Sparsity and Morphological Diversity in Source Separation

Jérôme Bobin
IRFU/SEDI-Service d'Astrophysique
CEA Saclay - France

Collaborators

- Yassir Moudden - CEA Saclay, France
- Jean-Luc Starck - CEA Saclay, France
- Jalal Fadili - Caen University, France
- Michael Elad - The Technion, Israel Institute of Technology
- David Donoho - Department of Statistics, Stanford University

Outline

Part I

Sparsity and Morphological Diversity in Source Separation

The Classical Mixture Model

Sources Mixtures

Examples :

EEG, ECG, Multispectral imaging, Astrophysical Component separation, etc...

A Short Introduction to Source Separation

Goal :

Estimating n unknown signals (the sources) from m linear mixtures (observations) with $m \geq n$.

$$\mathbf{X} = \sum_{k=1}^{n} a^k s_k + \mathbf{N} = \mathbf{AS} + \mathbf{N}$$

DATA

MIXING MATRIX

SOURCES

NOISE

A strenuous problem !

Both **S** and **A** are unknown

have to be estimated

➡ Blind Source Separation

BSS: A question of diversity

BSS is about devising quantitative measures of diversity to disentangle between the sources

Looking for :

- Decorrelated sources : equivalent to principal component analysis (PCA).

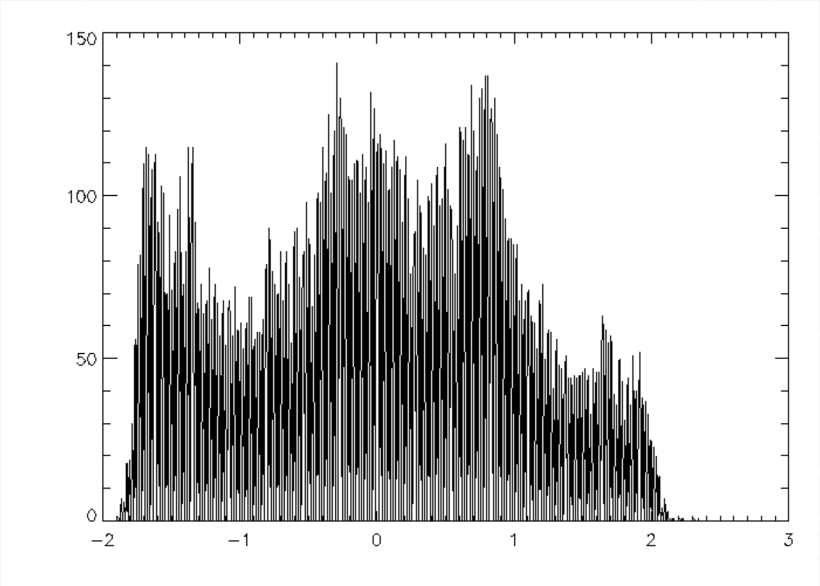
- Independent sources : Independent component analysis (ICA) emphasizes on statistical independence to distinguish between the sources.

Don't provide stable solutions in noisy environment.
Often strongly dependent on sources' statistical model.

Sparsity or how to better distinguish between the signals

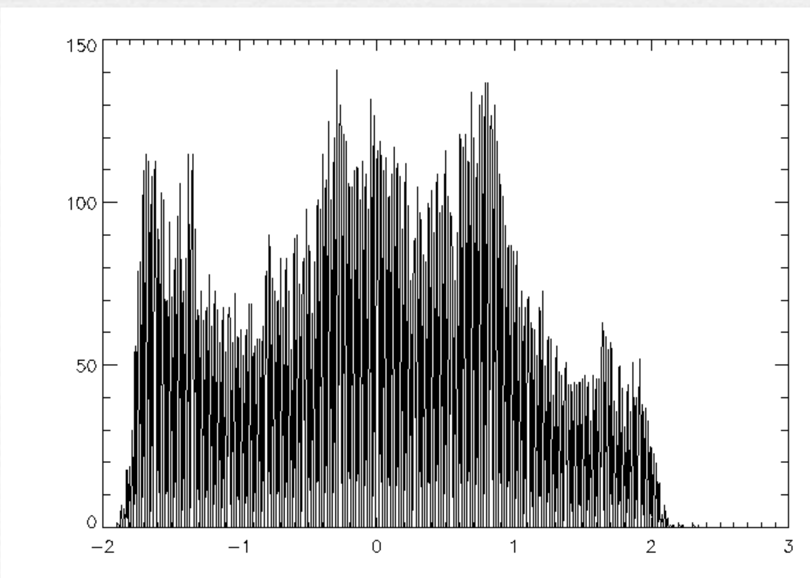
Sparsity or how to better distinguish between the signals

Pixel Domain

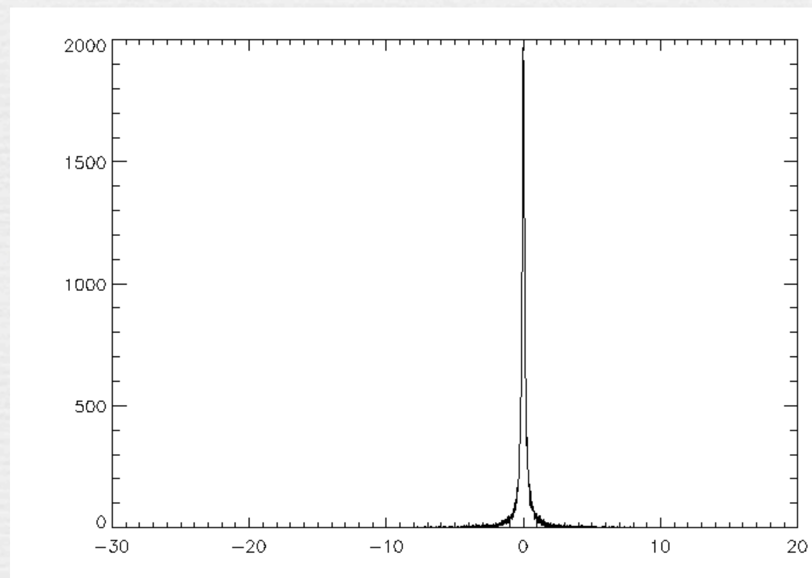


Sparsity or how to better distinguish between the signals

Pixel Domain



Curvelet Domain



How to choose a representation ?

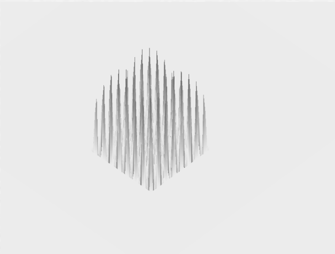
Basis

How to choose a representation ?

~~Basis~~ Dictionary

How to choose a representation ?

~~Basis~~ Dictionary



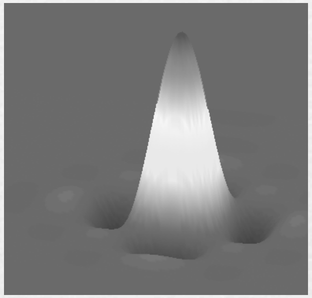
Local DCT

How to choose a representation ?

~~Basis~~ Dictionary

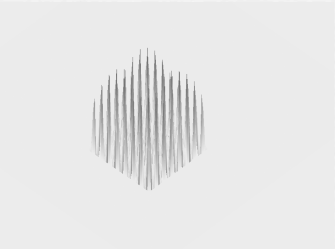
Local DCT

Wavelets



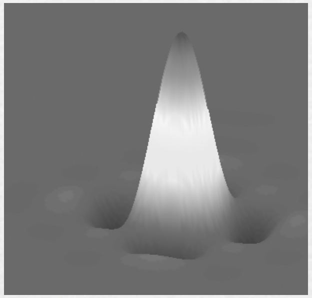
How to choose a representation ?

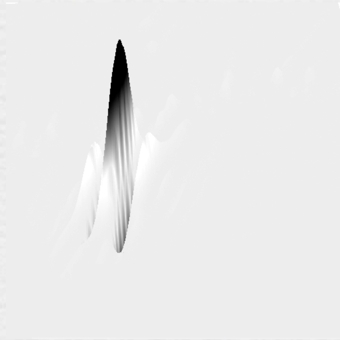
~~Basis~~ → Dictionary



Local DCT

Wavelets

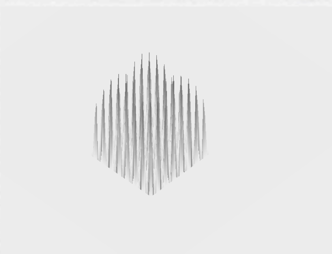




Curvelets

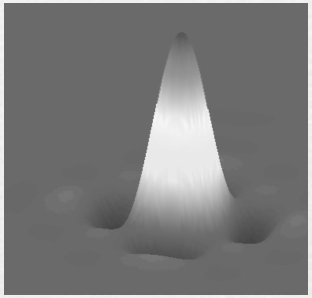
How to choose a representation ?

~~Basis~~ → Dictionary



Local DCT

Wavelets



Curvelets

Others

A question of diversity

s_1

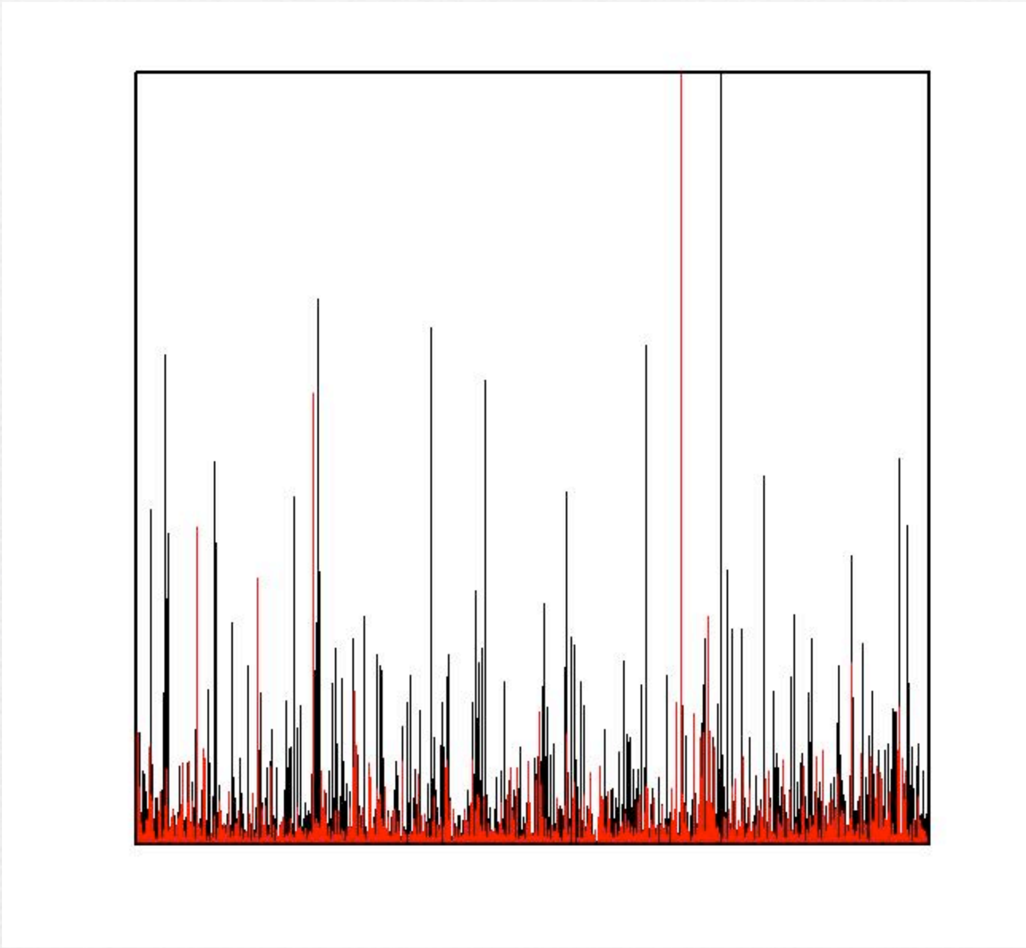
s_2

Projection coefficients in a wavelet basis

A question of diversity

s_1

s_2

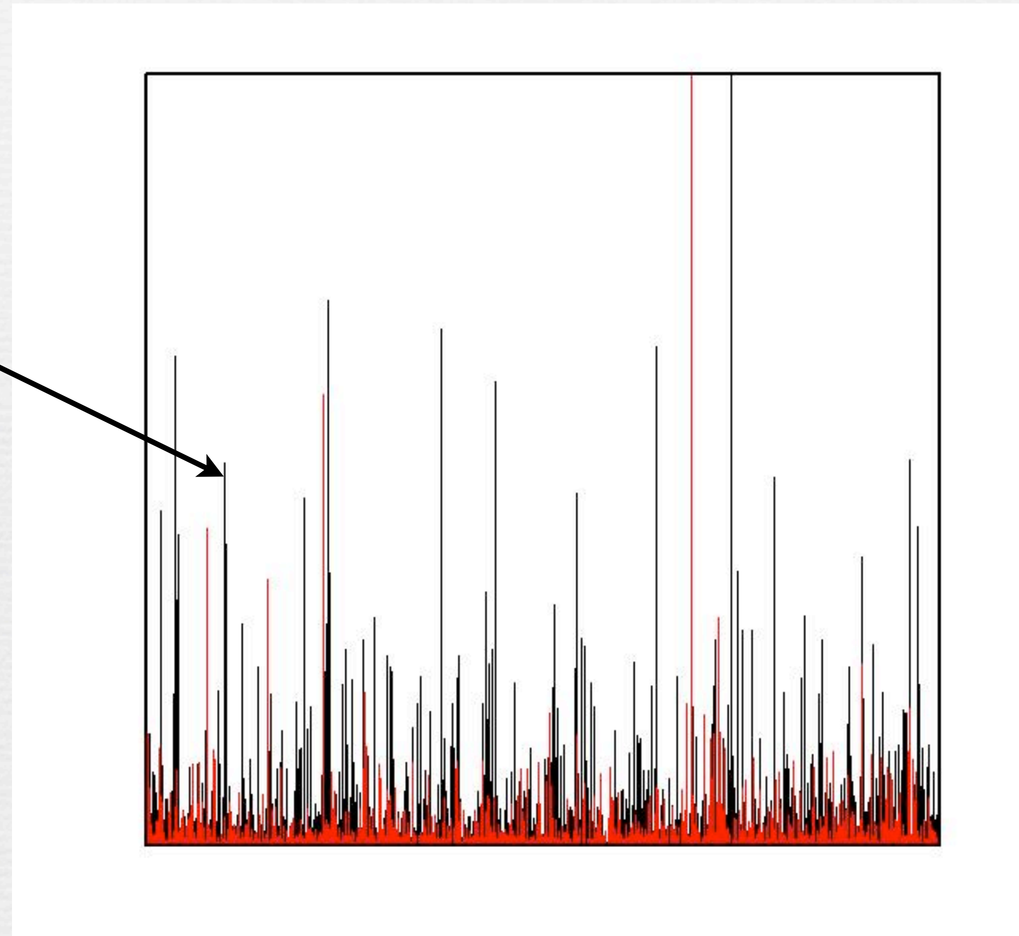


Projection coefficients in a wavelet basis

A question of diversity

s_1

s_2

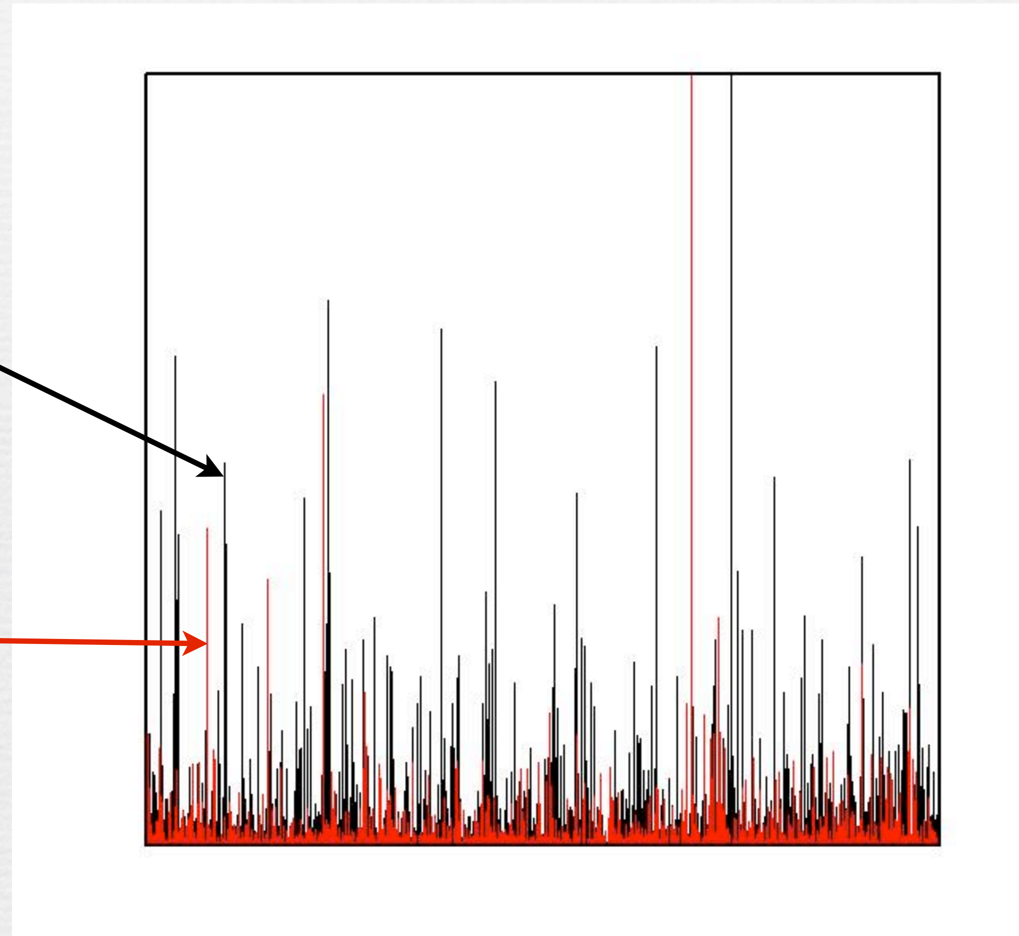


Projection coefficients in a wavelet basis

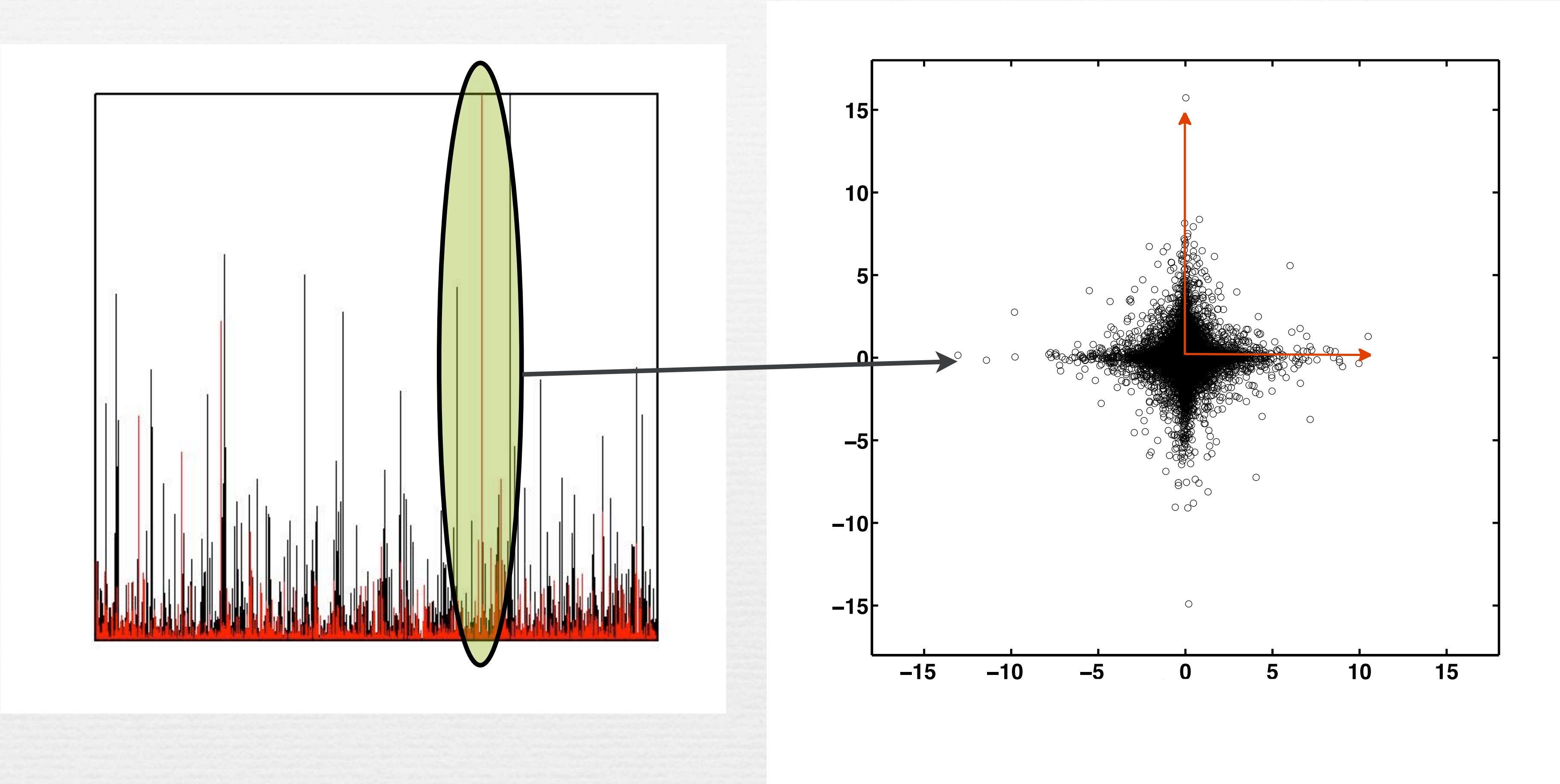
A question of diversity

s_1

s_2

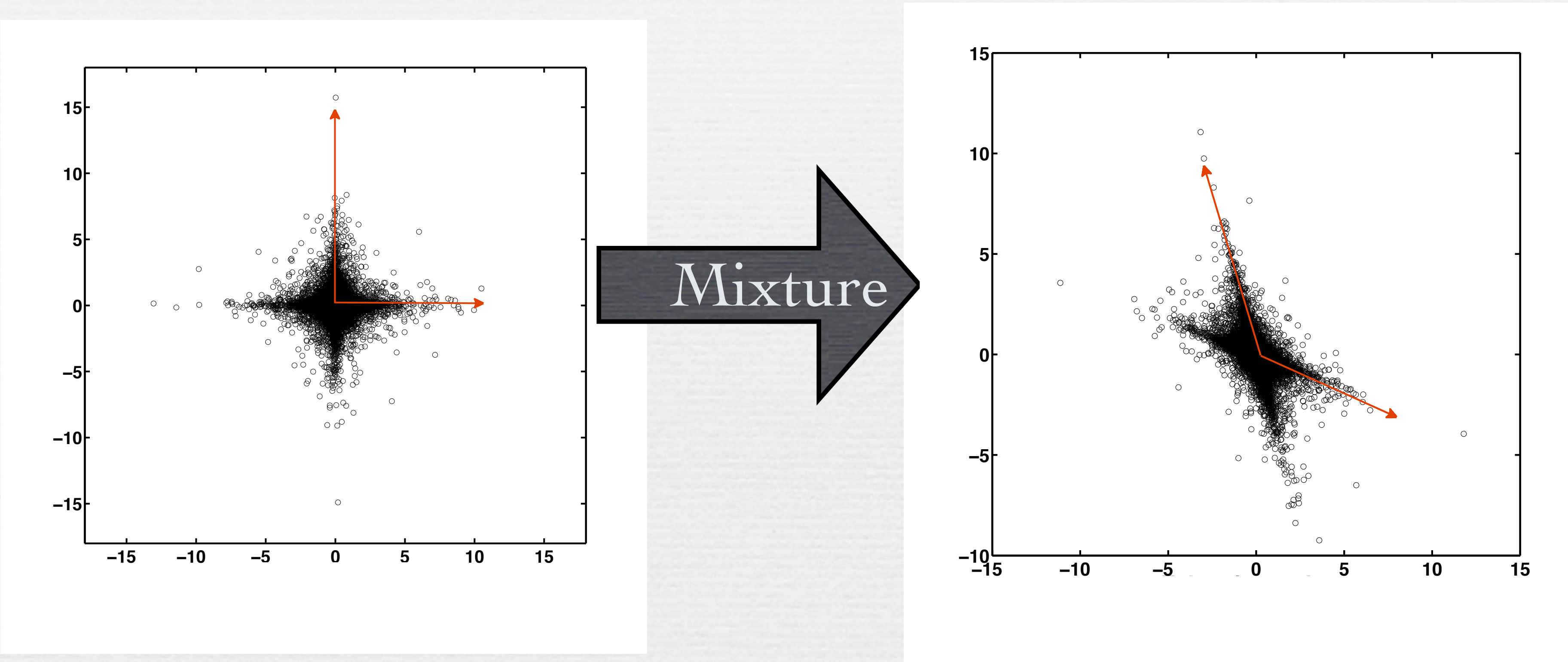


Projection coefficients in a wavelet basis



<u>Morphological diversity :</u> signals with different morphologies have disjoint significant coefficients in a sparsifying dictionary.

What about mixtures ?



Mixed signals are likely to be less sparse

Part II

Generalized Morphological Component Analysis

Generalized Morphological Component Analysis (GMCA)

Estimating A and S such that the sources are sparse in Φ

$$\min_{\mathbf{A},\mathbf{S}} \|\alpha\|_{\ell_1} \;\; \text{s.t.} \;\; \|\mathbf{X} - \mathbf{AS}\|_{\ell_2} < \epsilon \quad ; \mathbf{S} = \alpha \boldsymbol{\Phi}$$

Sparsity
"measure"

Solved via an iterative thresholding algorithm

A simple experiment

Original Sources



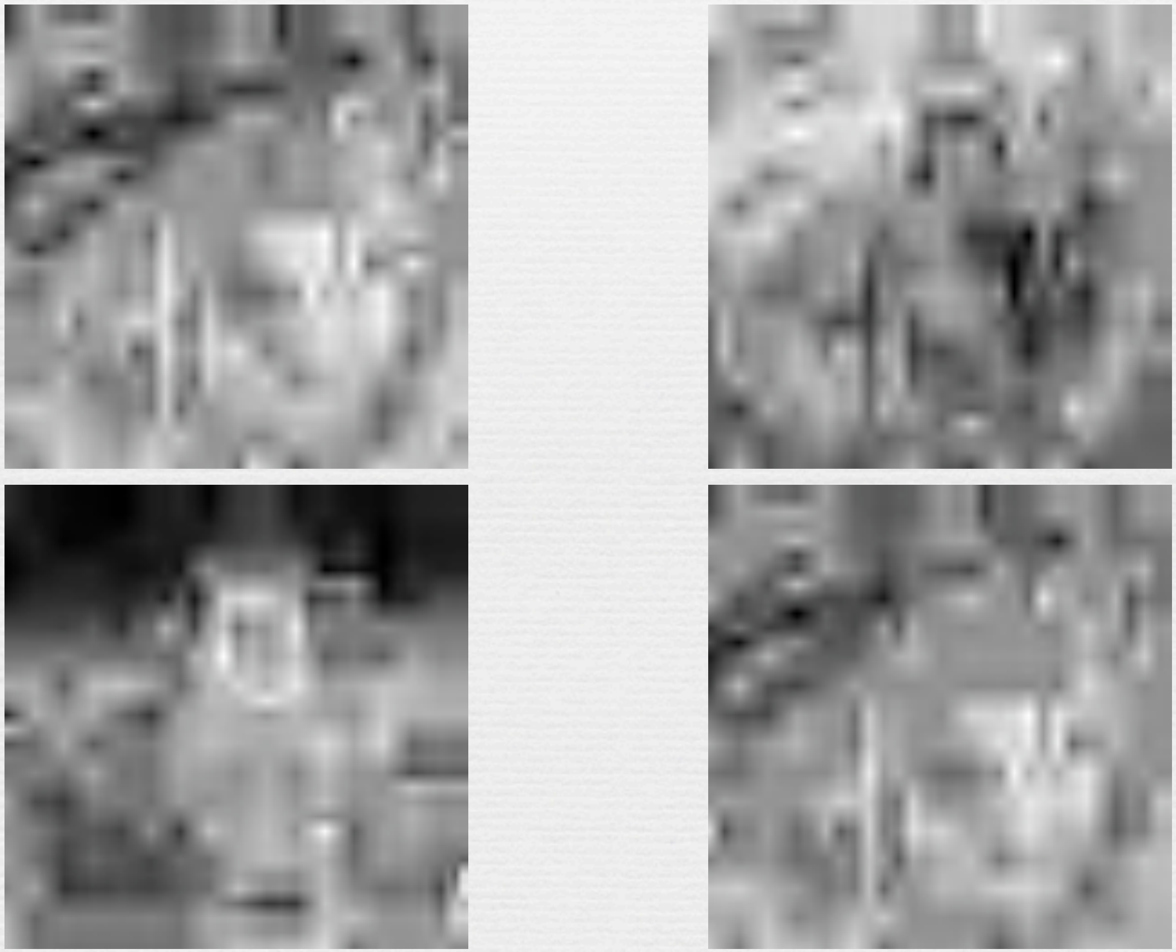

Mixtures


Noiseless experiment, 4 random mixtures, 4 sources

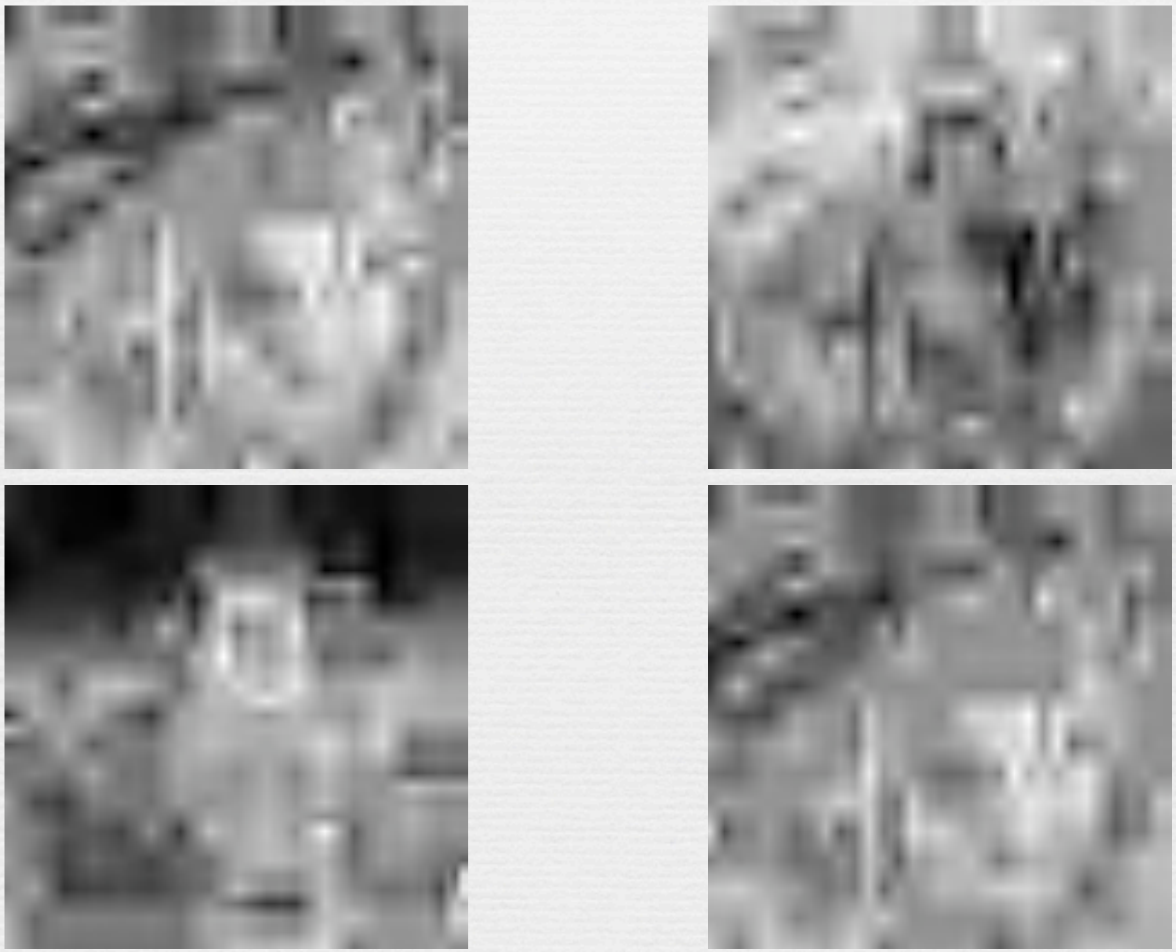
Single Wavelet basis

Result

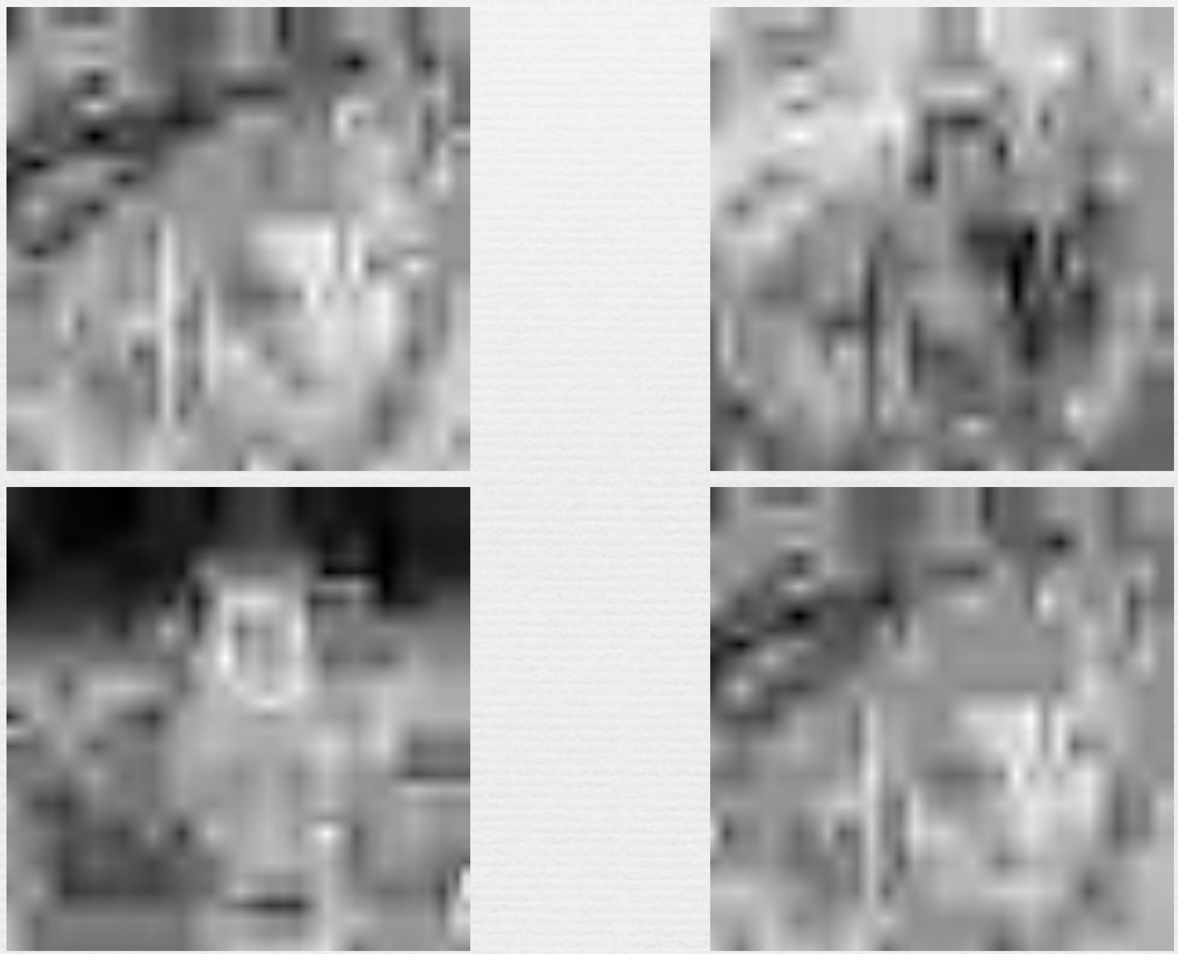
Result



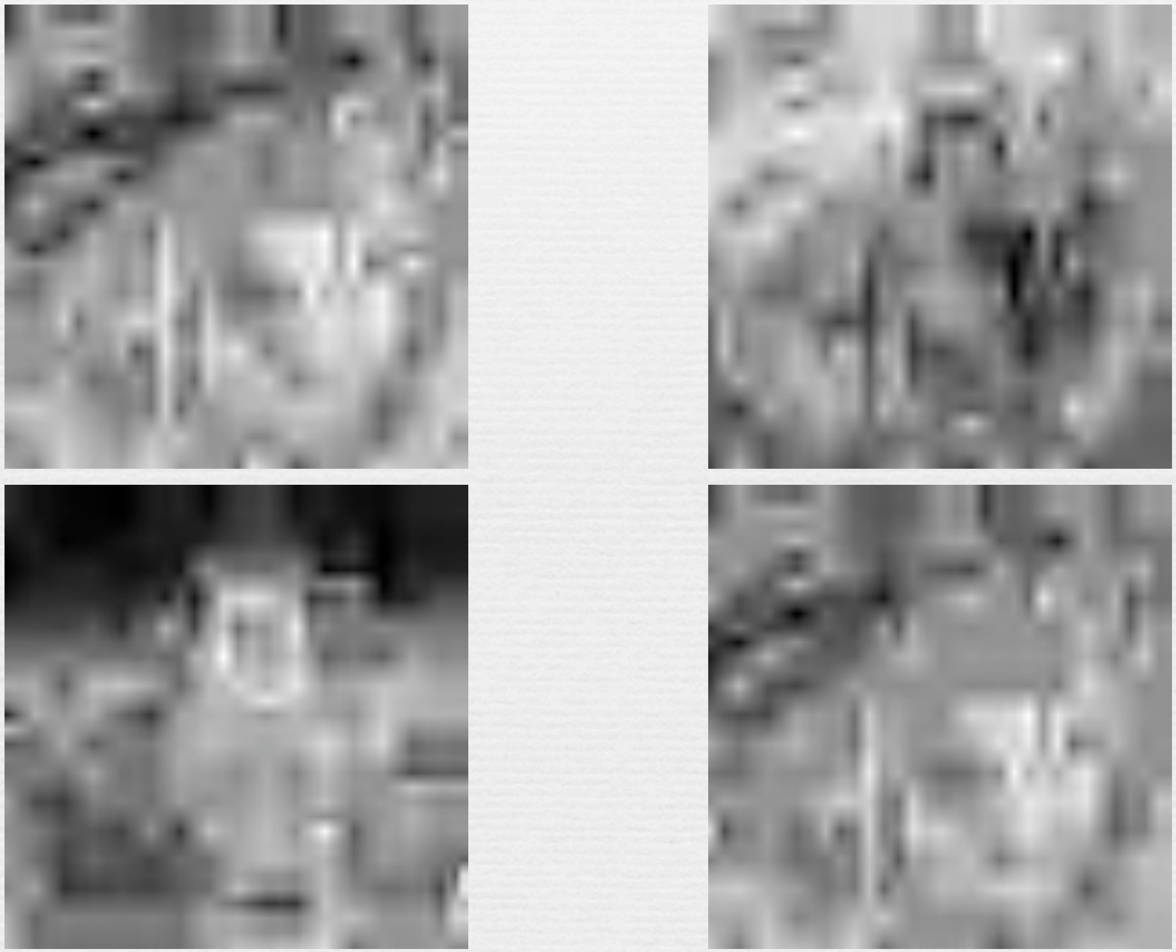
Result



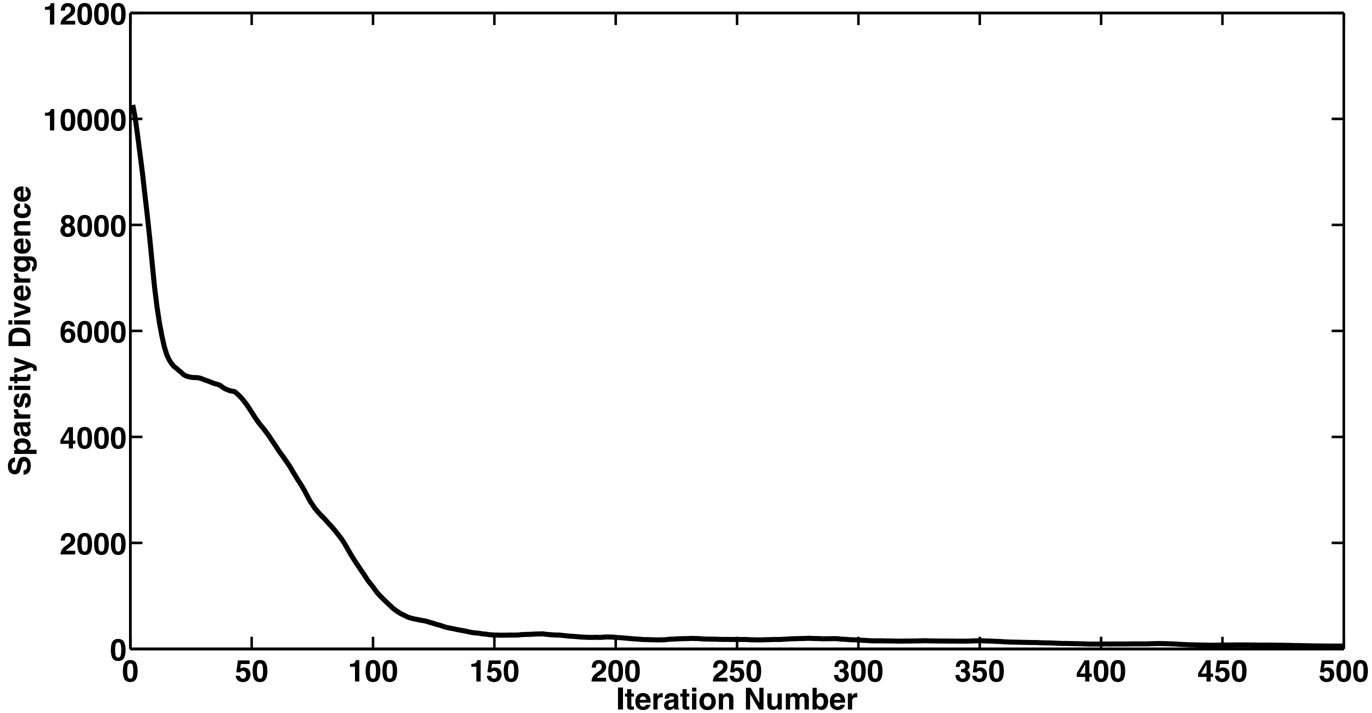
Result



Result



The solution is indeed sparser



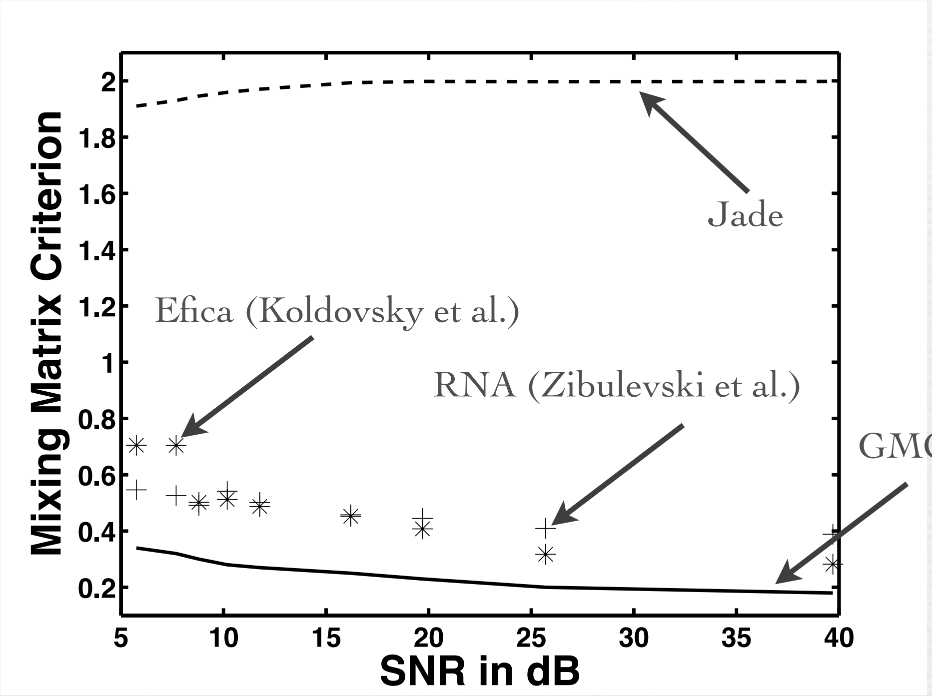
Sparsity divergence : $\|\tilde{\alpha}\|_1 - \|\alpha\|_1$

GMCA is robust to noise

Random mixtures with additive Gaussian noise

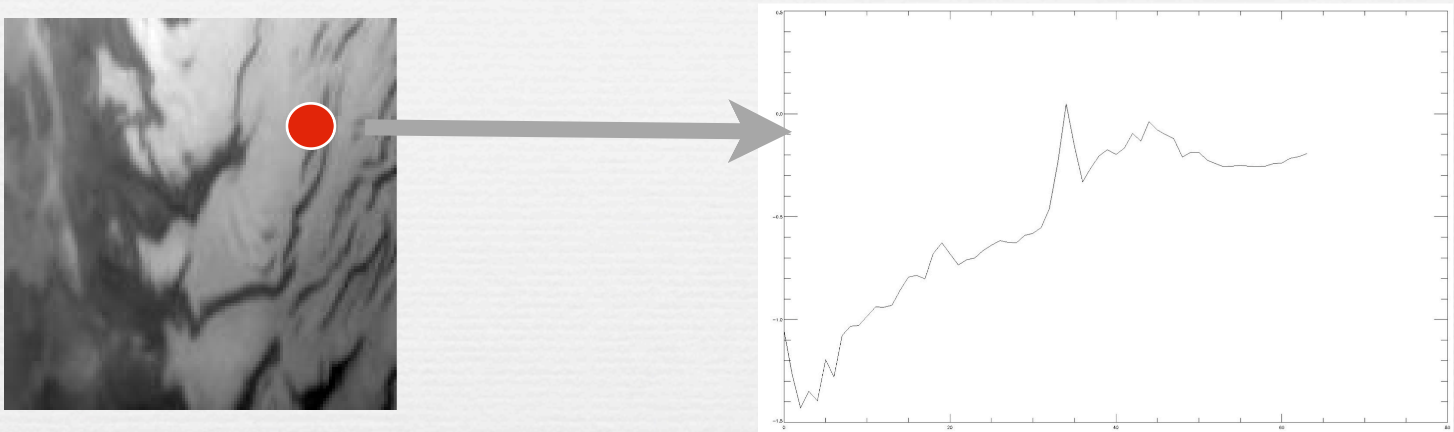
The mixing matrix criterion measures a deviation between the true mixing matrix and its estimate

$\mathbf{\Phi}$: Curvelets + Local DCT



GMCA provide more robust solutions in the presence of noise

Extension to Spatial and Spectral Sparsity Constraints



Extending GMCA to better account for spatial and spectral structures

GMCA with spectral and spatial sparsity constraints

$$\min_{\{a^i, s_i\}} \sum_{i=1}^{n} \left\| \mathbf{\Xi}^T a^i s_i \mathbf{\Phi}^T \right\|_1 \quad \text{s.t.} \quad \mathbf{X} = \sum_{i=1}^{n} a^i s_i$$

Spectral representation

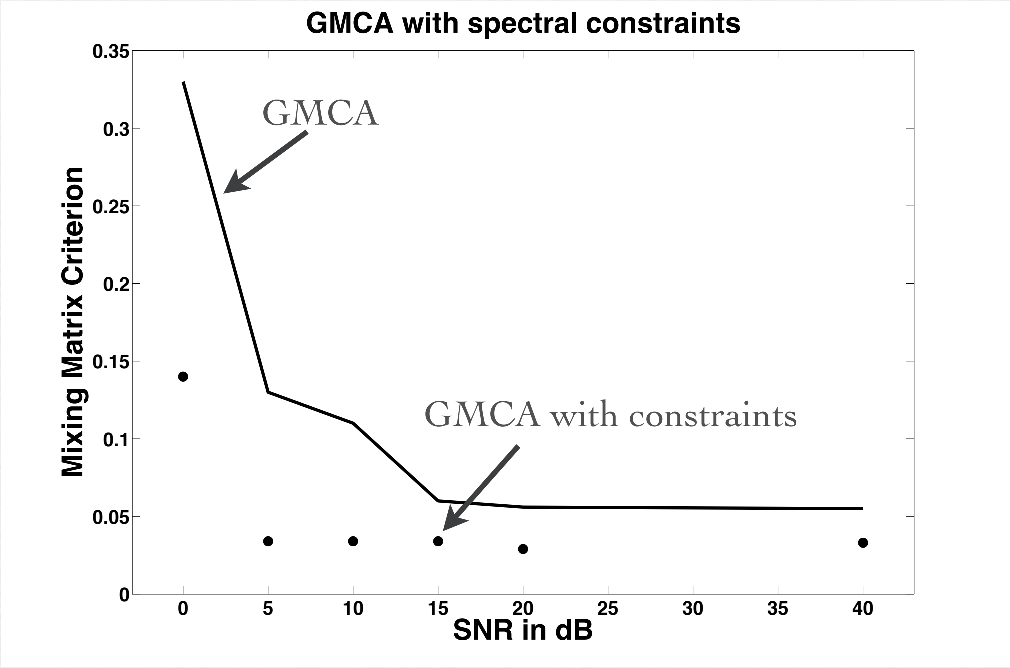
Spatial representation

It amounts to decomposing the data X into the linear combination of n rank-1 matrices that are sparse in $\mathbf{\Xi} \otimes \mathbf{\Phi}$

Preliminary Results

64 Channels, additive Gaussian noise
Each source is a 128x128 image
Random laplacian mixing matrix in the wavelet domain

Sparse representation - spatial : curvelets - spectral : wavelets



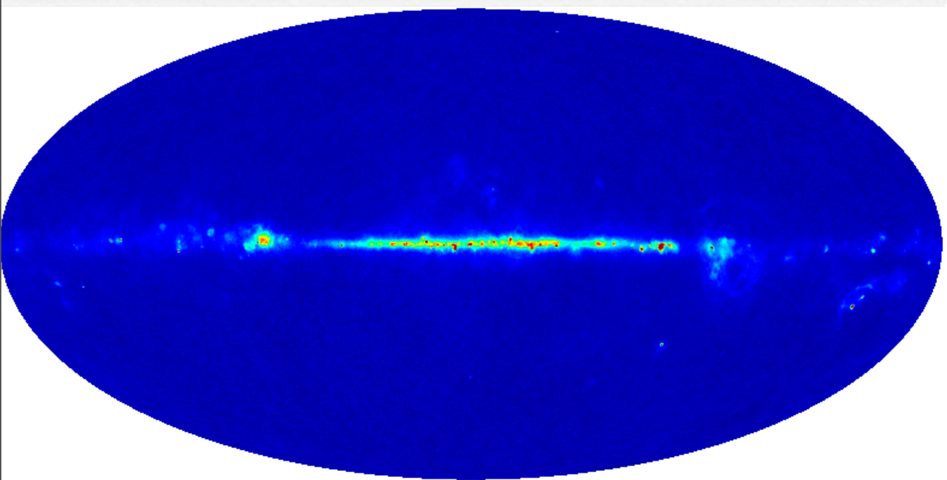
Collaboration with the CESR/Toulouse on hyperspectral Mars Observer data.

Application to Planck data

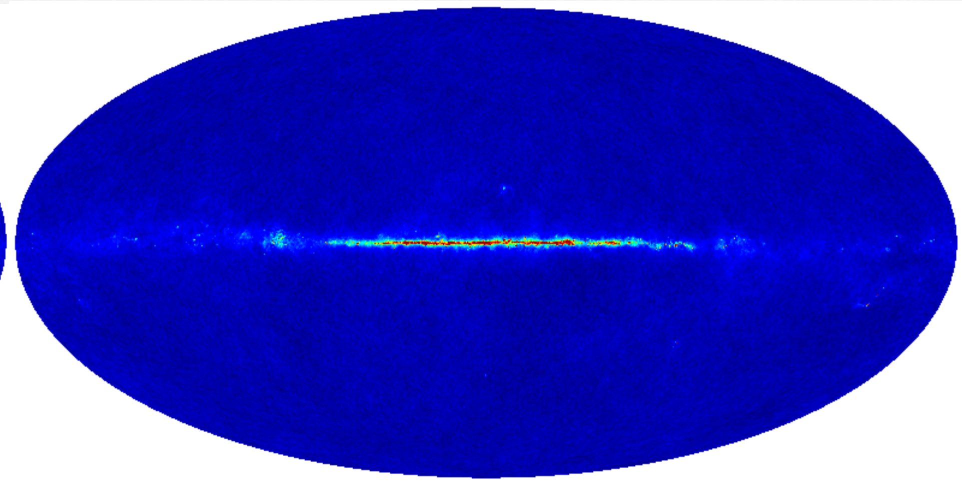
- 9 Channels : 30GHz to 857GHz
- Mixtures of : CMB, SZ, Dust, Synchrotron, Free-Free, ...

⇒ The mixture model is no more valid (Dust and Synchrotron)
GMCA is applied locally in the wavelet domain

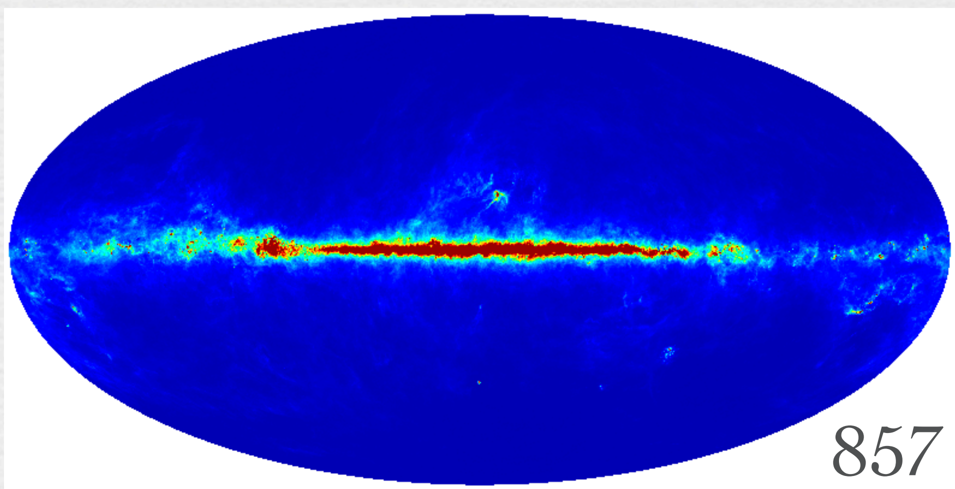
Planck data



30 GHz

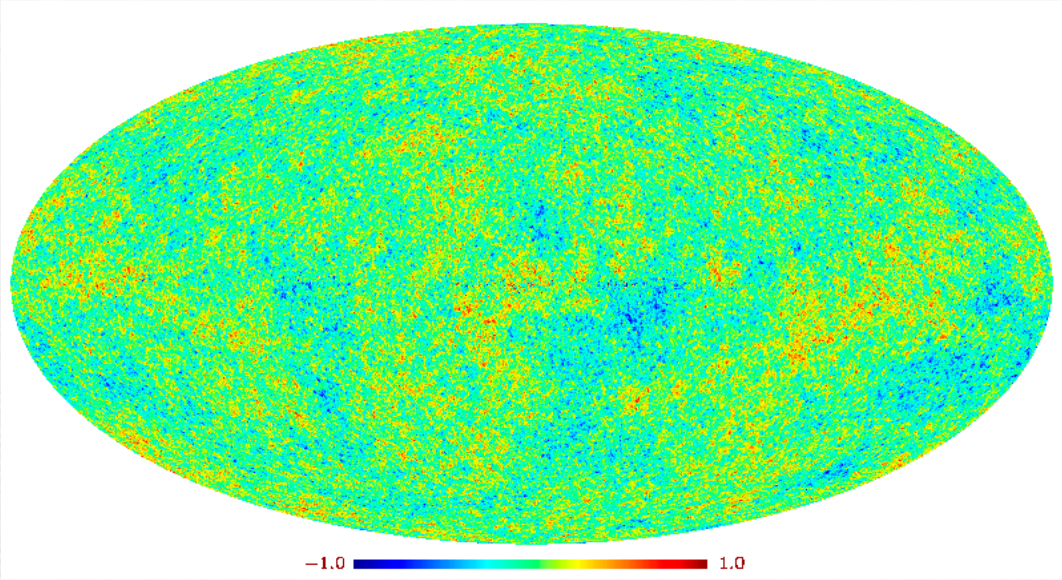


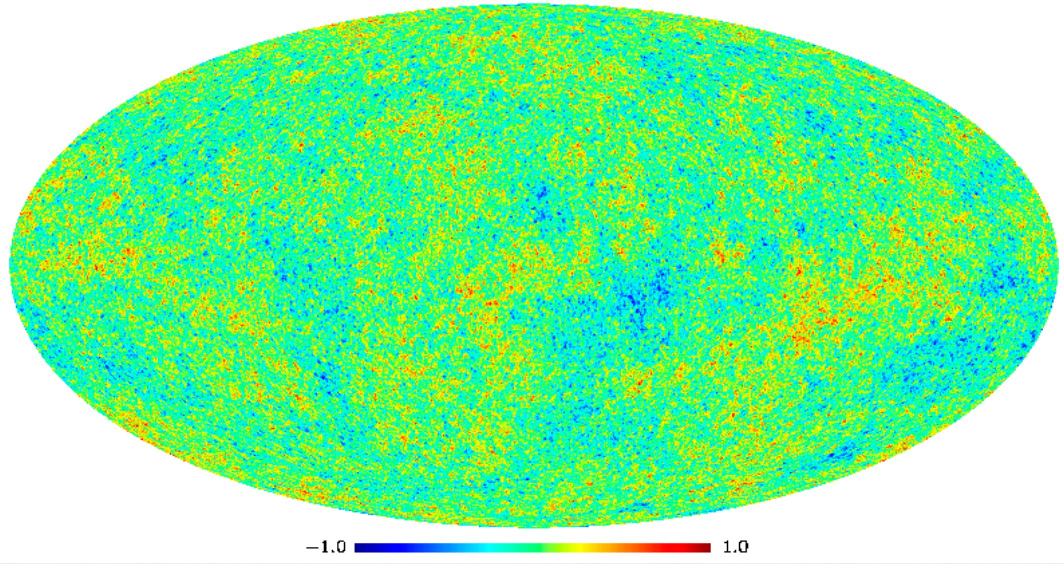
143 GHz



857 GHz

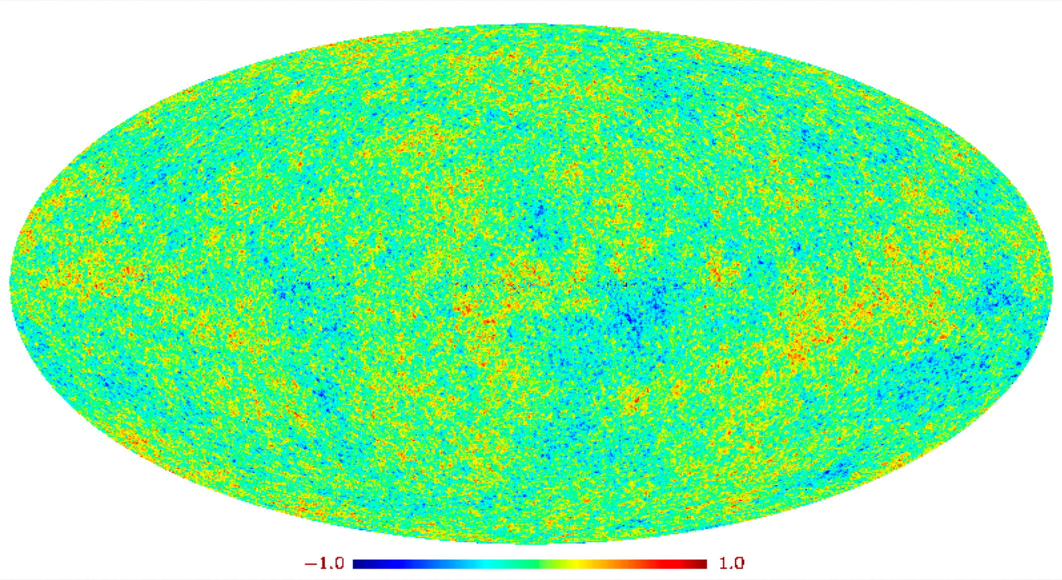
What about CMB ?

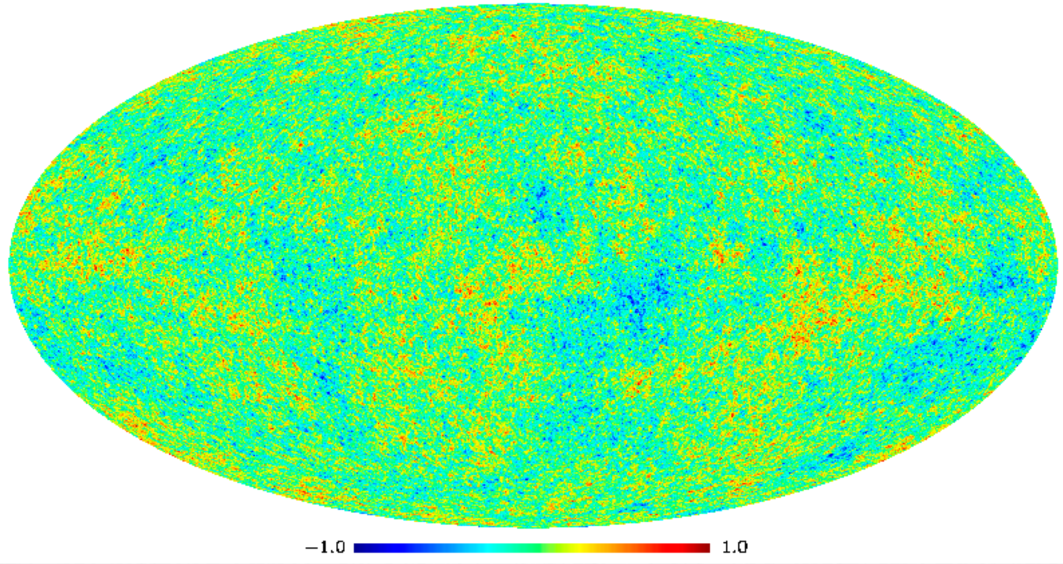




What about CMB ?

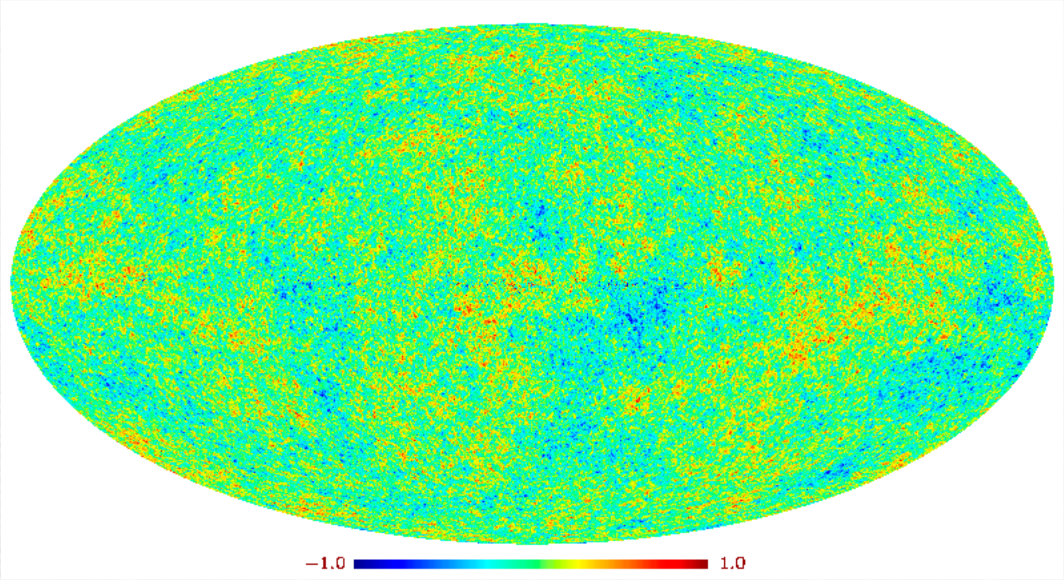
Estimate



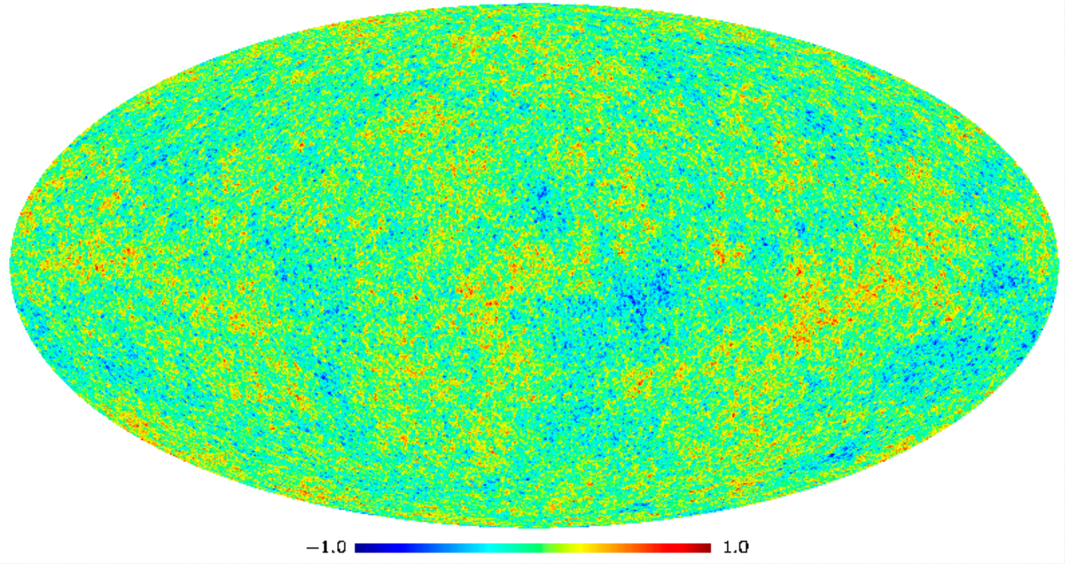


What about CMB ?

Estimate



Input Map



Part - III

Extensions to a wider range of inverse problems

The strength of GMCA

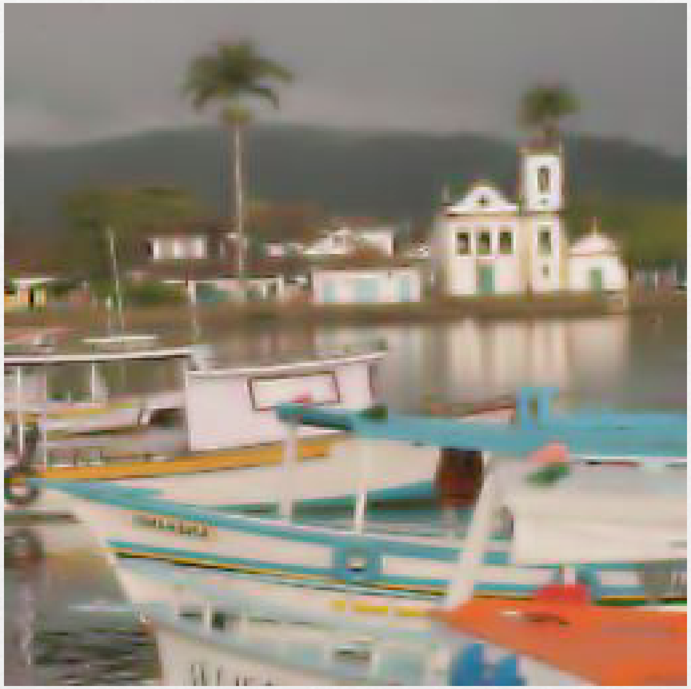
GMCA is able to find a multichannel representation in which signals are **jointly** the sparsest

GMCA framework can then be the backbone of new adaptive sparsity-based techniques in a wide range of multichannel inverse problems:

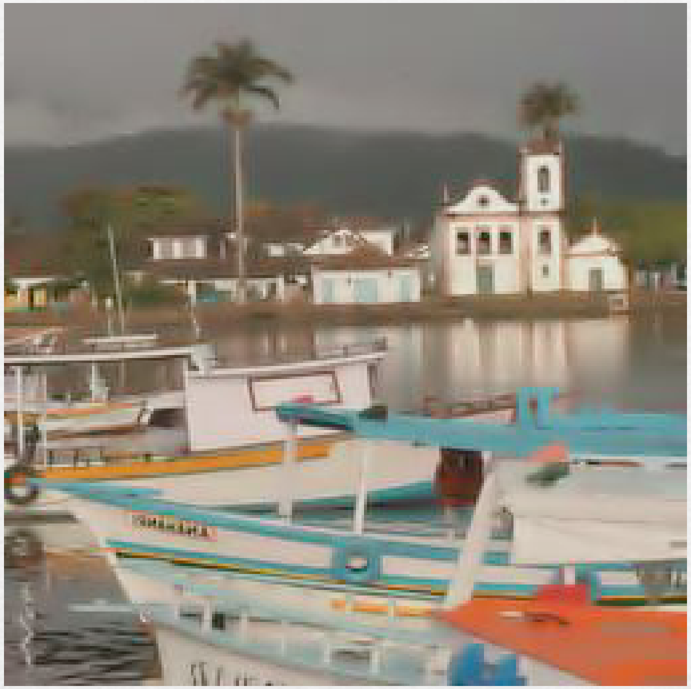
Denoising, deconvolution, inpainting, super-resolution, ...etc.

Adaptive multichannel denoising

SNR = 15 dB



Wavelet denoising in the RGB space

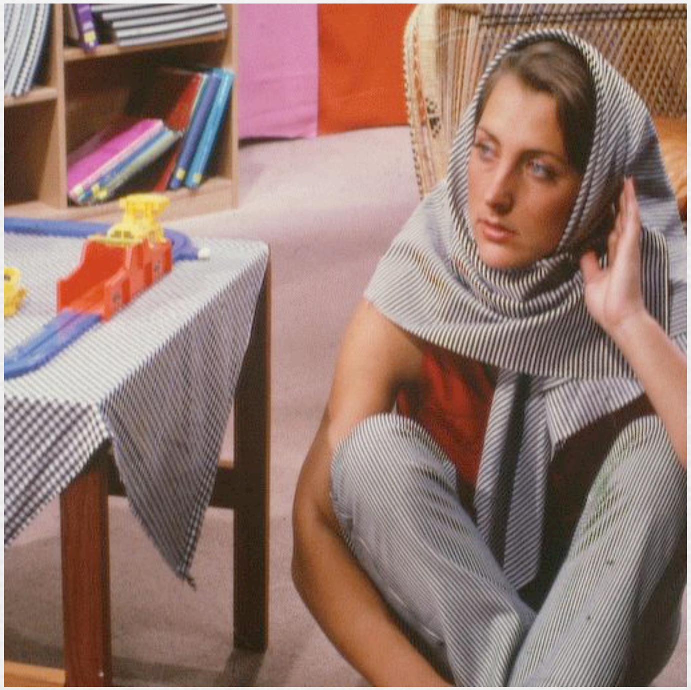


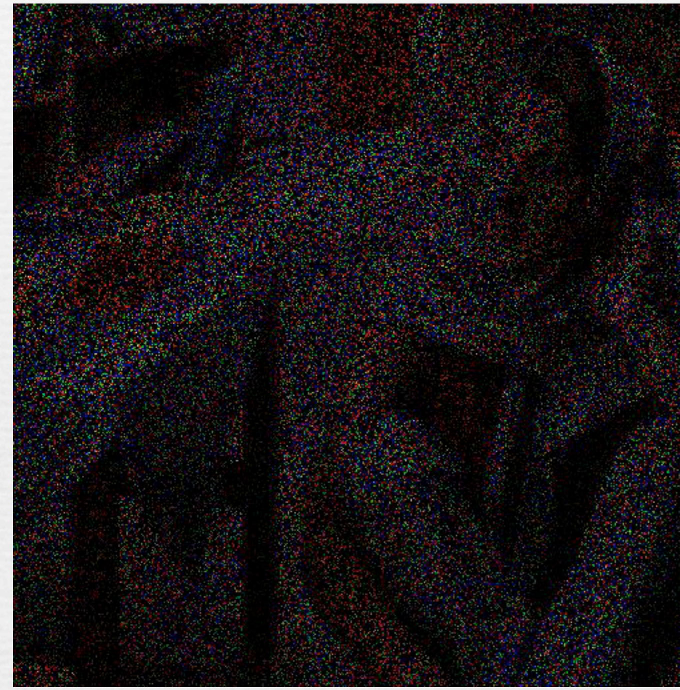
Denoising in the adaptive GMCA-Wavelet space

SNR enhancement up to 1.5 dB

- J.Bobin, J.-L. Starck, J. Fadili, Y.Moudden,"Sparsity and Morphological Diversity in Blind Source Separation", IEEE Transactions on Image Processing, Vol.16, N°11, p. 2662 - 2674, November 2007.

Adaptive multichannel inpainting





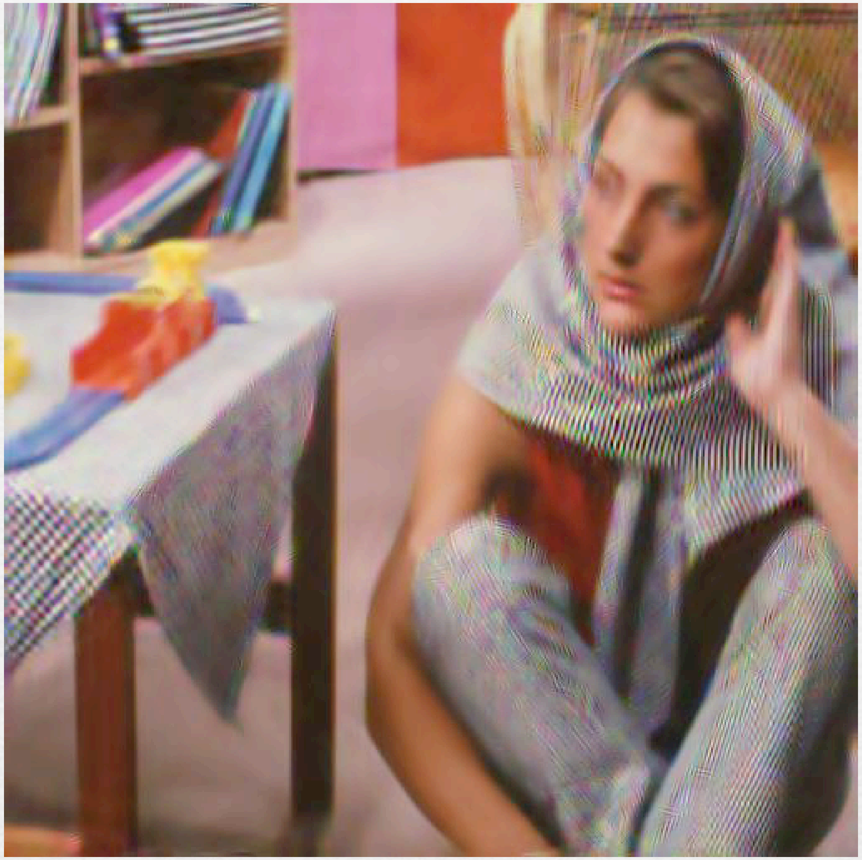
90% of the "color" pixels are missing !

Adaptive multichannel inpainting

$$\min_{\{a^i, s_i\}} \sum_i \|\alpha_i\|_{\ell_1} \;\text{ s.t. }\; \left\| \mathbf{X} - \mathcal{M} \odot \sum_{i=1}^{n} a^i s_i \right\|_{\ell_2} < \epsilon$$

Where $\mathcal{M}$ is a binary mask $\quad s_i = \alpha_i \mathbf{\Phi}$

Φ = Curvelet + Local DCT



MCA-based solution

GMCA-based solution

SNR enhancement up to 1 dB

- *J.Bobin, Y.Moudden, J. Fadili, J.-L. Starck,"Morphological Diversity and Sparsity for Multichannel Data Restoration", Journal of Mathematical Imaging and Vision, 2007 - in press.*

Take-away messages

- **A general framework :** GMCA provides a general framework for multi-valued data analysis.

- **Flexibility :** flexible framework with a wide range of potential extensions (multi/hyper-spectral component separation, joint estimation of the number of components, ...).

- **Applicability :** already applied to Astrophysical component separation; we are looking forward to applying GMCA to hyper-spectral imaging (Mars Observer data) and biomedical data analysis (NMR and Mass spectroscopy).

Going further ...

A software - GMCALab - is also available at http://pagesperso-orange.fr/jbobin/

- *J.Bobin, J.-L. Starck, Y.Moudden, J. Fadili, "Blind Source Separation: the Sparsity Revolution", submitted to Advances in Imaging and Electron Physics - in press*
- *J.Bobin, J.-L. Starck, J. Fadili, Y.Moudden,"Sparsity and Morphological Diversity in Blind Source Separation", IEEE Transactions on Image Processing, Vol.16, N°11, p. 2662 - 2674, November 2007.*
- *J. Bobin, Y. Moudden, J.-L. Starck and M. Elad, "Morphological Diversity and Source Separation", IEEE Signal Processing Letters, Vol.13, N°7, p. 409-412, July 2006.*

Application to the ESA Planck Mission

- *J.Bobin, Y.Moudden, J.-L. Starck, J. Fadili, N. Aghanim, "SZ and CMB reconstruction using GMCA", Statistical Methodology - Special Issue on Astrostatistics - in press - 2007.*

More experiments available at : www.morphologicaldiversity.org