Dictionaries used for the analysis



Similar detection efficiency but greatly reduced number of spurious detections

Learning and Blind Source Separation

•Part 1: Introduction:

•Part 2: Monochannel Source Separation

•Part 3: Multichannel Blind Source Separation

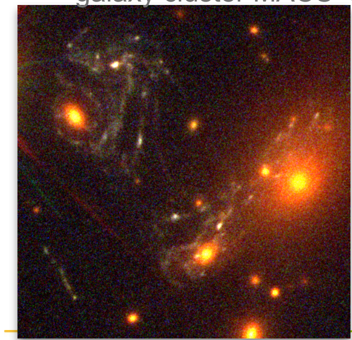
•Part 4: Dictionary Learning

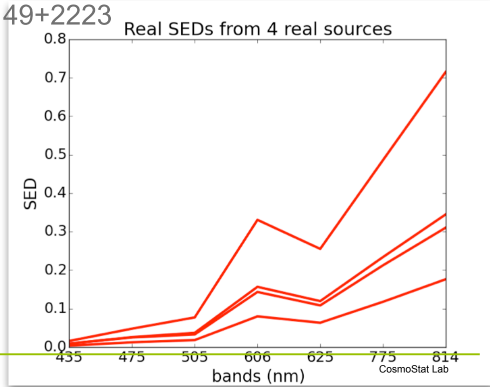
Multichannel data

GOAL: separate the foreground cluster galaxies (red) from the background lensed galaxy (blue).

$$Y_i = H_i * \sum_{s=1}^{S} a_{i,s} X_s + N$$

galaxy cluster MACS~J1149+2223





Morpho-Spectral Diversity

$$Y_i = H_i * \sum_{s=1}^{S} a_{i,s} X_s + N$$

$\mathrm{H}_i = \mathrm{Id}$

The fixing matrix A is assumed to be known

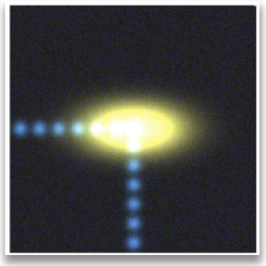
X_s is sparse in $\Phi_s = \mathcal{S}_s \Psi_s$

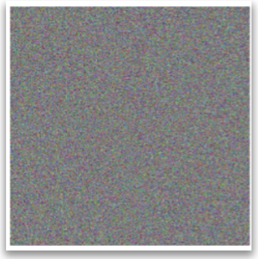
$\forall s, \Psi_s = \Psi$, where Ψ is the starlet transform.

$$\min_X \parallel Y - AX \parallel^2 + \sum_{j=1}^{J} \lambda_j \parallel \Psi^* x_j \parallel_0$$

R. Jospeh, F. Courbin and J.-L. Starck, "Multi-band morpho-Spectral Component Analysis Deblending Tool (MuSCADeT): deblending colourful objects", A&A, submitted, 2015

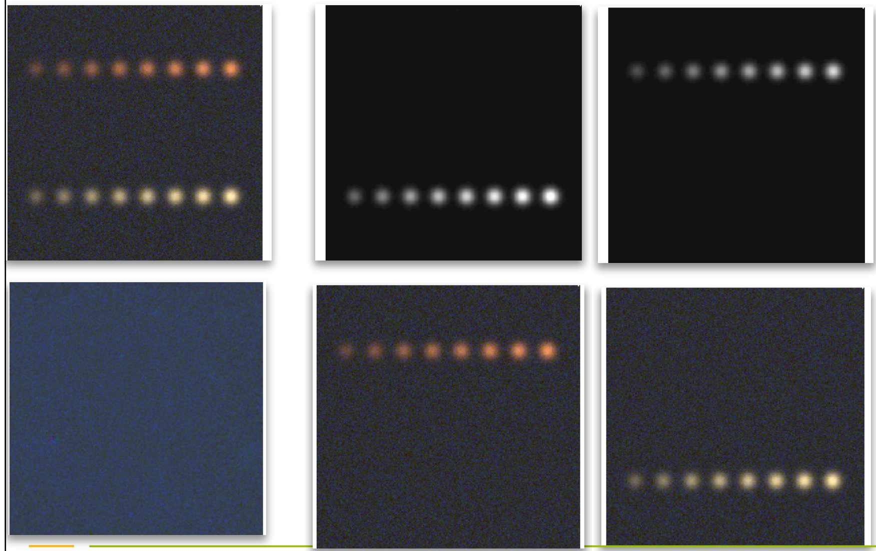
Multichannel data





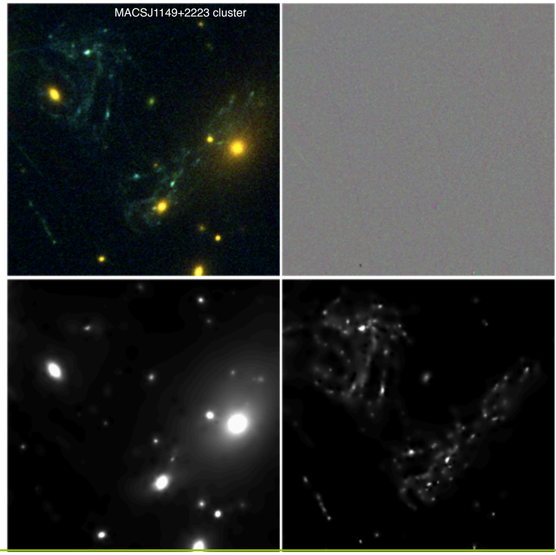
No SED variation.

Multichannel data



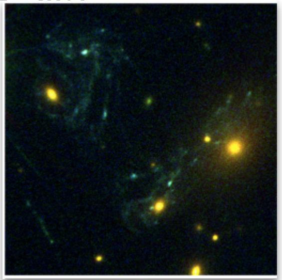
realistic SED variation.

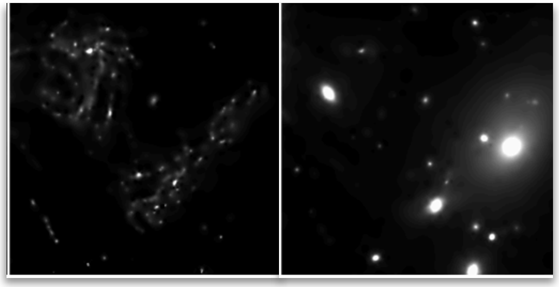
Multichannel data

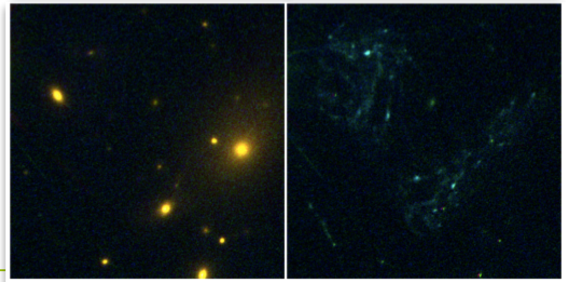


Realistic SEDs variation.

Multichannel data

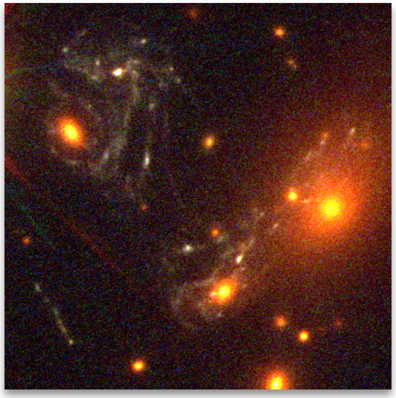


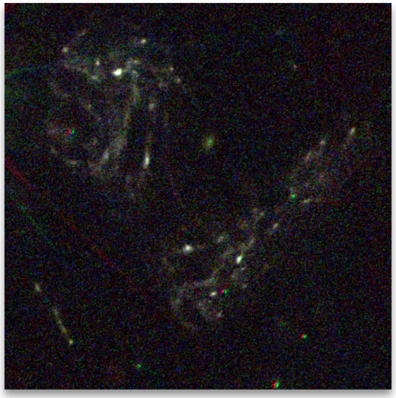


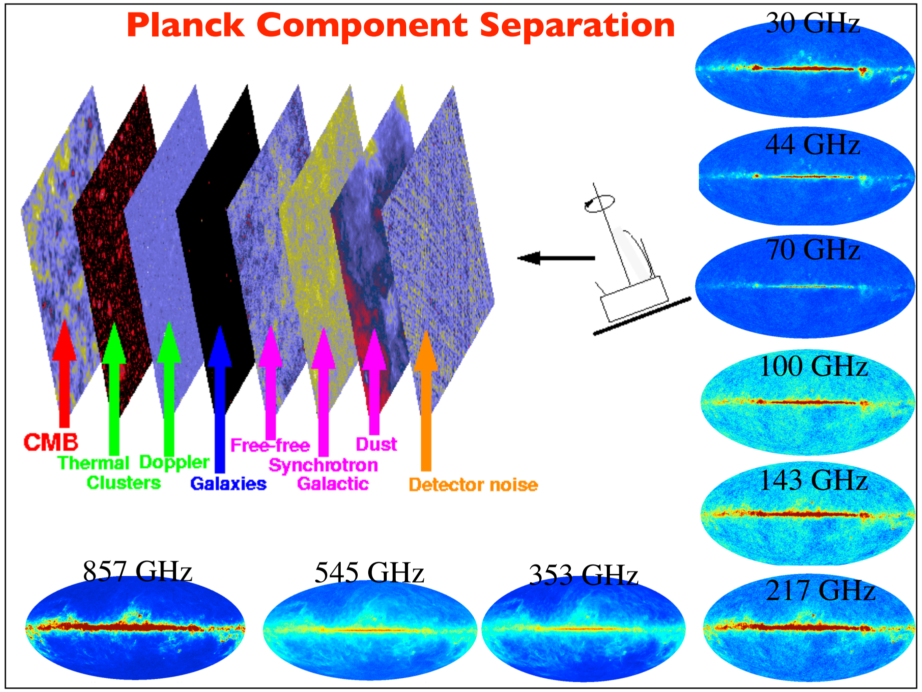


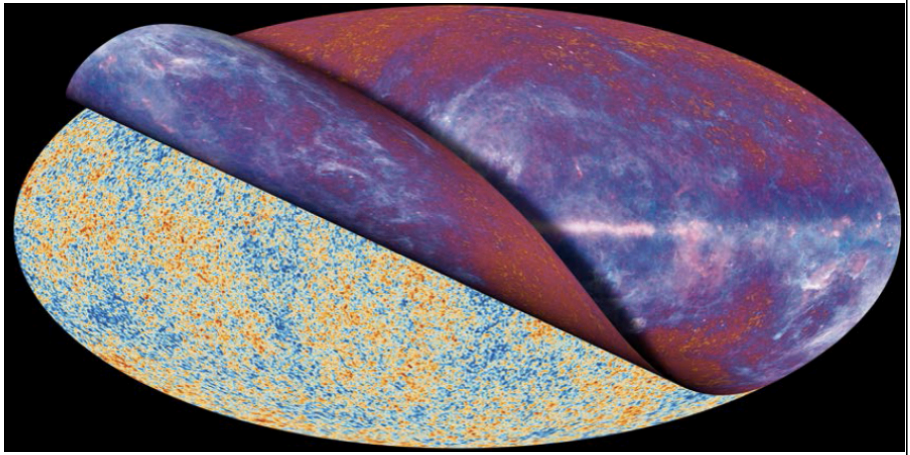
Multichannel data

galaxy cluster MACS~J1149+2223






Planck Component Separation
30 GHz
44 GHz
70 GHz
100 GHz
143 GHz
217 GHz
CMB
Thermal
Clusters
Doppler
Galaxies
Free-free
Synchrotron
Galactic
Dust
Detector noise
857 GHz
545 GHz
353 GHz



Caché dans les autres emissions du ciel

Sparse Component Separation: the GMCA Method

A and S are estimated alternately and iteratively in two steps :

•J. Bobin, J.-L. Starck, M.J. Fadili, and Y. Moudden, "Sparsity, Morphological Diversity and Blind Source Separation", IEEE Trans. o Image Processing, Vol 16, No 11, pp 2662 - 2674, 2007.
•.J. Bobin, J.-L. Starck, M.J. Fadili, and Y. Moudden, "Blind Source Separation: The Sparsity Revolution", Advances in Imaging and Electron Physics , Vol 152, pp 221 -- 306, 2008.

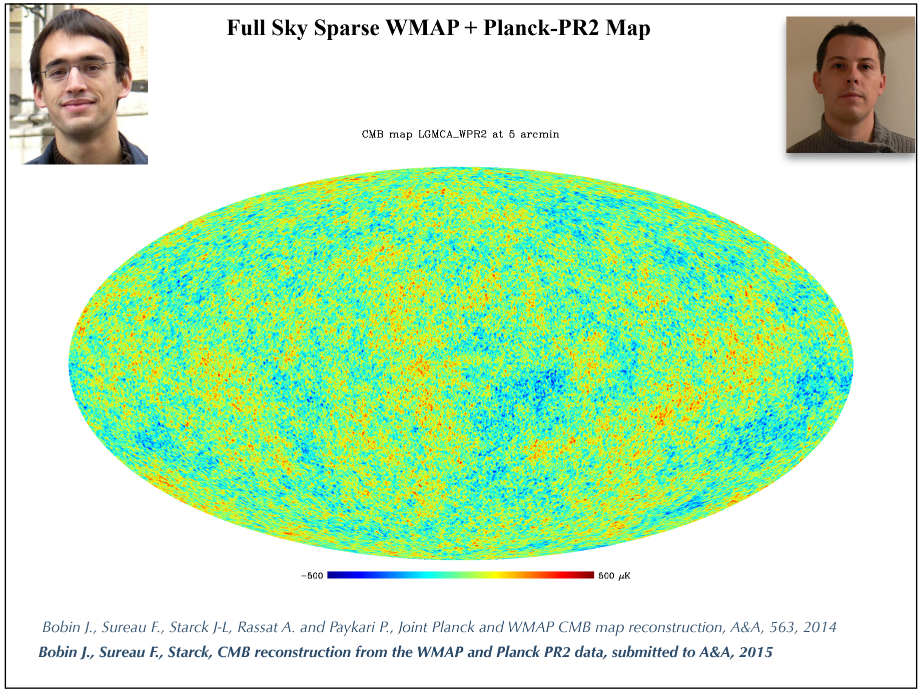
$$\mathbf{X} = \mathbf{AS}$$

1) Estimate S assuming A is fixed (iterative thresholding) :

$$\{S\} = \text{Argmin}_S \sum_j \lambda_j \|s_j \mathbf{W}\|_1 + \|\mathbf{X} - \mathbf{AS}\|^2_{F,\mathbf{\Sigma}}$$

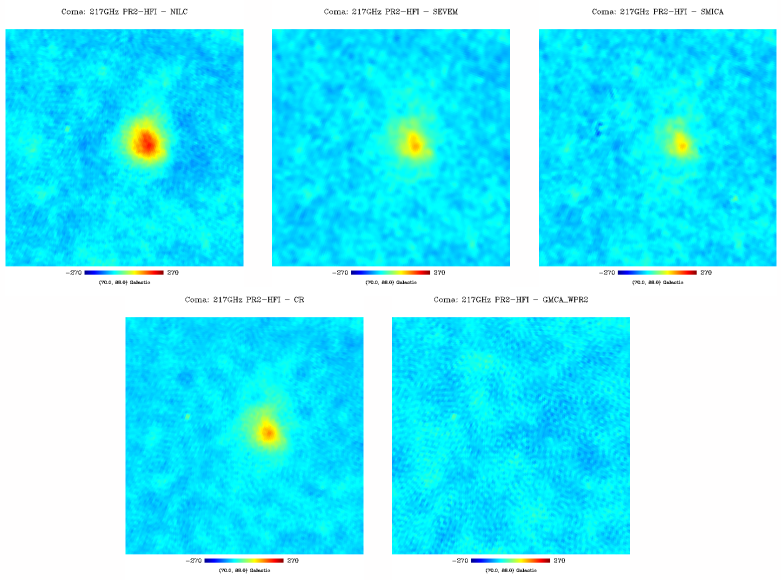
2) Estimate A assuming S is fixed (a simple least square problem) :

$$\{A\} = \text{Argmin}_A \|\mathbf{X} - \mathbf{AS}\|^2_{F,\mathbf{\Sigma}}$$

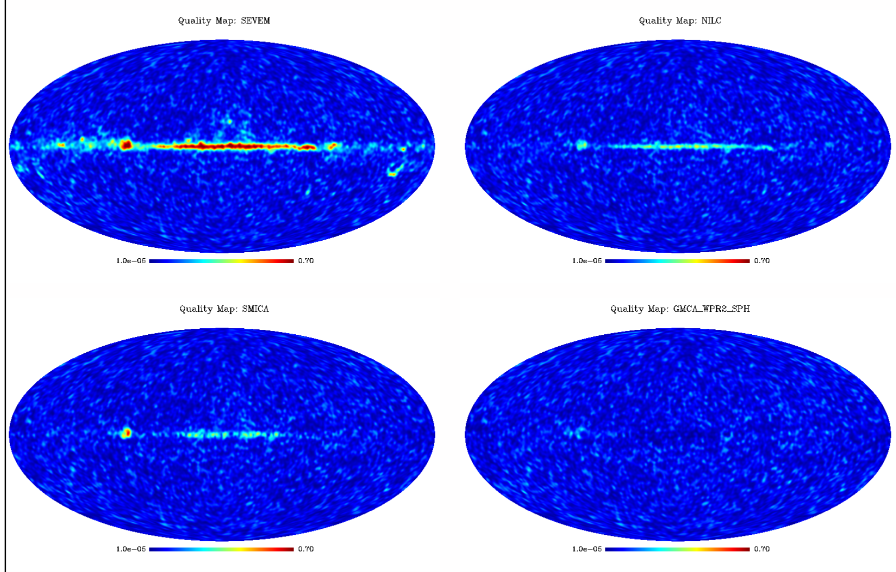


The anisotropies of the Cosmic microwave background (CMB) as observed by Planck. The CMB is a snapshot of the oldest light in our Universe, imprinted on the sky when the Universe was just 380 000 years old. It shows tiny temperature fluctuations that correspond to regions of slightly different densities, representing
Credits: ESA and the Planck Collaboration
the seeds of all future structure: the stars and galaxies of today.
La plus belle carte du fond diffus cosmologique

Traces of tSZ effect



Quality map


Quality Map: SEVEM
1.0e-05
0.70
Quality Map: NILC
1.0e-05
0.70
Quality Map: SMICA
1.0e-05
0.70
Quality Map: GMCA_WPR2_SPH
1.0e-05
0.70

Learning and Blind Source Separation

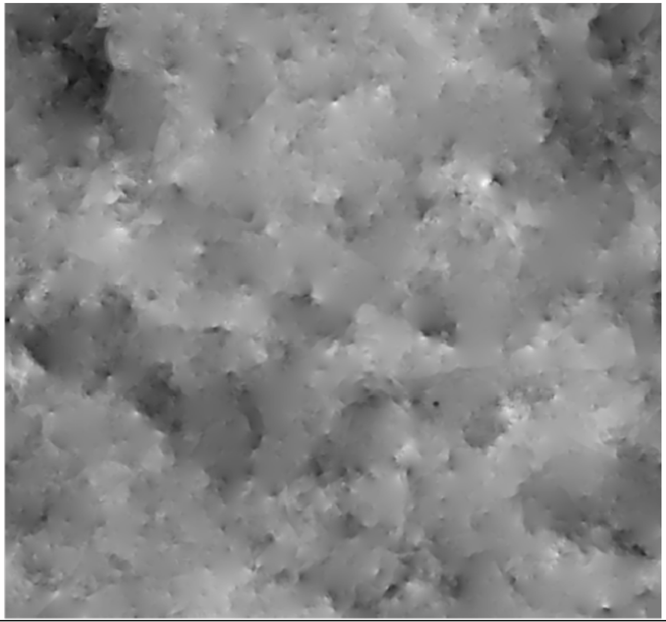
•Part 1: Introduction:

•Part 2: Monochannel Source Separation

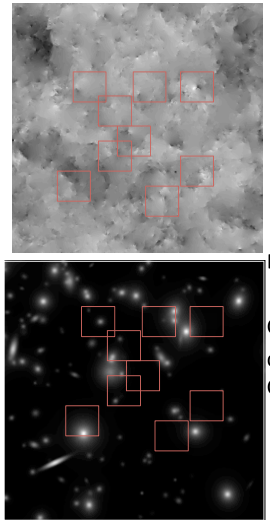
•Part 3: Multichannel Blind Source Separation

•Part 4: Dictionary Learning

Simulated Cosmic String Map



Dictionary Learning



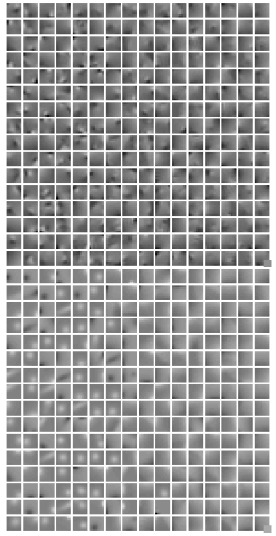
Training basis.

$$\left(\hat{D},\hat{A}\right) = \underset{\substack{D\in C_1 \\ A\in C_2}}{\arg\min}\left(Y = DA\right)$$

DL: Matrix Factorization problem

C_1: Constraints on the Sparsifying dictionary D

C_2: Constraints on the Sparse codes



Dictionary Learning

1 - Introduction to dictionary learning : How ?

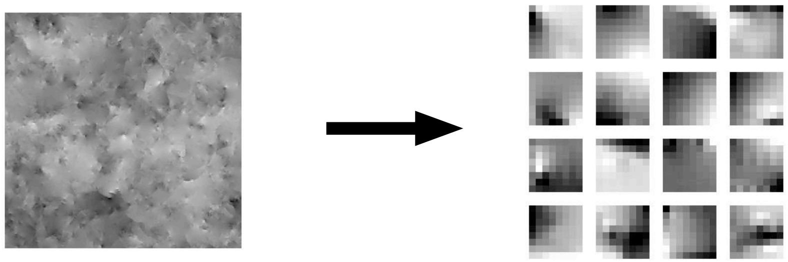
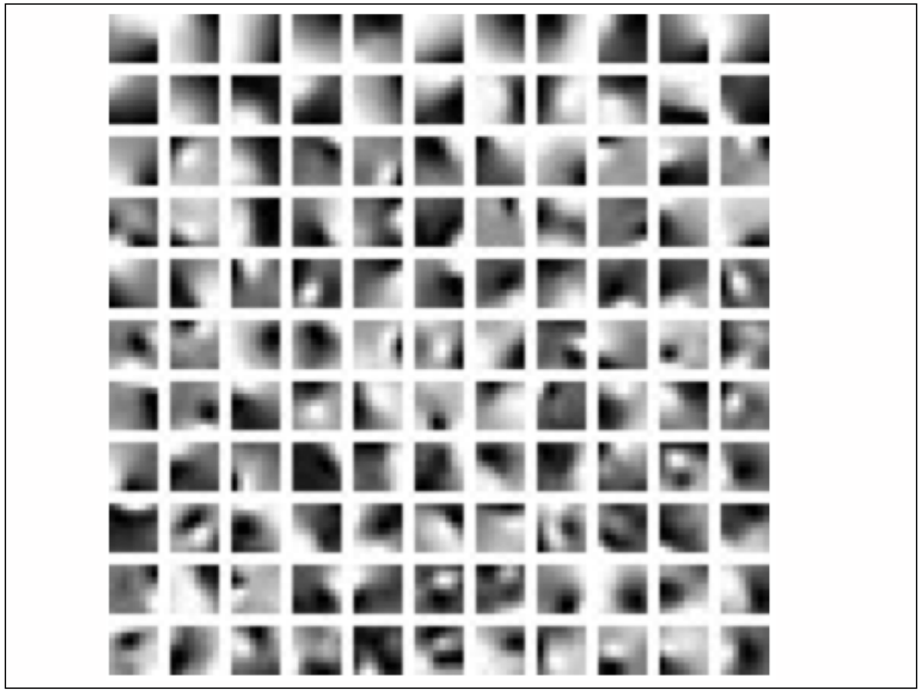


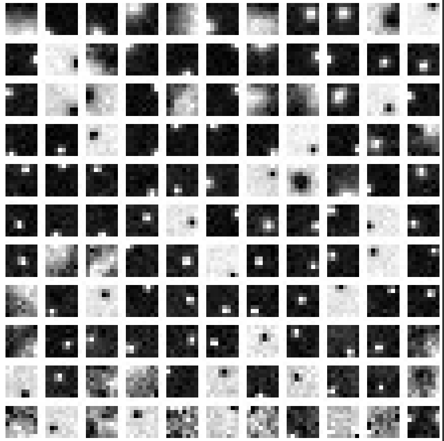
Image to learn from

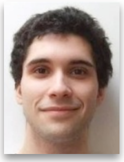
N training patches Y

$$\arg\min_{\mathbf{D},\mathrm{X}_i} \sum_{i=1}^{N} \min_{\mathbf{x}_i} \left\{ \|\mathbf{D}\mathbf{x}_i - \mathbf{y}_i\|^2 + \lambda \|\mathbf{x}_i\|_1 \right\}$$



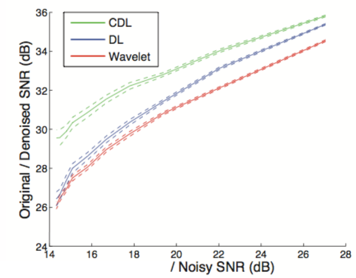
Dictionary Learning

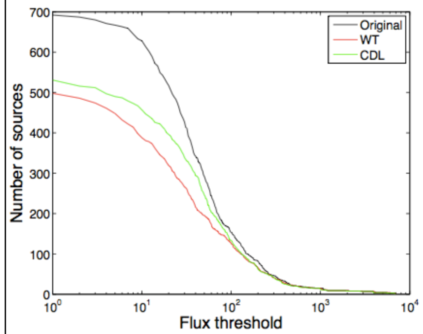


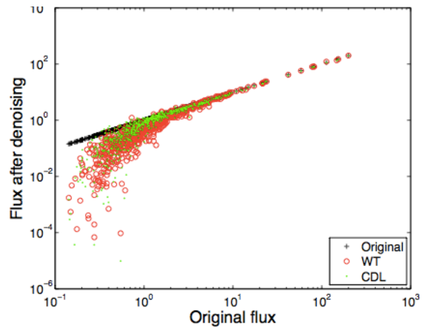


S. Beckouche

Astronomical Image Denoising Using Dictionary Learning, S. Beckouche, J.L. Starck, and J. Fadili, A&A, 556, id.A132, pp 14, 2013.







- ✓ **Sparse Learning techniques are very efficient for**
  - ➡ **Component separation**
    http://www.cosmostat.org/research/statistical-methods/gmca/
  - ➡ **Joint CMB map reconstruction from WMAP and Planck data**
    http://www.cosmostat.org/research/cmb/planck_wpr2/
- ✓ **Reproducible Research**
  http://www.cosmostat.org/software.html
- ✓ **Perspective**
  - ➡ **Extend the sparse component separation to polarized data.**
  - ➡ **Develop sparsity techniques for SKA and Euclid.**