Méthode de séparation de sources
Modèles et algorithmes

Applications en Astrophysique

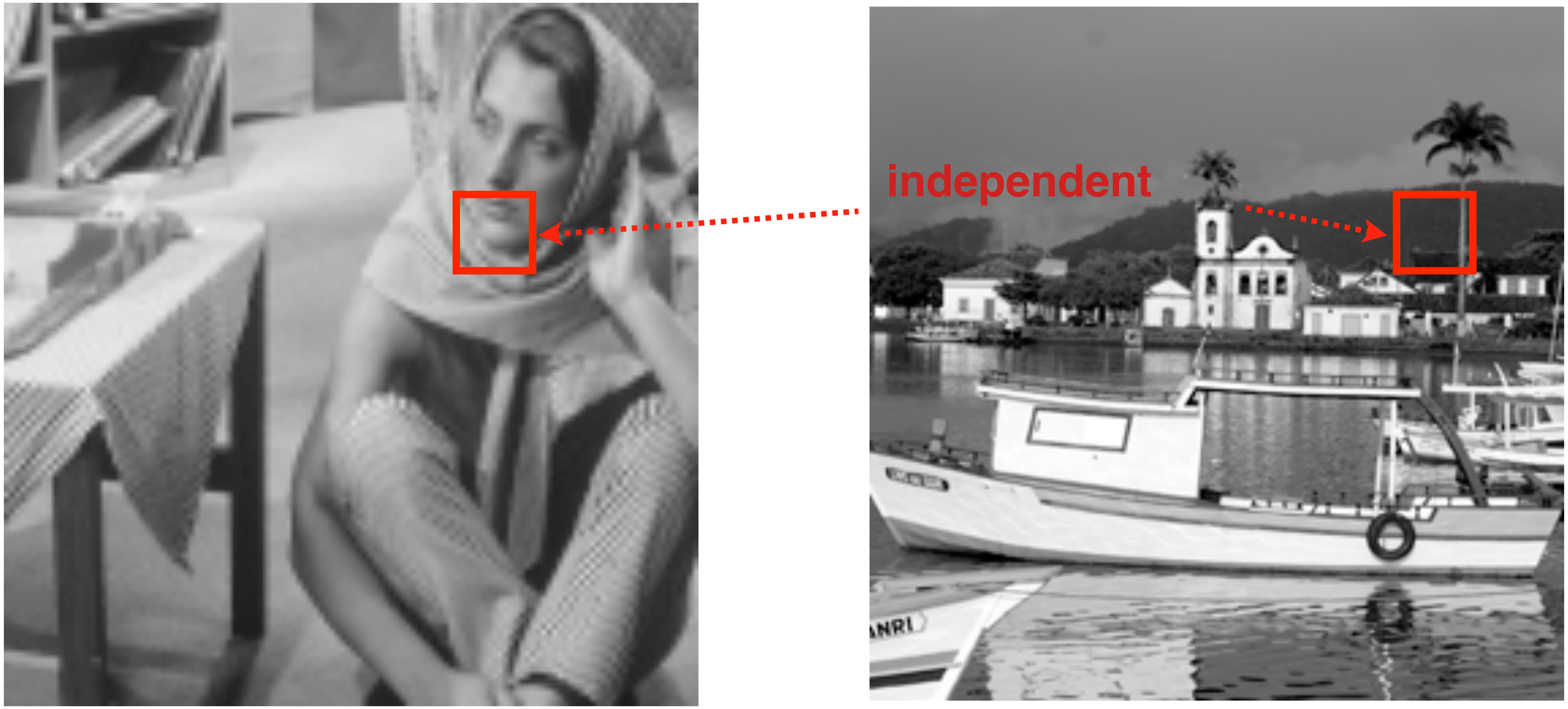
Sparse representations and BSS

The sparse way of tackling BSS

J.Bobin/C. Kervazo

jerome.bobin@cea.fr - christophe.kervazo@telecom-paris.fr

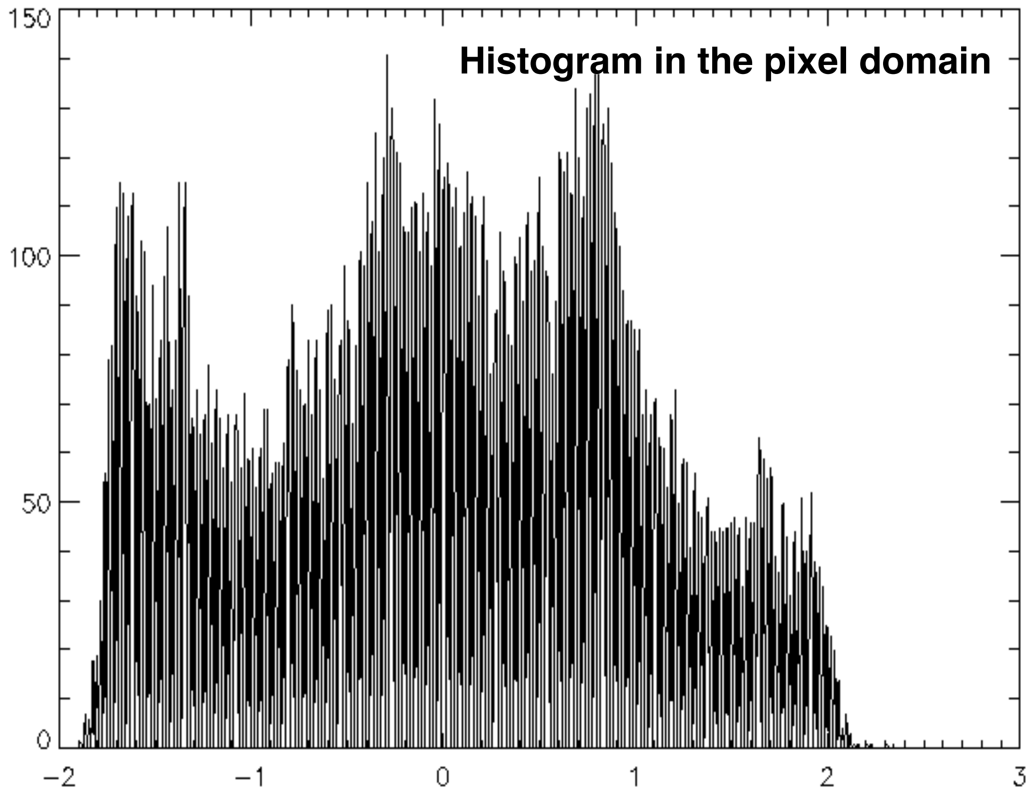
BSS reloaded - the sparse way



How to account for signal's inner structures/morphologies ?

How to model signal's structures ?

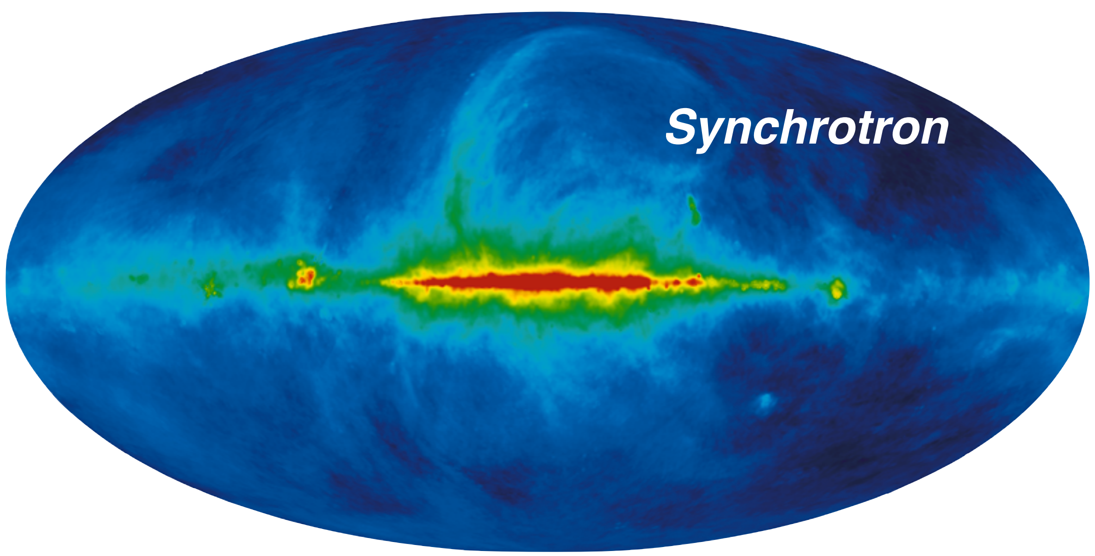
Rationale: source separation highly depends on the ability to increase the contrast between the sources

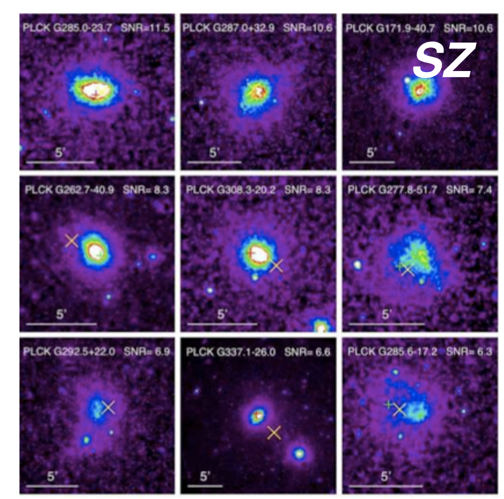


- The i.i.d. assumption is too simplistic
- The probability density function of natural images is highly complex

Building a "universal"/flexible BSS framework

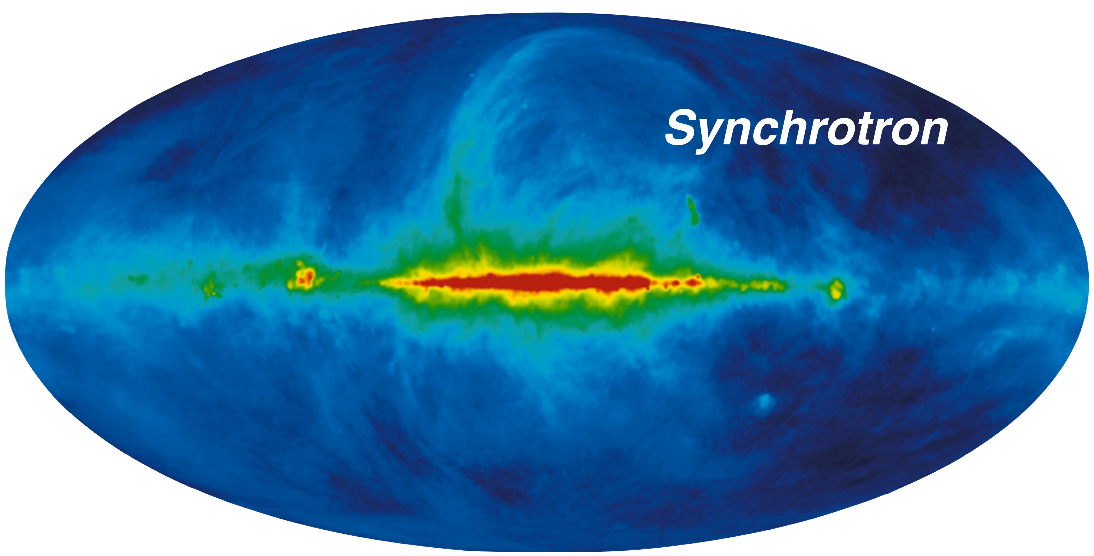
Morphological diversity

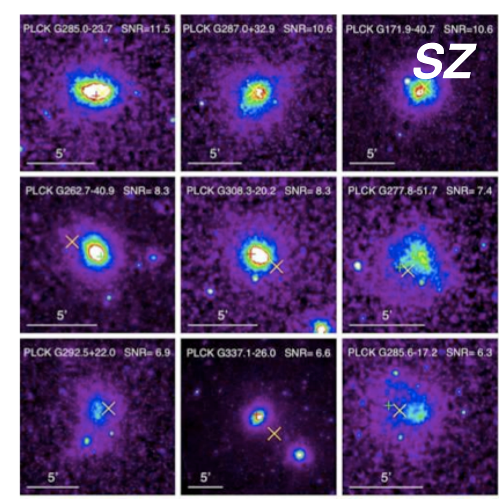




How to better characterize such signals ?

Morphological diversity

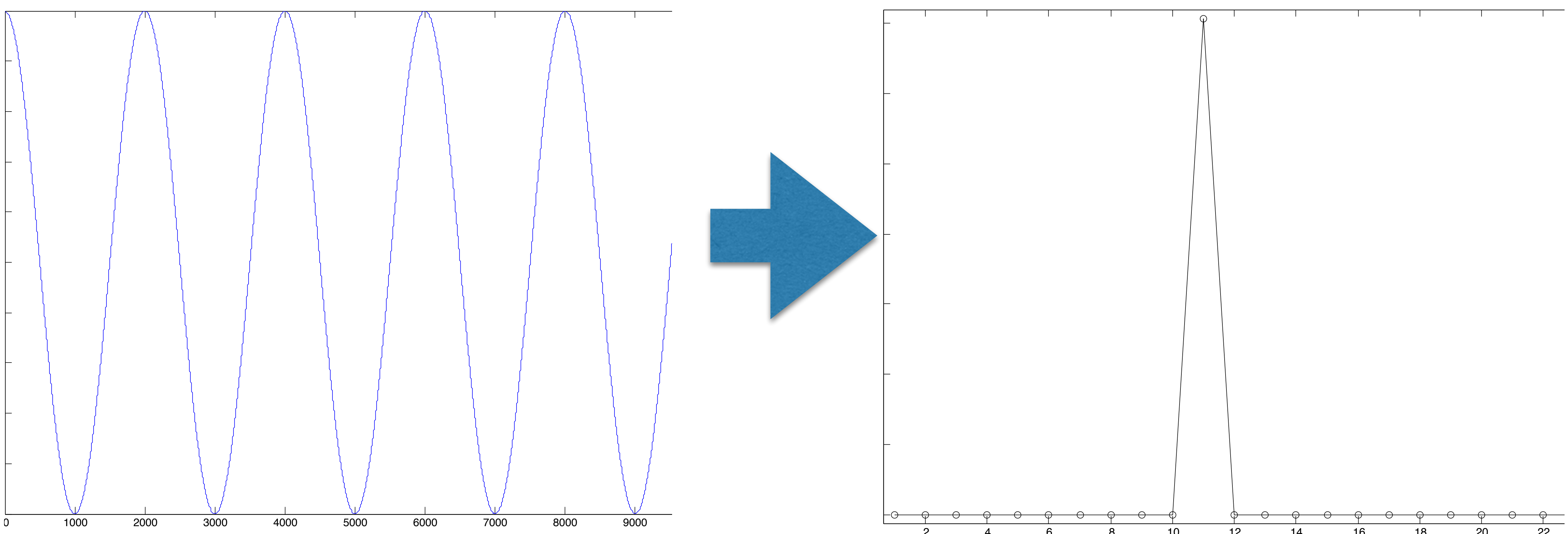




Sources can be distinguished based on their morphologies

In the last two decades, the most dramatic advances in signal estimation have focused on using prior information enforcing signal properties based on desired geometrical/morphological properties.

Gist of the sparsity : signals can be sparsely represented in representations (basis, etc.) that efficiently encode their geometrical/morphological properties.



x

$$\alpha = \mathbf{\Phi} x$$

Discrete cosine transform

Sparse signal representations

Basis, frame, dictionary
(Fourier, wavelets, curvelets ....)

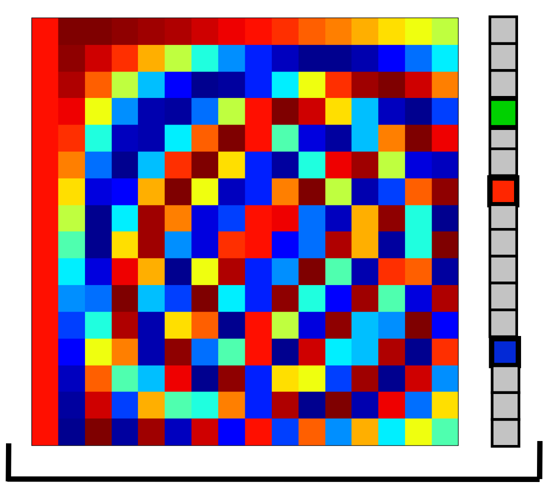
$$\Phi = \{\phi_1, \cdots, \phi_t\}$$

coefficients

$$x = \sum_{j=1}^{t} \alpha_j \phi_j$$

Φ α

$x =$



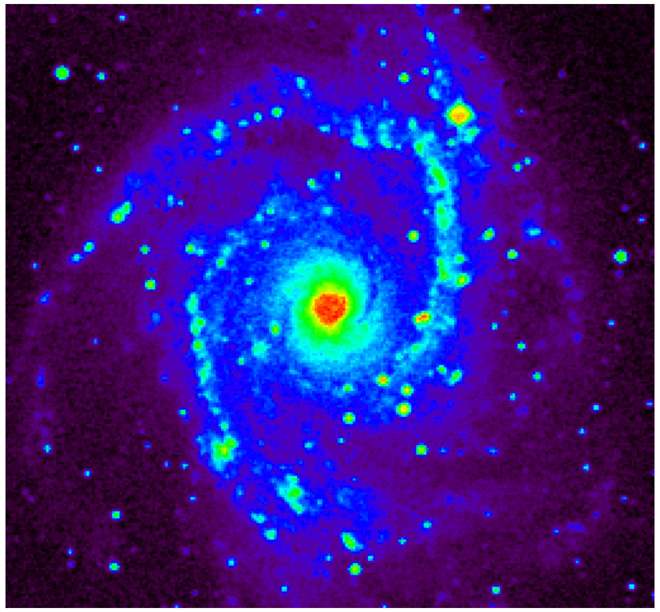
Prior: Data Representation

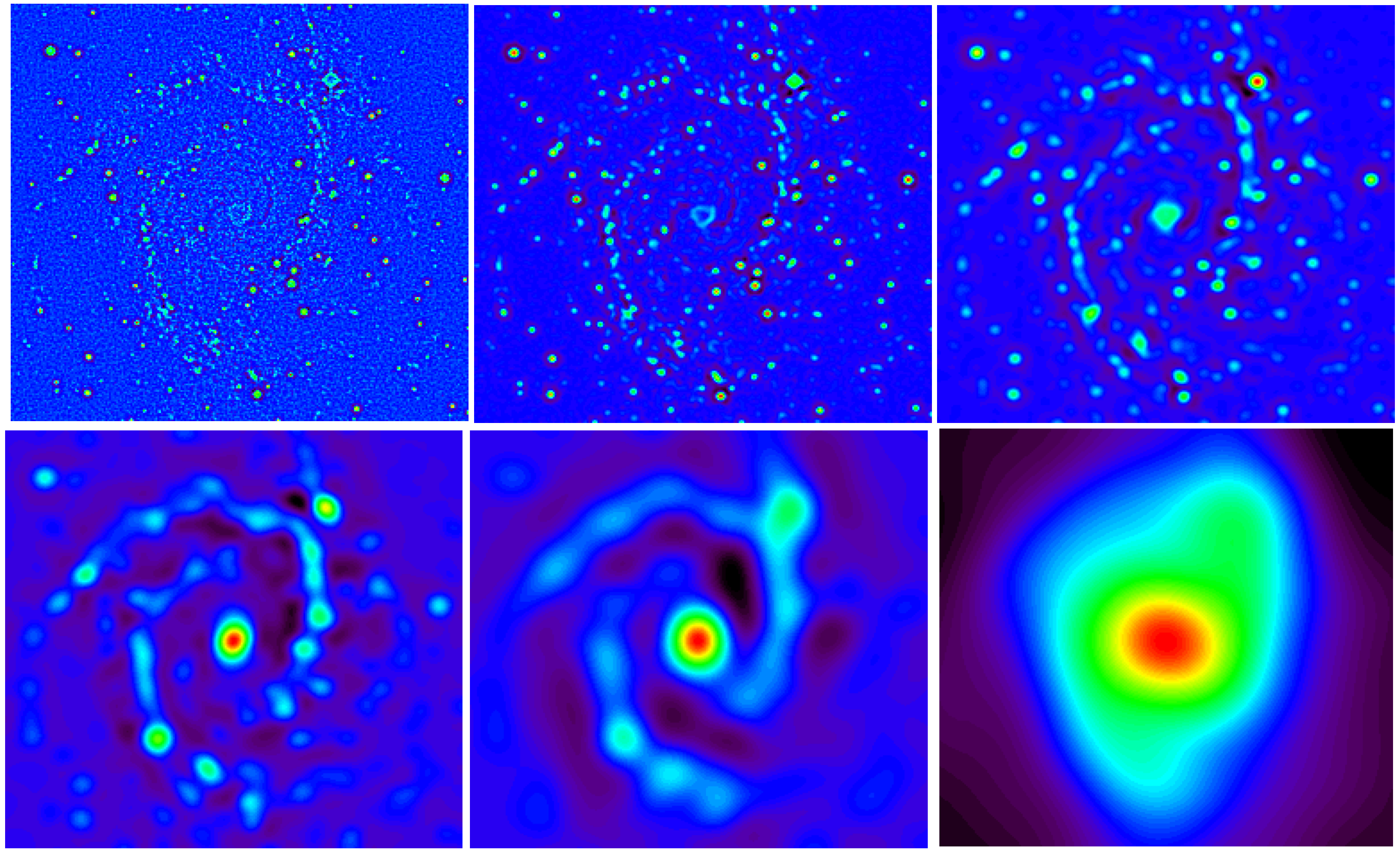
Sparse Model 1 :

x is assumed to have only k non zero entries

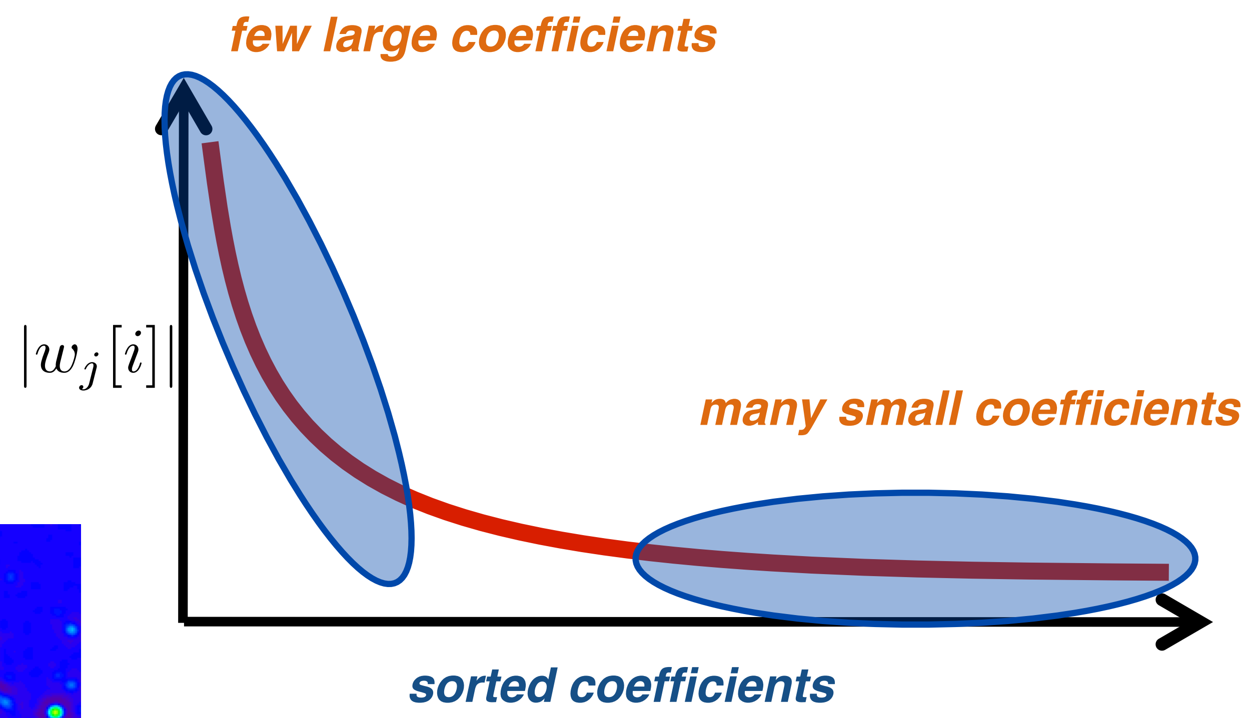
x is said to be exactly k-sparse in Φ

Sparse signal representations





Starlet transform
(isotropic undecimated wavelet transform)



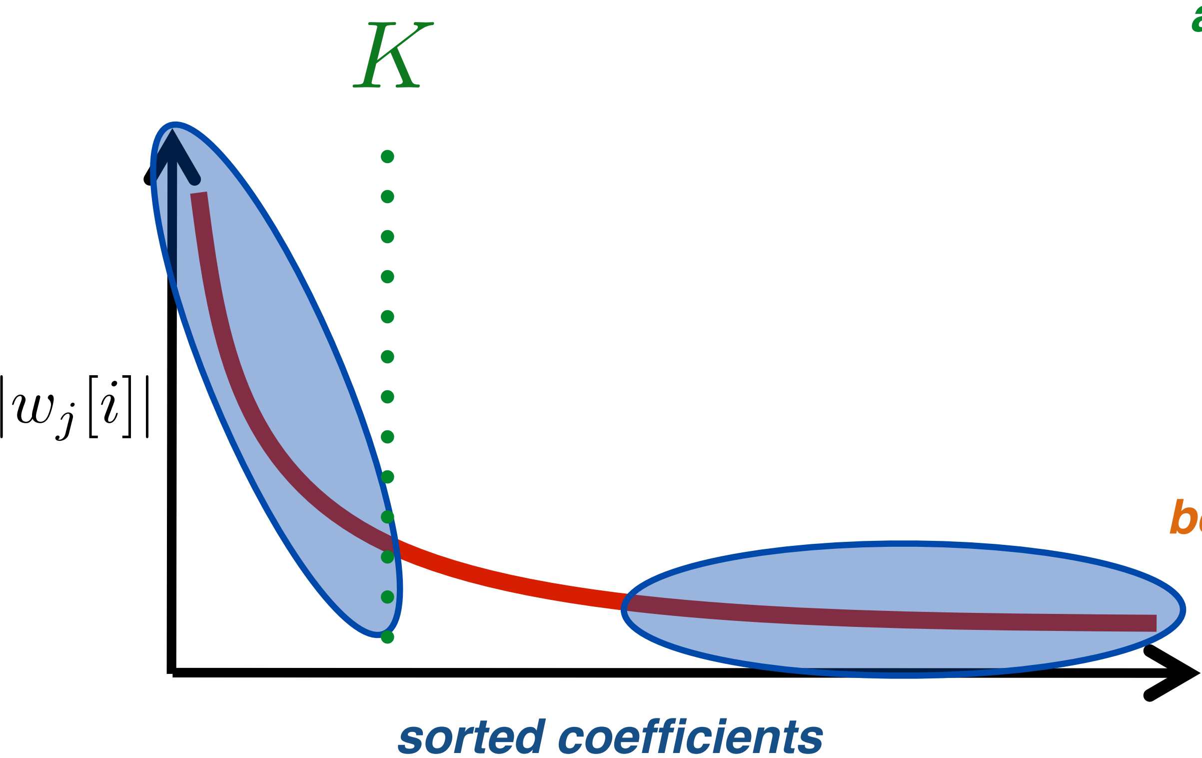
for clarity $w_j[i] \rightarrow \alpha[k]$

$$|\alpha[k]| \leq C k^{-1/q}$$

Sparse Model 2 :

x is approximately sparse in Φ

The sparse distribution of wavelet coefficients is at the origin of a fundamental data characterization tool: ***non-linear approximation***



$$\overbrace{\|\alpha - \alpha_K\|_2}^{\text{approximation error}} \leq C_q K^{1/2-1/q} \text{ for } q < 2$$

best K-term approximation

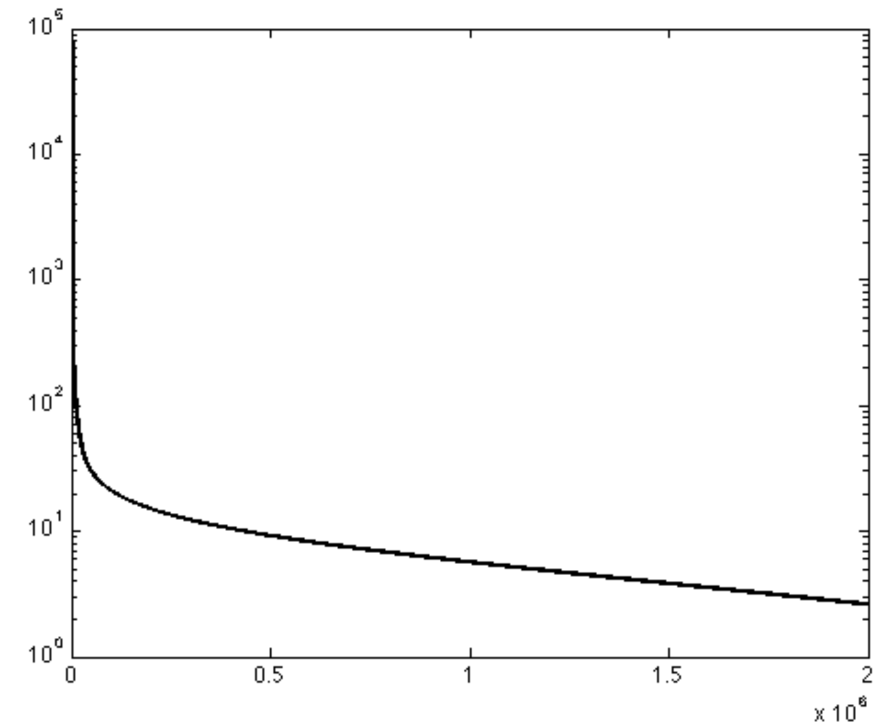
The smaller q, the better the approximation

for clarity $w_j[i] \rightarrow \alpha[k]$

$$|\alpha[k]| \leq C k^{-1/q}$$

This property is at the origin of the main applications of wavelets

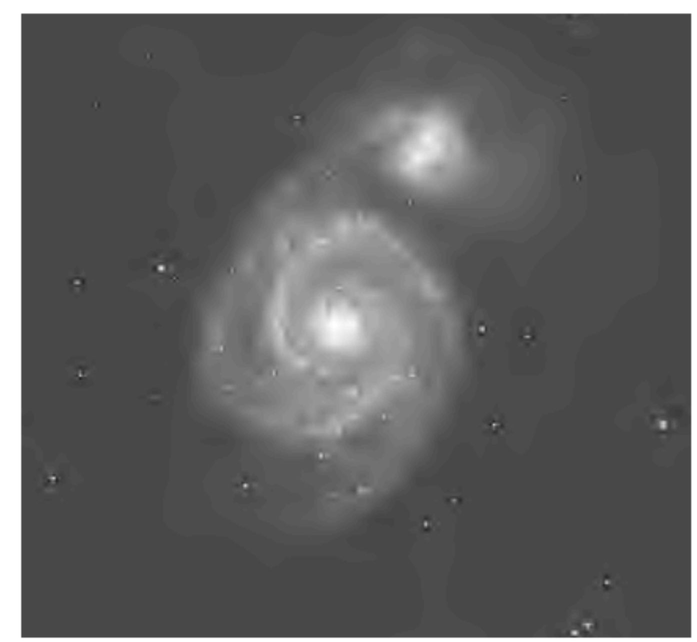
wavelet transform



sorted wavelet coefficients

1%

0.1%

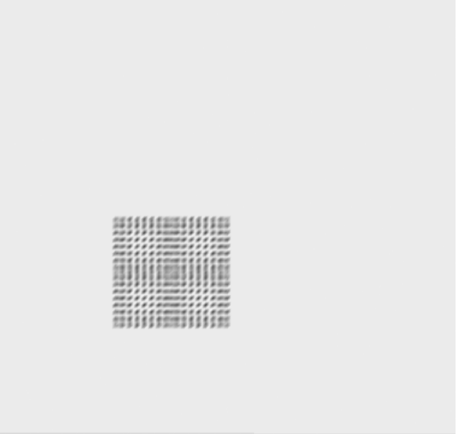


0.01%

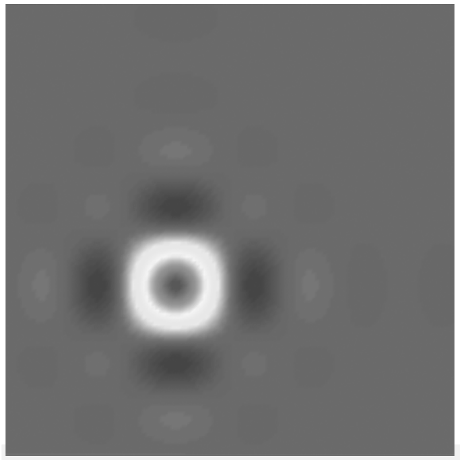
Sparse signal representations

In general, sparse representations should be chosen based on the desired morphology

Cosine transform — Oscillating/periodic structures; Textures



Wavelets — Singularities, point-like structures



Curvelets — Contours

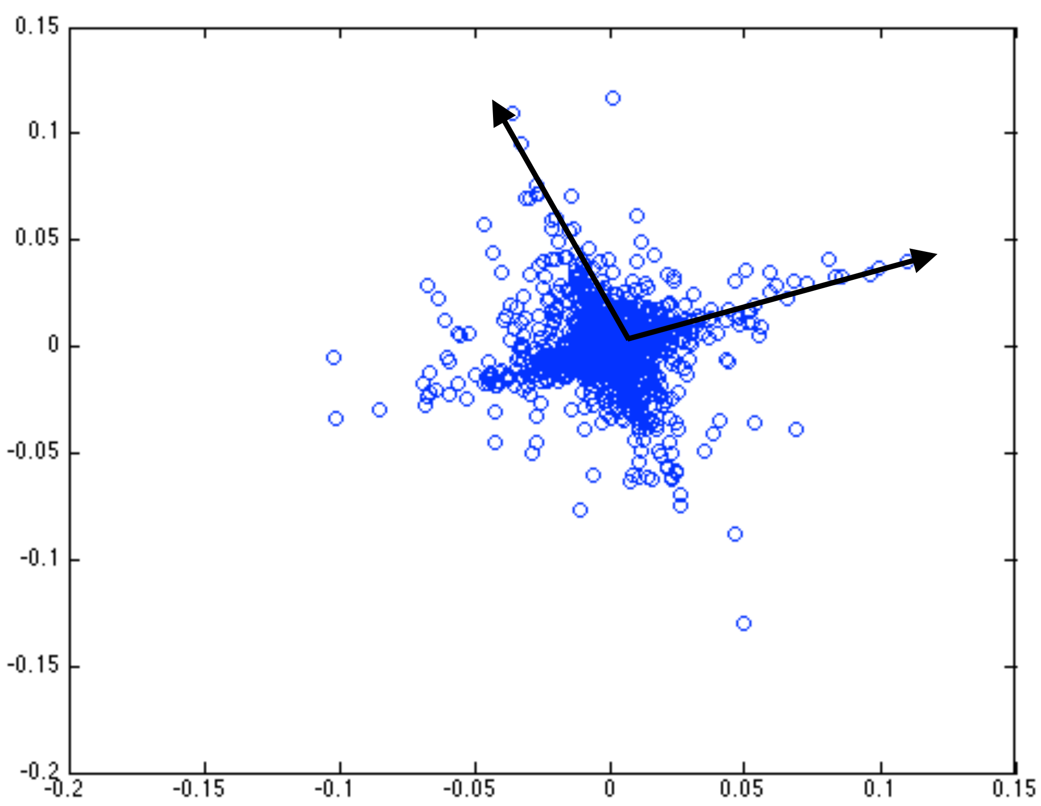
Or others (see Course #6)

Sparse BSS

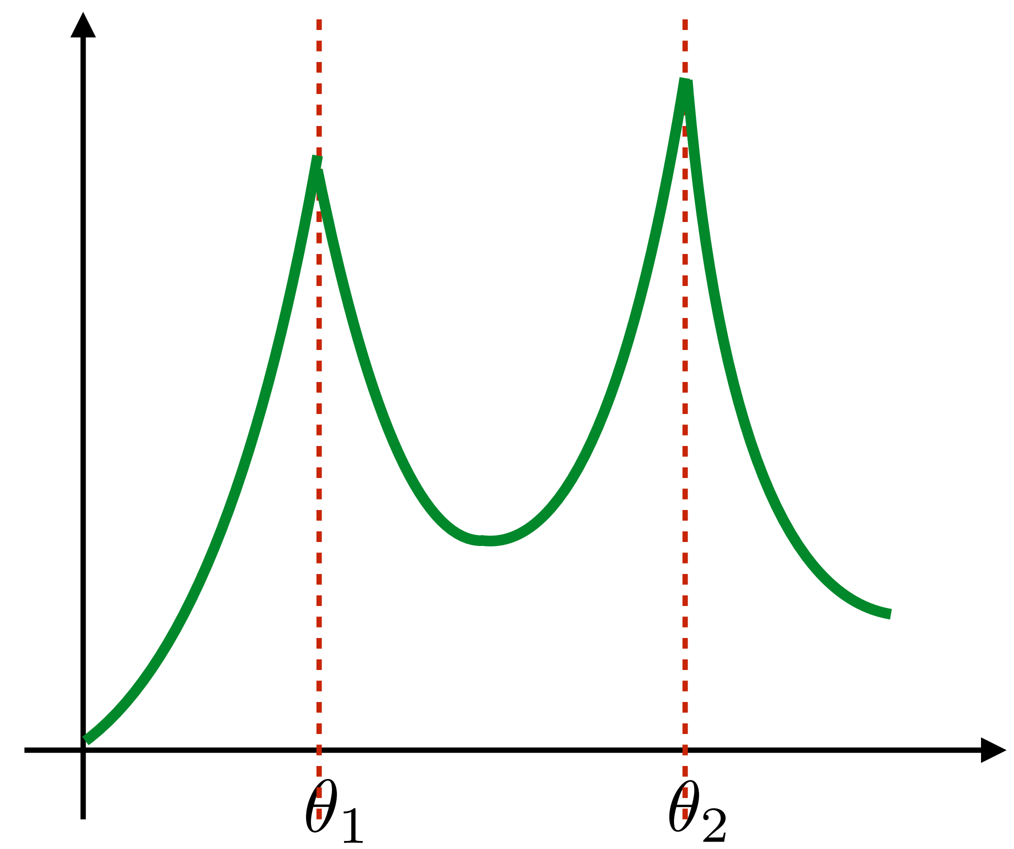
Sparse component analysis

Original SCA methods are based on a two-step estimation of the mixing matrix and the sources.

Estimating the mixing matrix is generally made via clustering techniques:



Scatter plot of 2 mixtures



Angle

which is performed by looking for clusters in an angular representation of the data point cloud.

Sparse component analysis

The sources are then estimated by minimizing their L_p norm:

$$\{\hat{\mathbf{S}}\} = \mathrm{Argmin}_{\mathbf{S}} \sum_i \lambda_i \|\alpha_i\|_{\ell_1} + \frac{1}{2}\|\mathbf{X} - \mathbf{A}\mathbf{S}\|_F^2 \text{ with } s_i = \alpha\mathbf{\Phi}$$

One big advantage of the SCA is its ability to potentially perform well in the under-determined case (the number fo sources is larger than the number of observations).

As well, the number of sources can be jointly estimated via clustering.

Clustering is highly sensitive to noise contamination.

It is limited to low-dimensional data (*e.g.* audio data processing, etc.).

The first method that enforced the sparsity in an orthogonal transform has been introduced by Zibulevsky et al.ien 2001 (RNA).

The objective of RNA is to estimate the unmixing matrix B (inverse of the mixing matrix) so that :

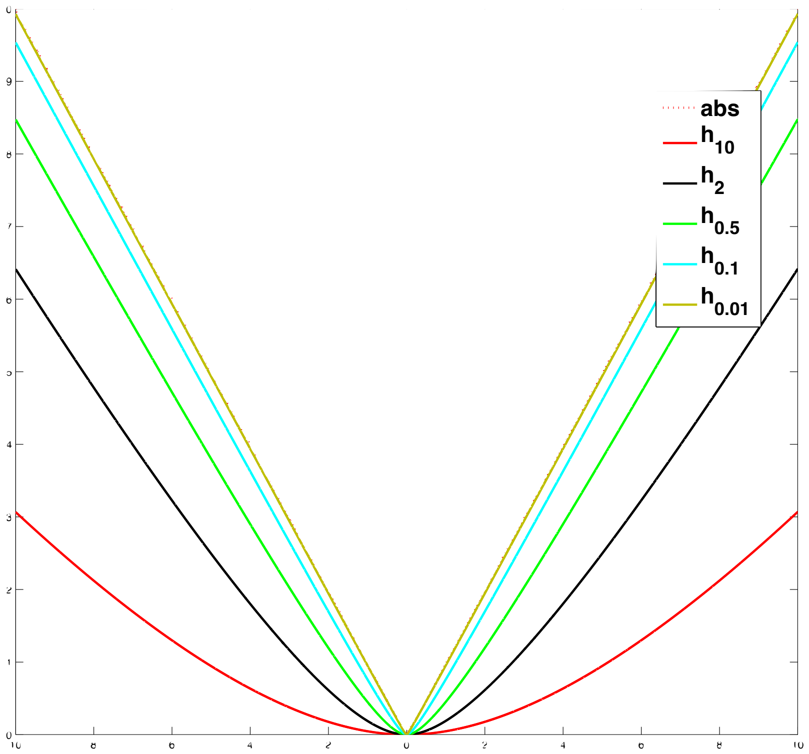
$$\min_{\mathbf{B}} \quad \gamma \sum_{i,j} h(S_{ij}) - \log(\det(\mathbf{B})) \text{ où } \mathbf{S} = \mathbf{BX}$$

This is equivalent to minimizing the anti log-likelihood of B assuming that the samples of the sources in the sparse domain are i.i.d. and :

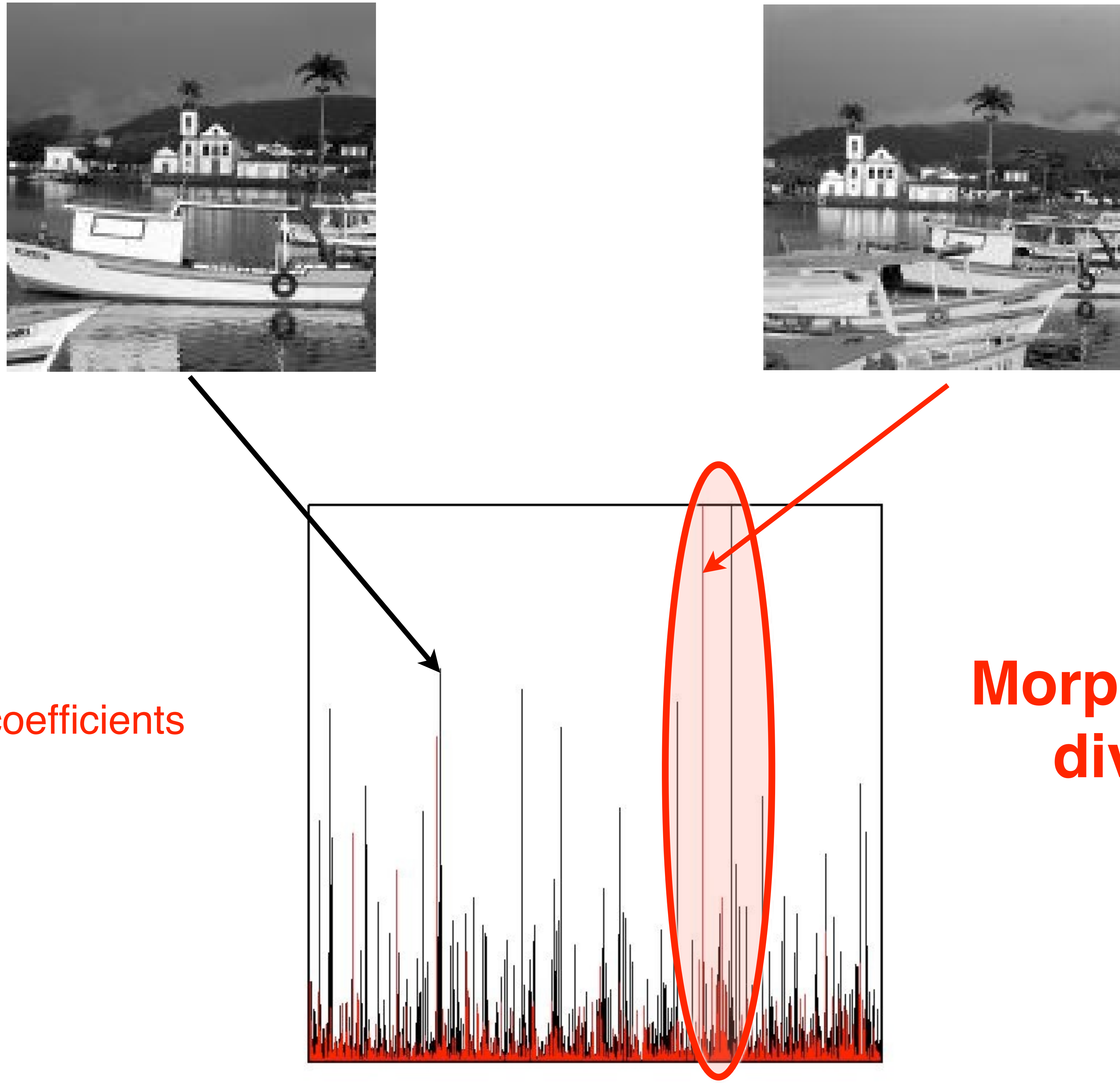
$$f_{S_{ij}} \propto \exp \gamma h_\lambda(S_{ij})$$

$$h_\lambda(z) = |z| - \lambda \log\left(1 + |z|/\lambda\right)$$

differentiable approximation of abs(z)



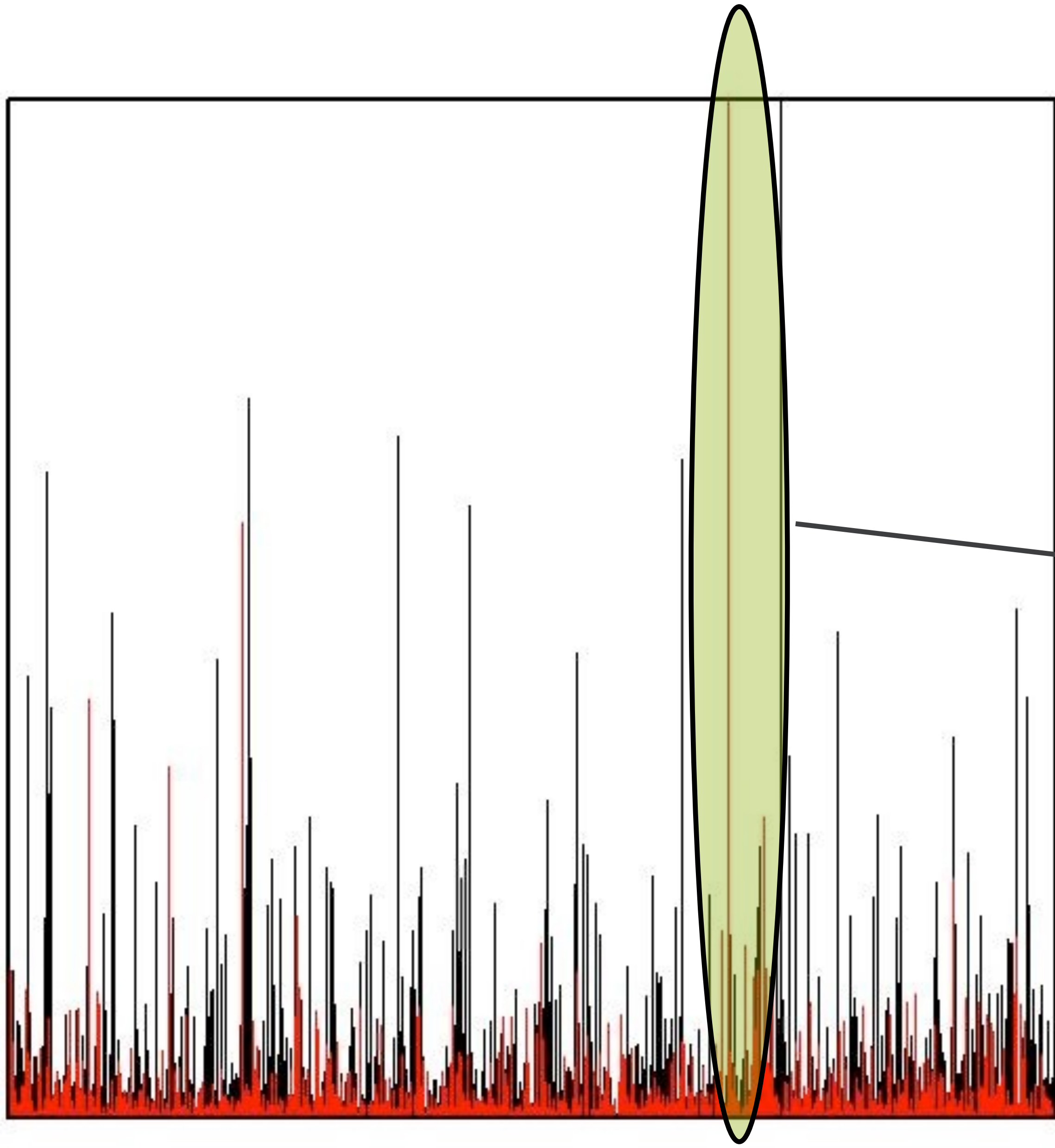
Sparse component analysis

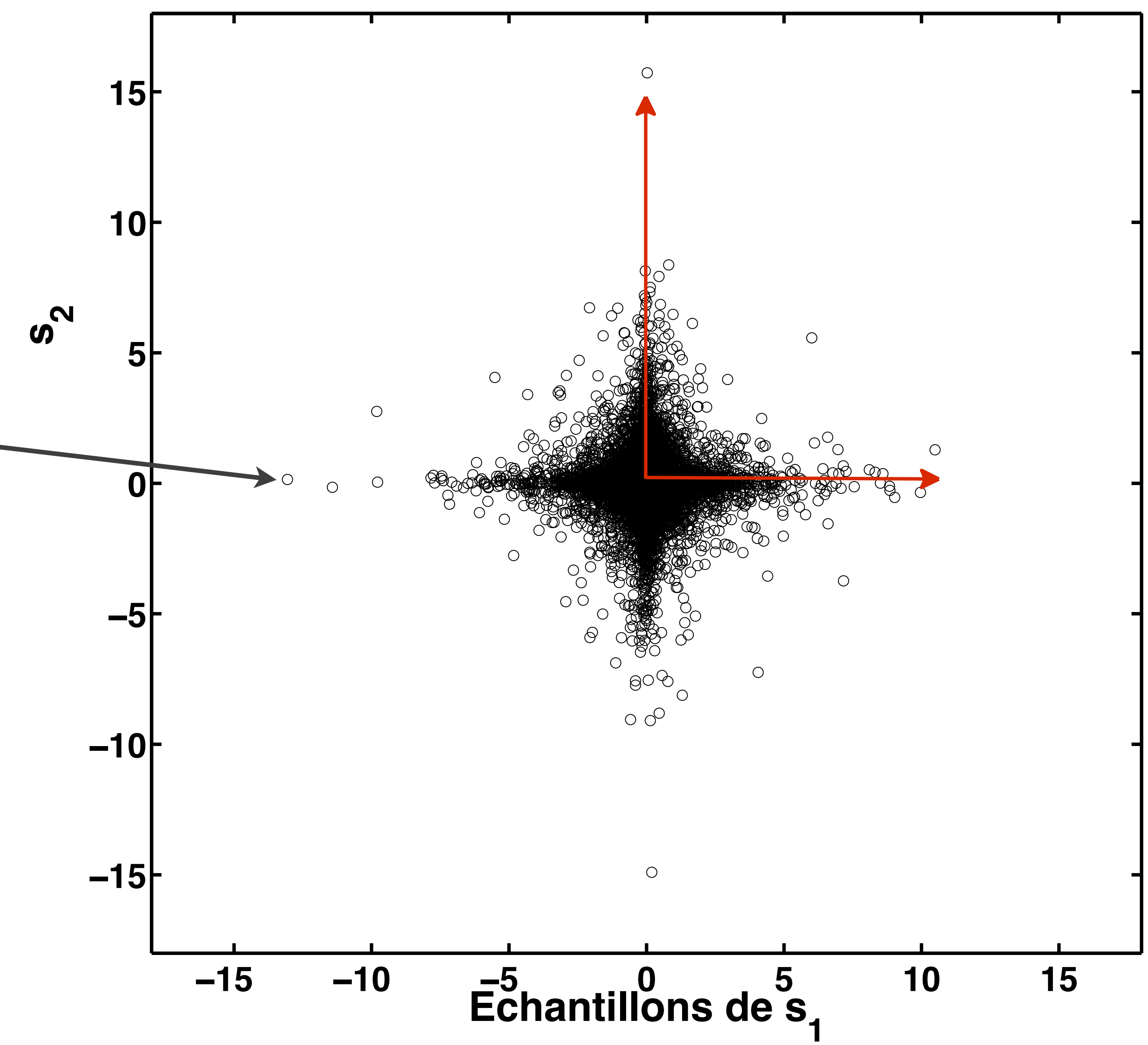


Wavelet coefficients

Morphological diversity

Sparse component analysis

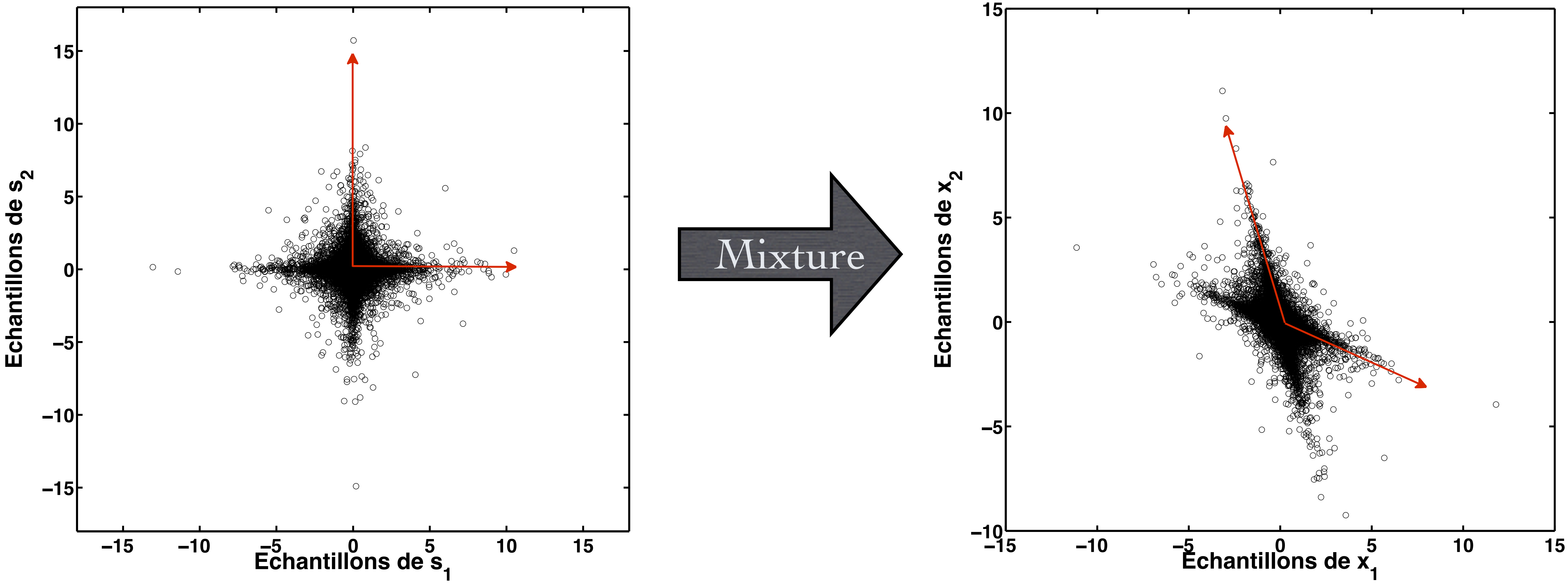




Different sources have different sparsity patterns on the transformed domain

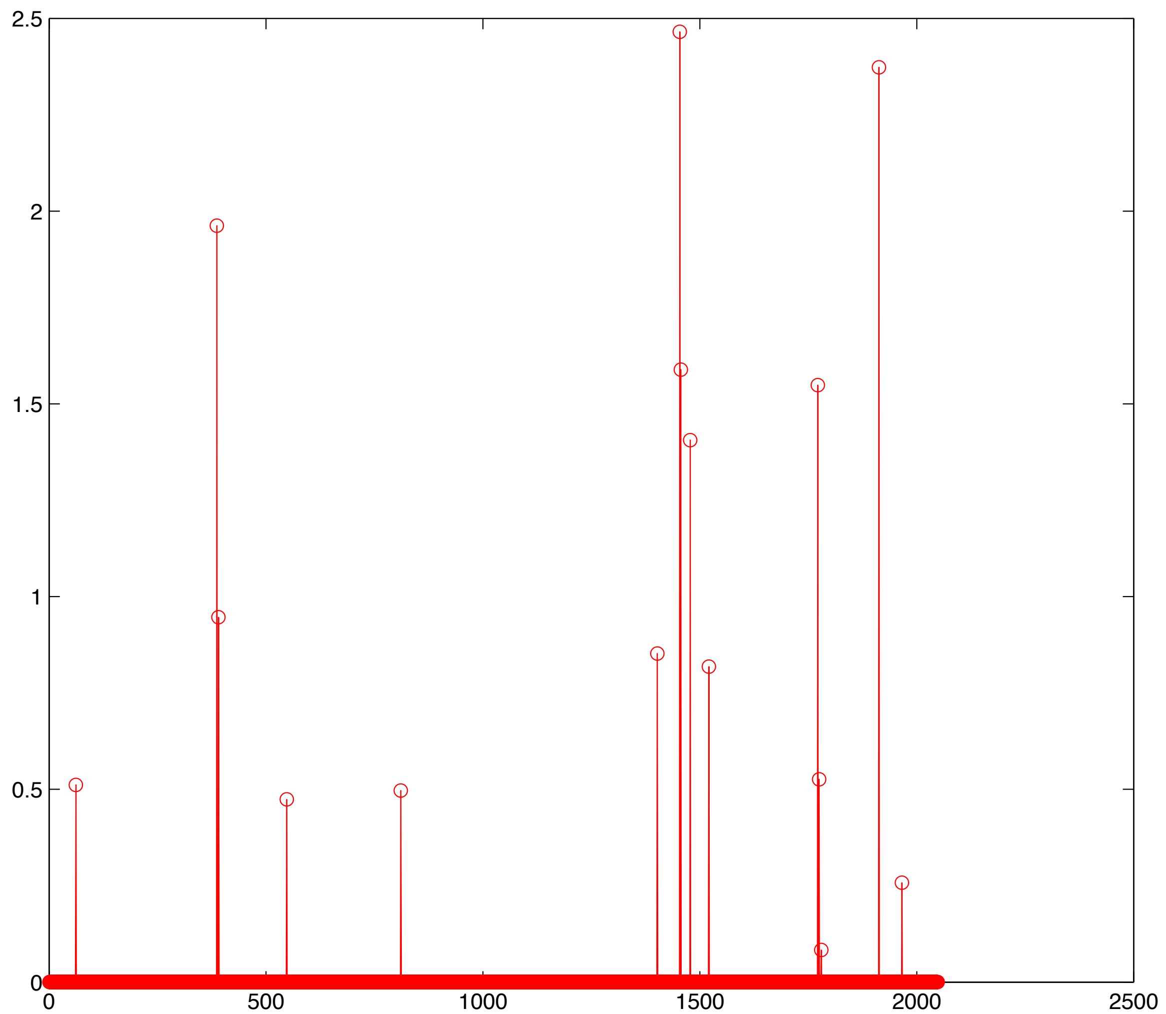
significant coefficients for one source ⟷ slightly significant for the others

Sparse component analysis

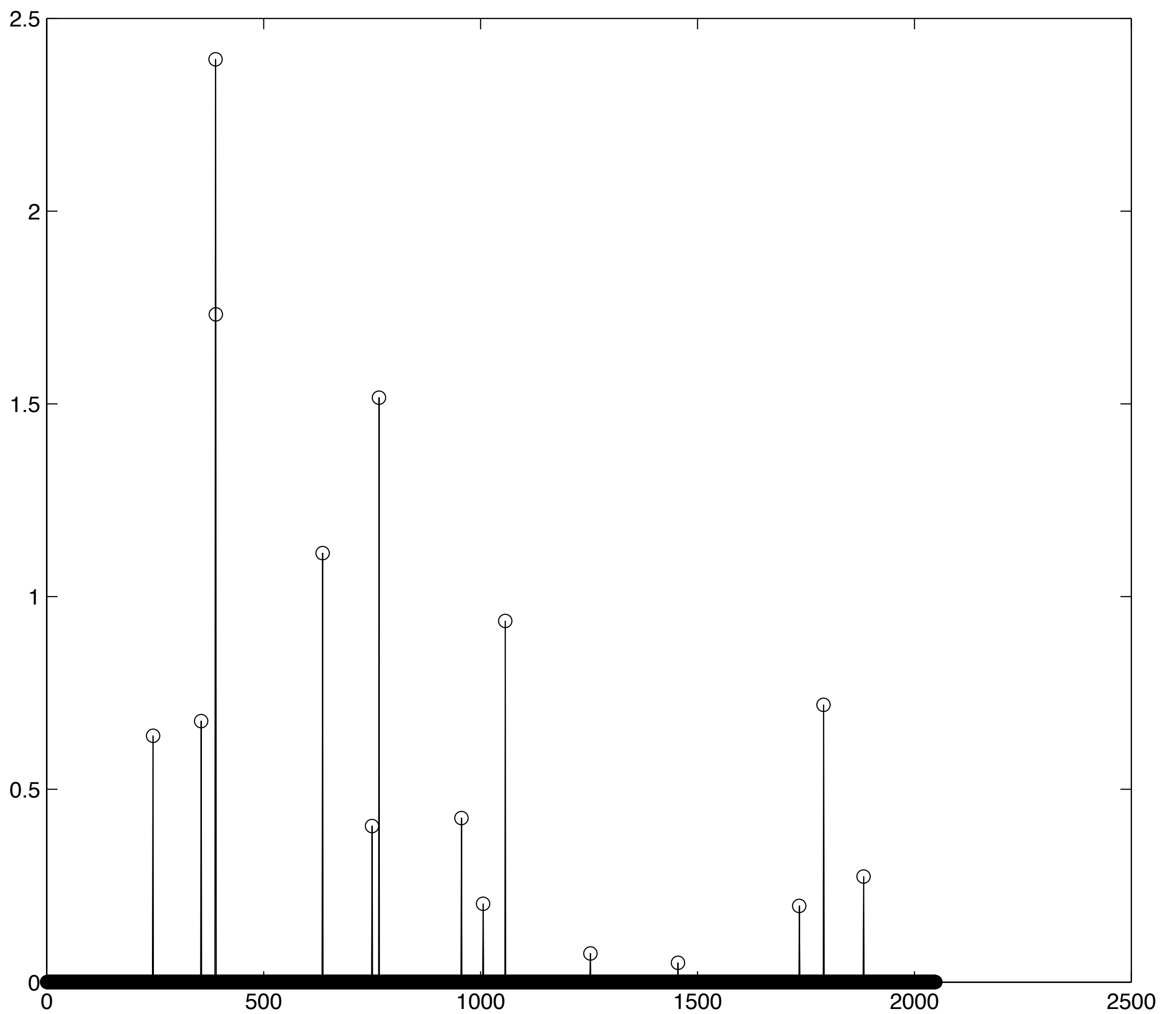


Gist: mixture of the sources are less sparse than the sources

Exactly sparse case:



s1

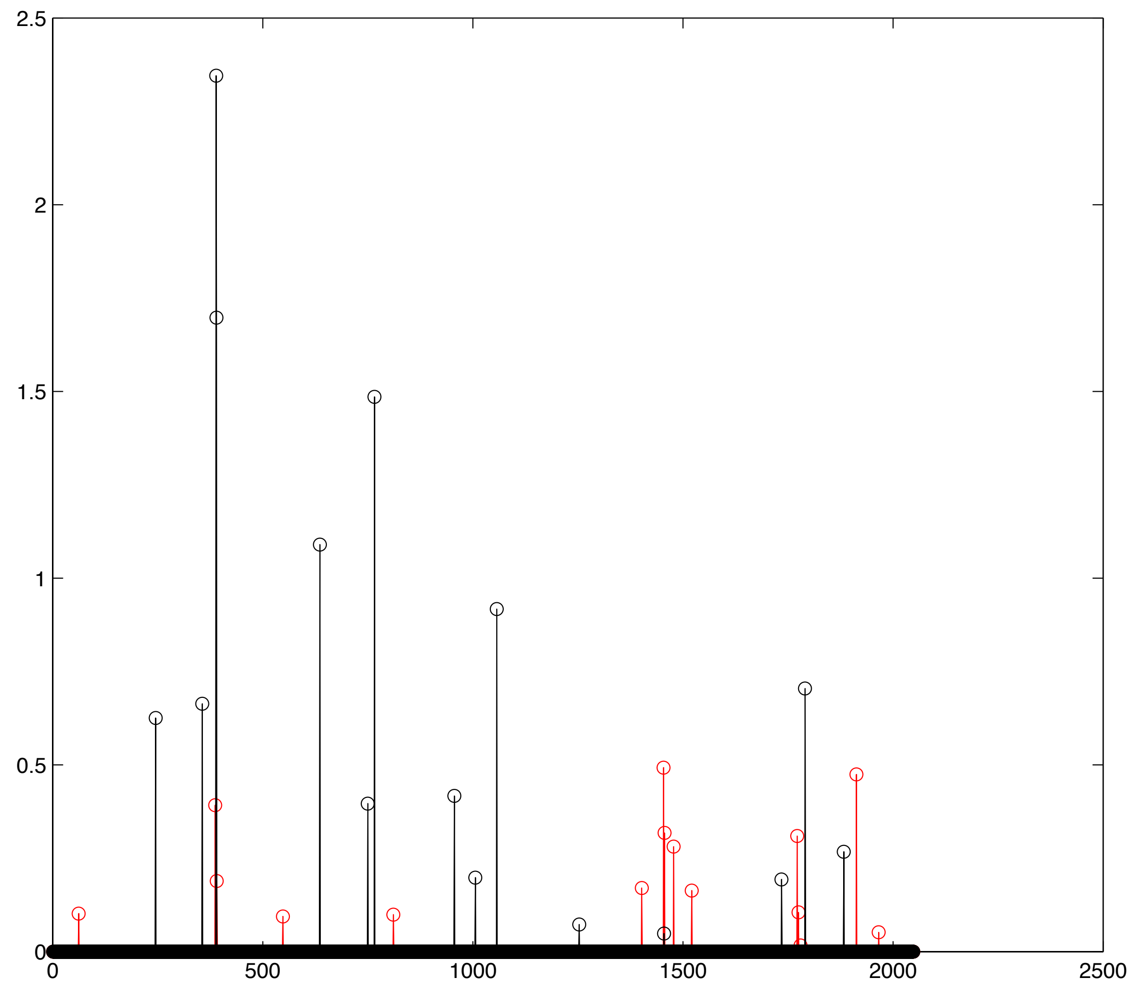


s2

K_1 **non-zero coefficients**

K_2 **non-zero coefficients**

Exactly sparse case:

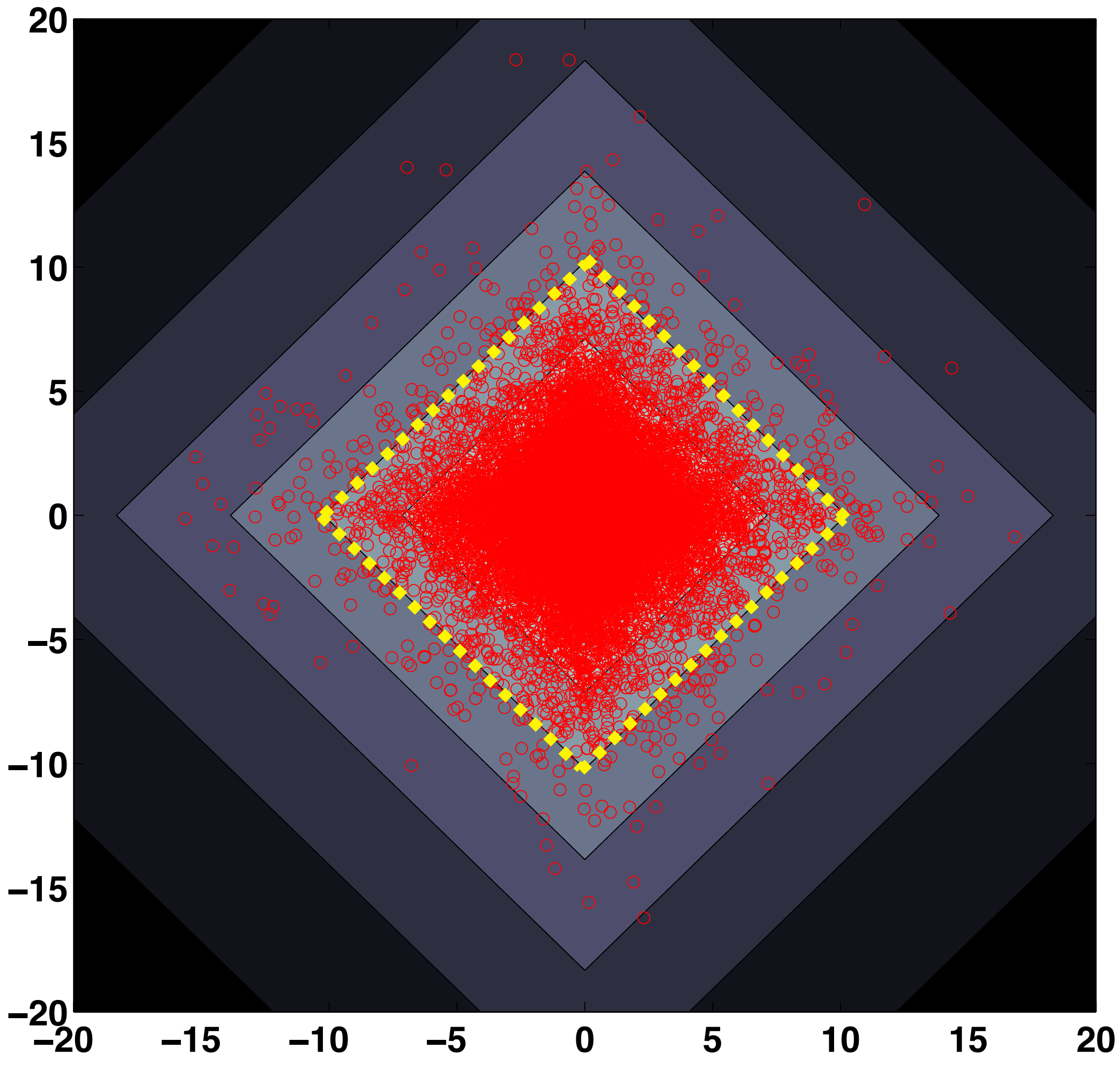


mixtures

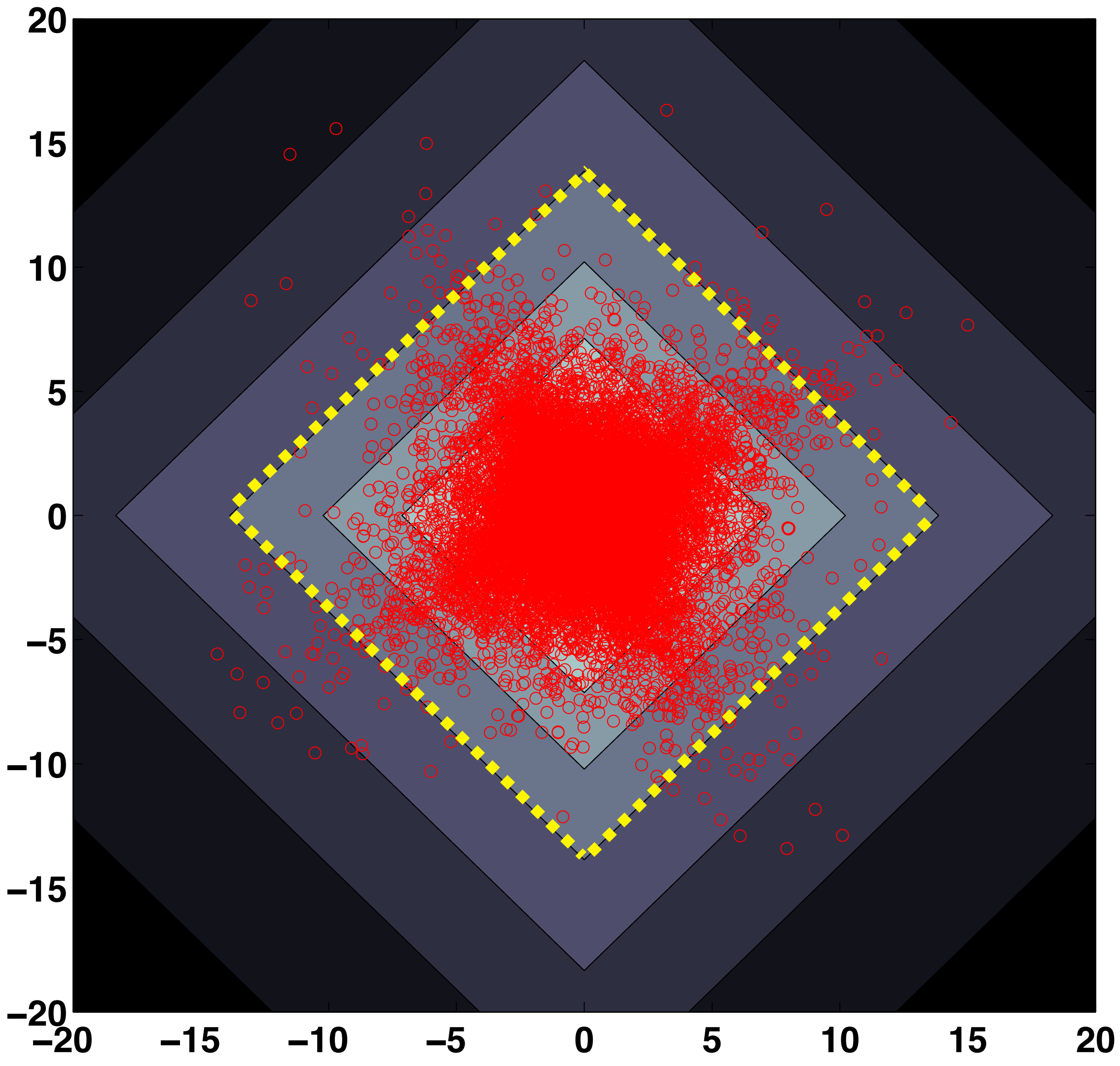
$$K > \max\{K_1, K_2\}$$

non-zero coefficients

Approximately sparse case:



Sources



Mixtures

$$\|\mathbf{AS}\|_1 > \|\mathbf{S}\|_1$$

Sparse BSS

Limitations of RNA:
- The sparse modeling of the sources is limited to orthogonal representations.
- No modeling of the noise.

In 2007, a more general method has been introduced: GMCA (Generalized Morphological Component Analysis).

It aims at solving:

$$\min_{\mathbf{A},\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_1 \text{ s.c. } \|\mathbf{X} - \mathbf{AS}\|_F < \epsilon \text{ avec } \mathbf{S} = \boldsymbol{\alpha}\boldsymbol{\Phi}$$

Coefficients of the sources

Data fidelity term

The sources are sparse in a representation that is not necessarily orthogonal

Sparse component analysis

Model :

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{N} = \mathbf{A}\boldsymbol{\alpha}\boldsymbol{\Phi} + \mathbf{N} \text{ avec } \boxed{\|\mathbf{N}\|_F < \epsilon}$$

Estimating the sources and the mixing matrix is done as follows:

$$\min_{\mathbf{A},\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_1 \text{ s.c. } \|\mathbf{X} - \mathbf{A}\mathbf{S}\|_F < \epsilon \text{ avec } \mathbf{S} = \boldsymbol{\alpha}\boldsymbol{\Phi}$$

In Lagragian formulation, this boils down to :

$$\min_{\mathbf{A},\boldsymbol{\alpha}} \gamma \|\boldsymbol{\alpha}\|_1 + \frac{1}{2}\|\mathbf{X} - \mathbf{A}\mathbf{S}\|_F^2 \text{ avec } \mathbf{S} = \boldsymbol{\alpha}\boldsymbol{\Phi}$$

GMCA algorithm

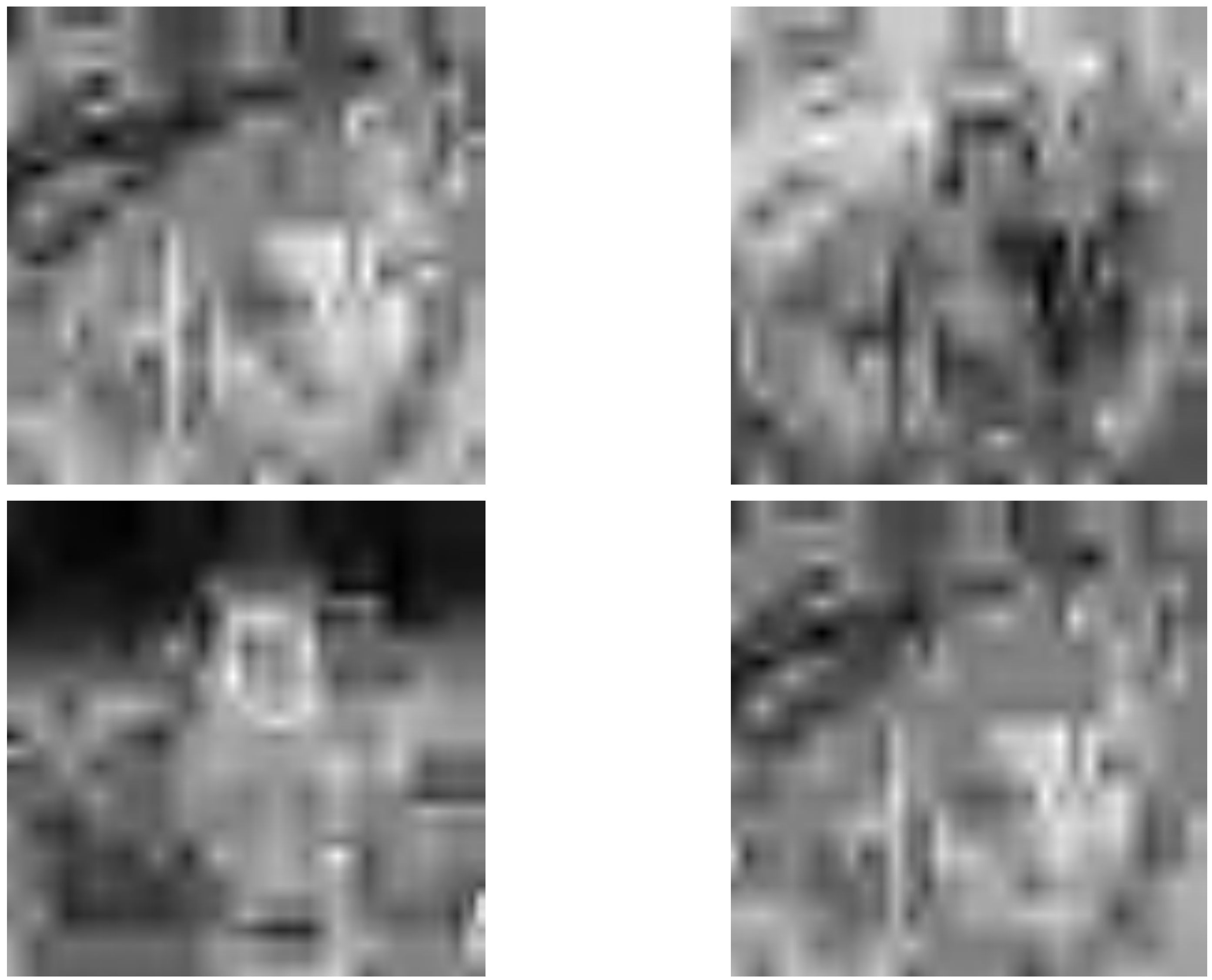
1. Decomposition of $\mathbf{X}$ in $\mathbf{\Phi}$: $\boldsymbol{\alpha}_{\mathbf{X}}$
2. Fix the number of iterations $P_{\max}$ and the threshold $\gamma^{(0)}$
3. While $\gamma^{(h)}$ is greater than $\gamma_{\min}$ (*e.g.* depends on the noise level),
   – Estimation of the sources $\boldsymbol{\alpha}_{\mathbf{S}}$ at iteration h assuming $\mathbf{A}$ fixed : $\boldsymbol{\alpha}_{\mathbf{S}}{}^{(h+1)} = \mathcal{S}_{\gamma^{(h)}}\left(\mathbf{A}^{\dagger(h)}\boldsymbol{\alpha}_{\mathbf{X}}\right)$:
   – Updating $\mathbf{A}$ assuming $\boldsymbol{\alpha}_{\mathbf{S}}$ is fixed :
   $$\tilde{\mathbf{A}}^{(h+1)} = \boldsymbol{\alpha}_{\mathbf{X}}\tilde{\boldsymbol{\alpha}}_{\mathbf{S}}^{(h+1)^T}\left(\tilde{\boldsymbol{\alpha}}_{\mathbf{S}}^{(h+1)}\tilde{\boldsymbol{\alpha}}_{\mathbf{S}}^{(h+1)^T}\right)^{-1}$$
   – Decreasing the threshold $\gamma^{(h)}$.

Decreasing the threshold allows:

- **to provide a greater robustness to local minima**
- **to provide more robustness to noise**

Sparse component analysis

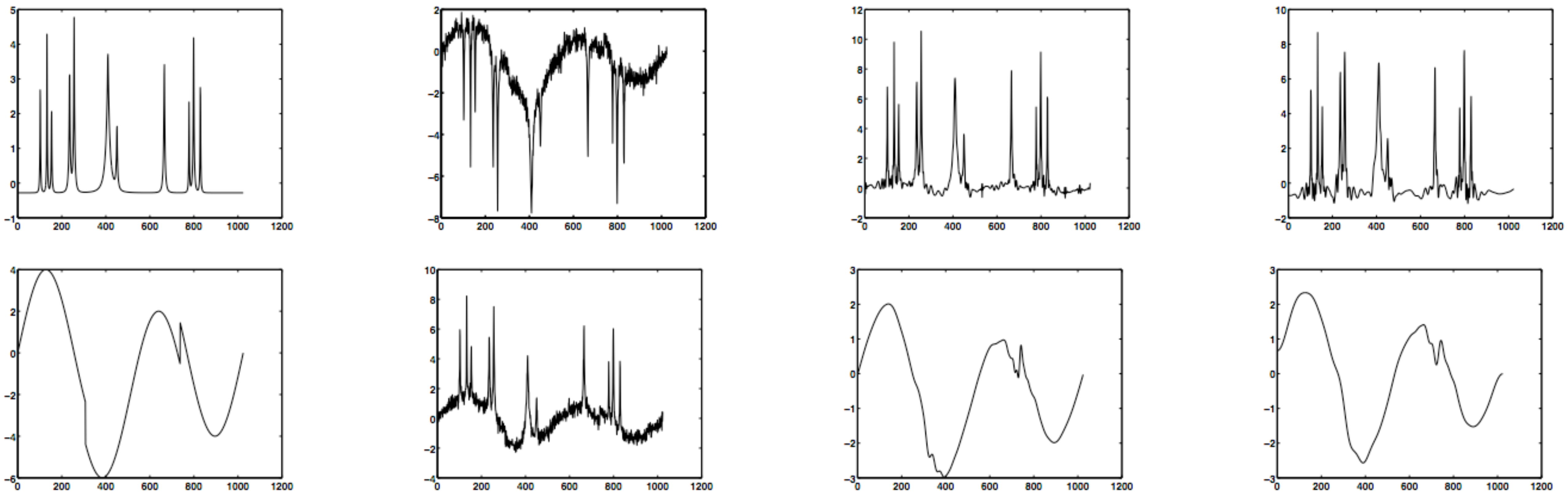
A and **S** are estimated from the most prominent elements of each sources



Evolution of the estimates along the iterative process

4 sources, 4 mixtures without bruit, sparsity in orthogonal wavelets

Sparse component analysis



Sources

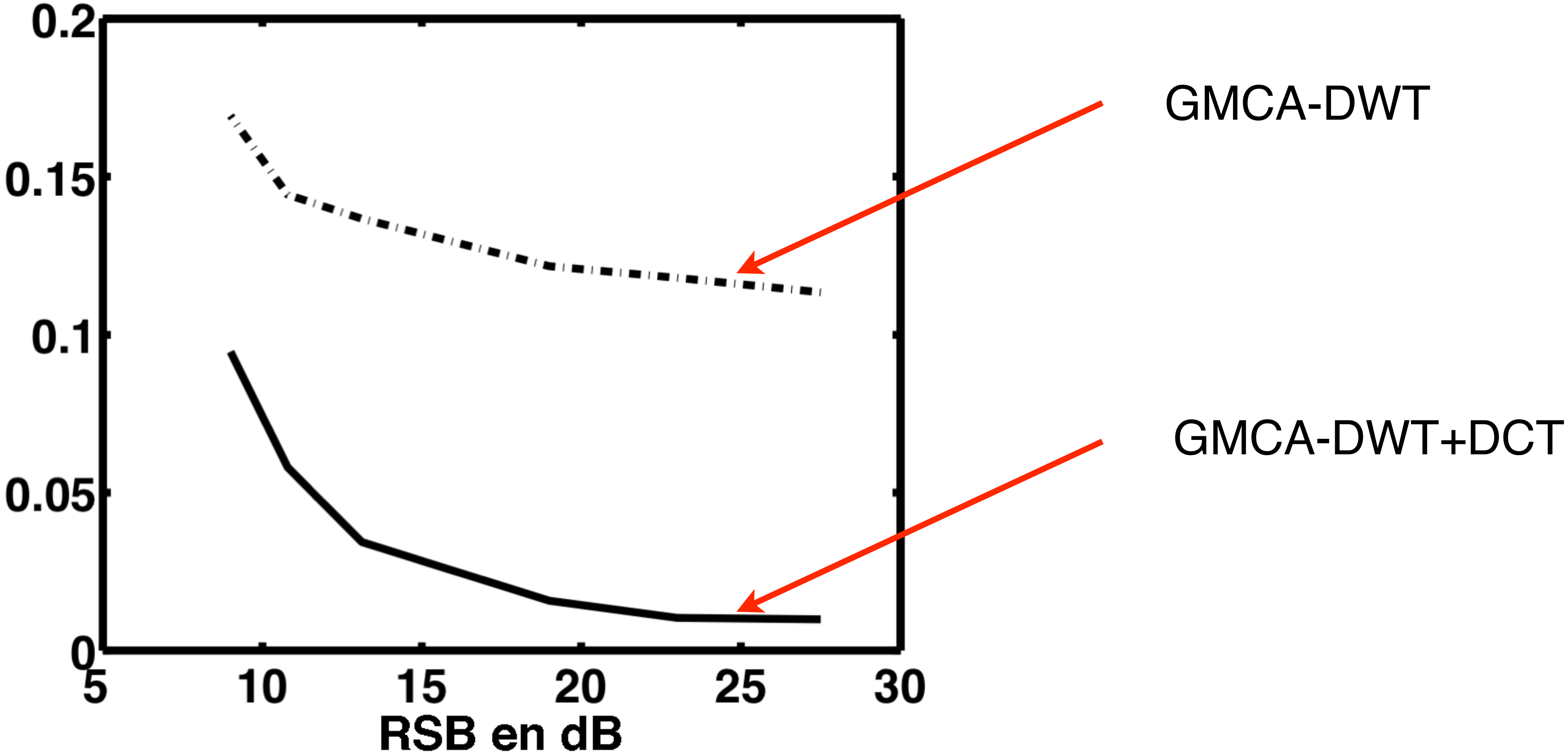
Mixtures
SNR : 19dB

Estimated sources
GMCA - DWT

Estimated sources
GMCA - DWT+DCT

Impact of the representation

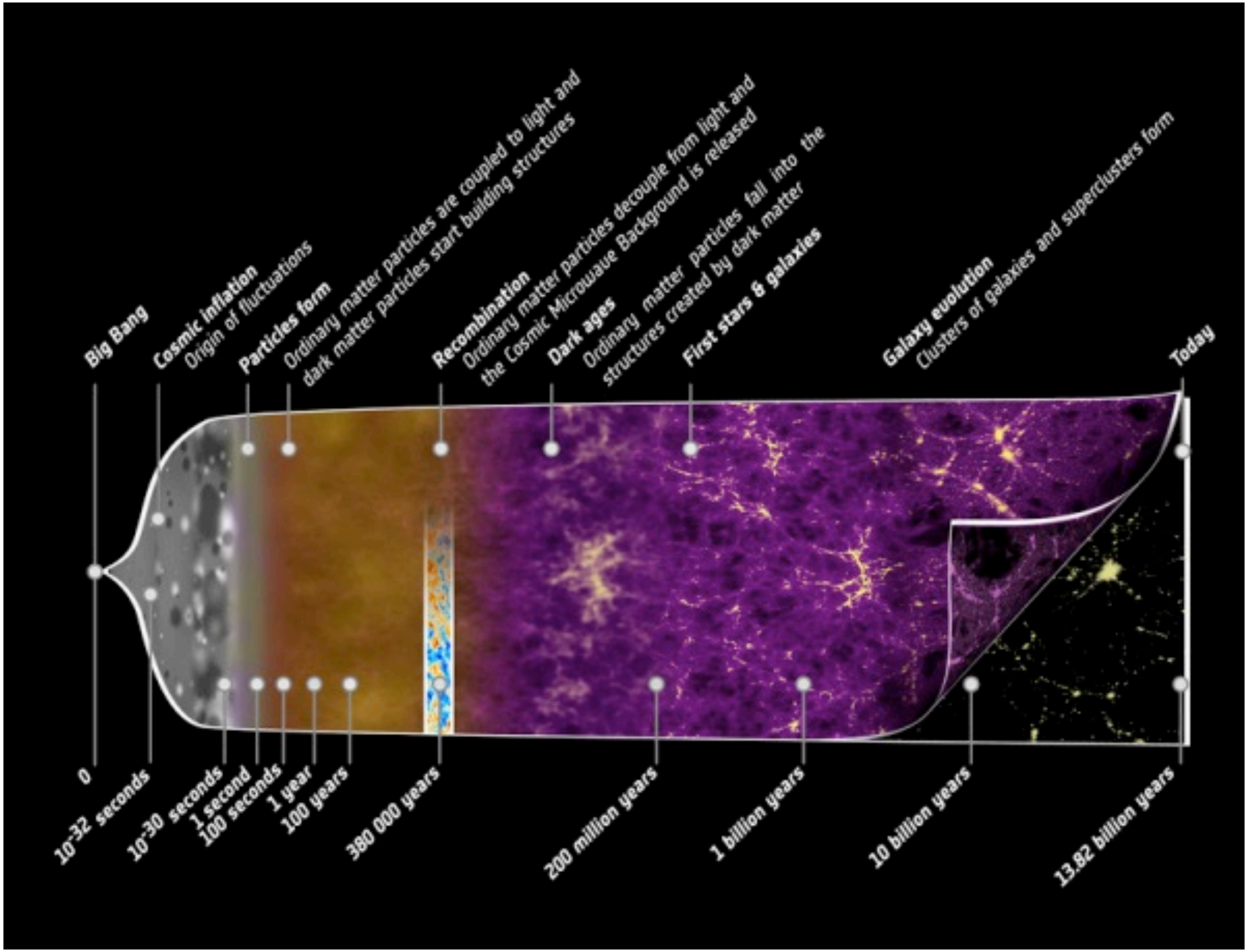
Mixing matrix criterion $\mathcal{C}_{\mathbf{A}} = \left\| \mathbf{I} - \tilde{\mathbf{A}}^{\dagger}\mathbf{A} \right\|_1$



Noise level

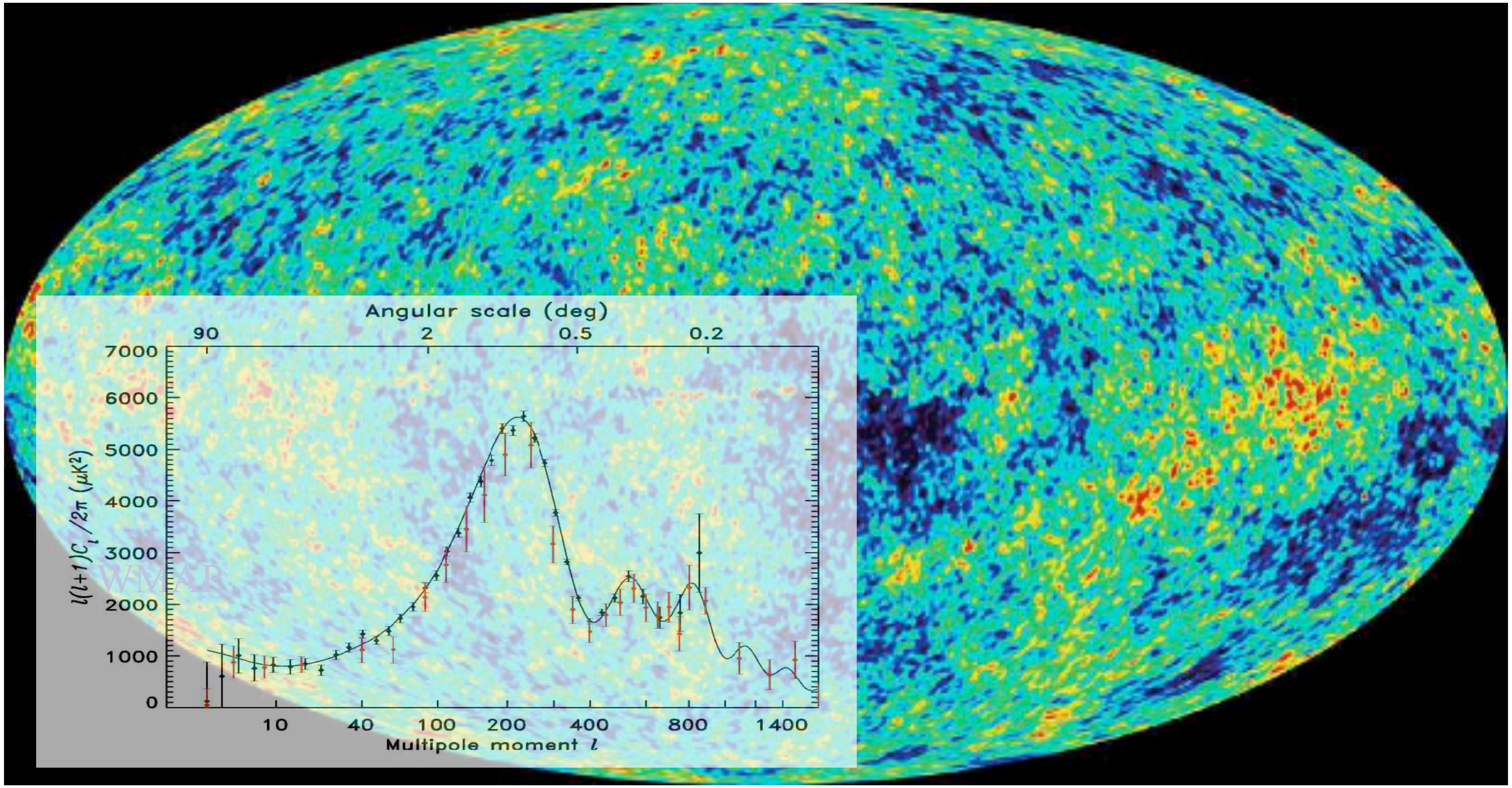
Application to astrophysics - take I

Application to CMB estimation

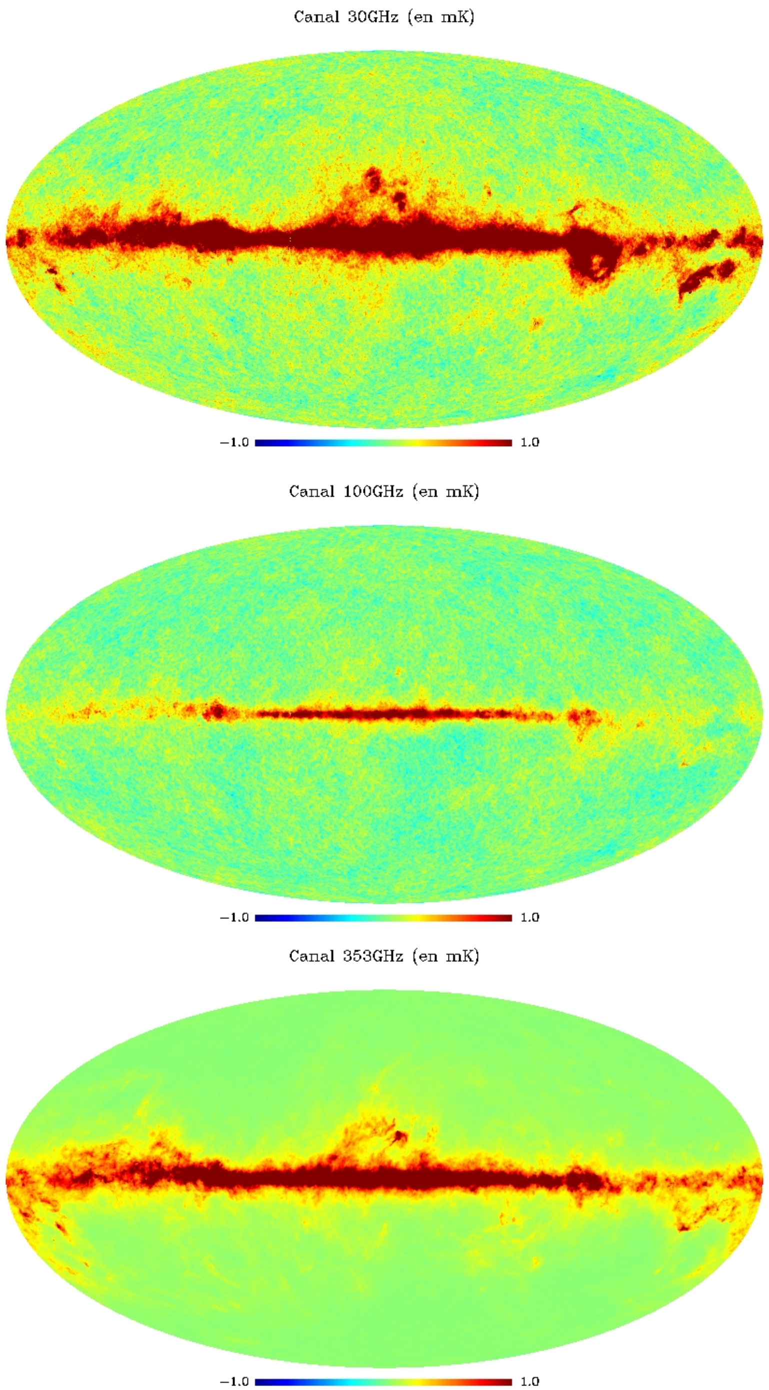


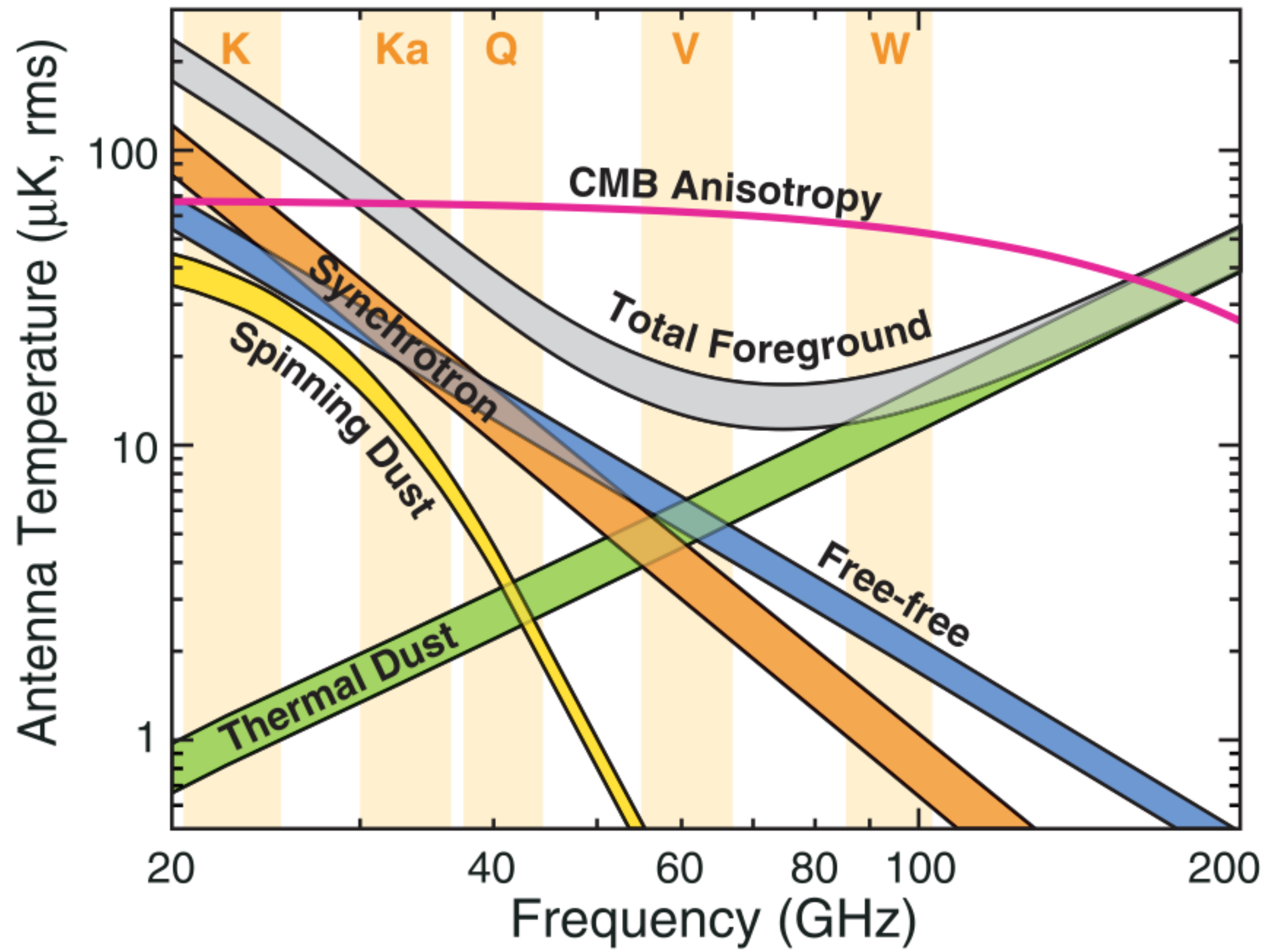
Cosmic Microwave Background (CMB)

The CMB exhibits Fluctuations



Application to CMB estimation

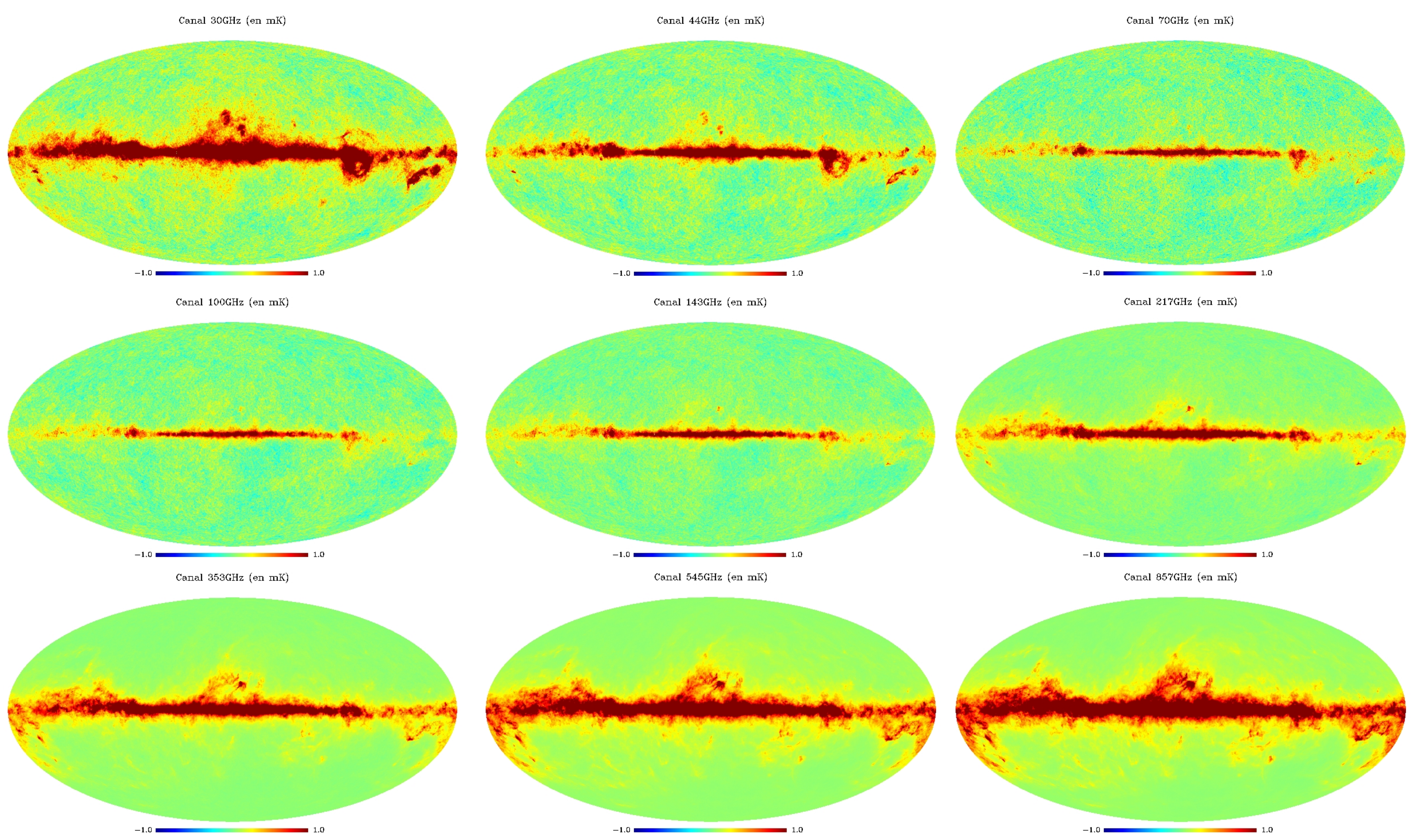




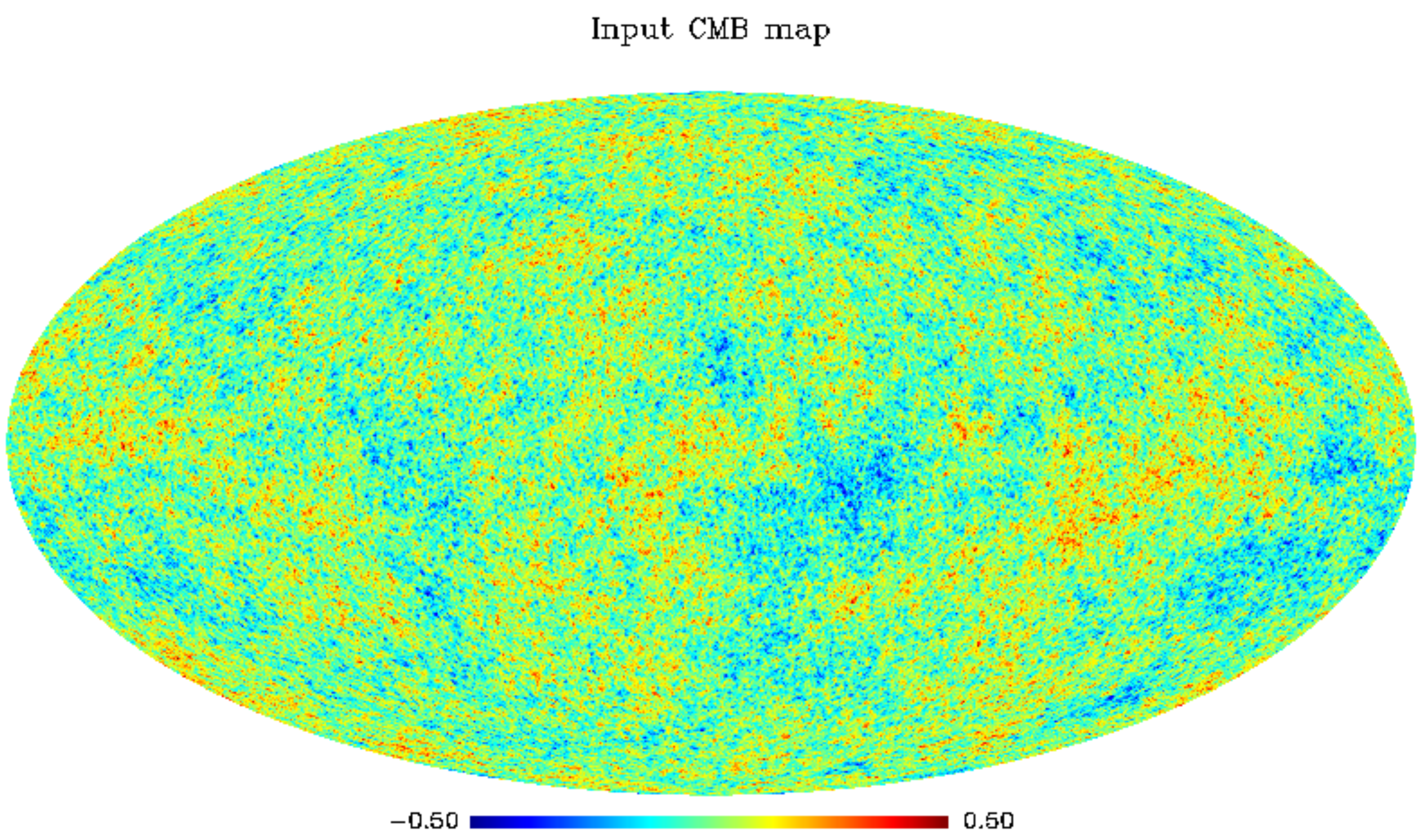
For the sake of evaluation, L-GMCA has been applied to simulated but realistic data (Leach et al. 2008)

Planck sky modeling : CMB, SZ, free-free, synchrotron and dust emission, spinning dust

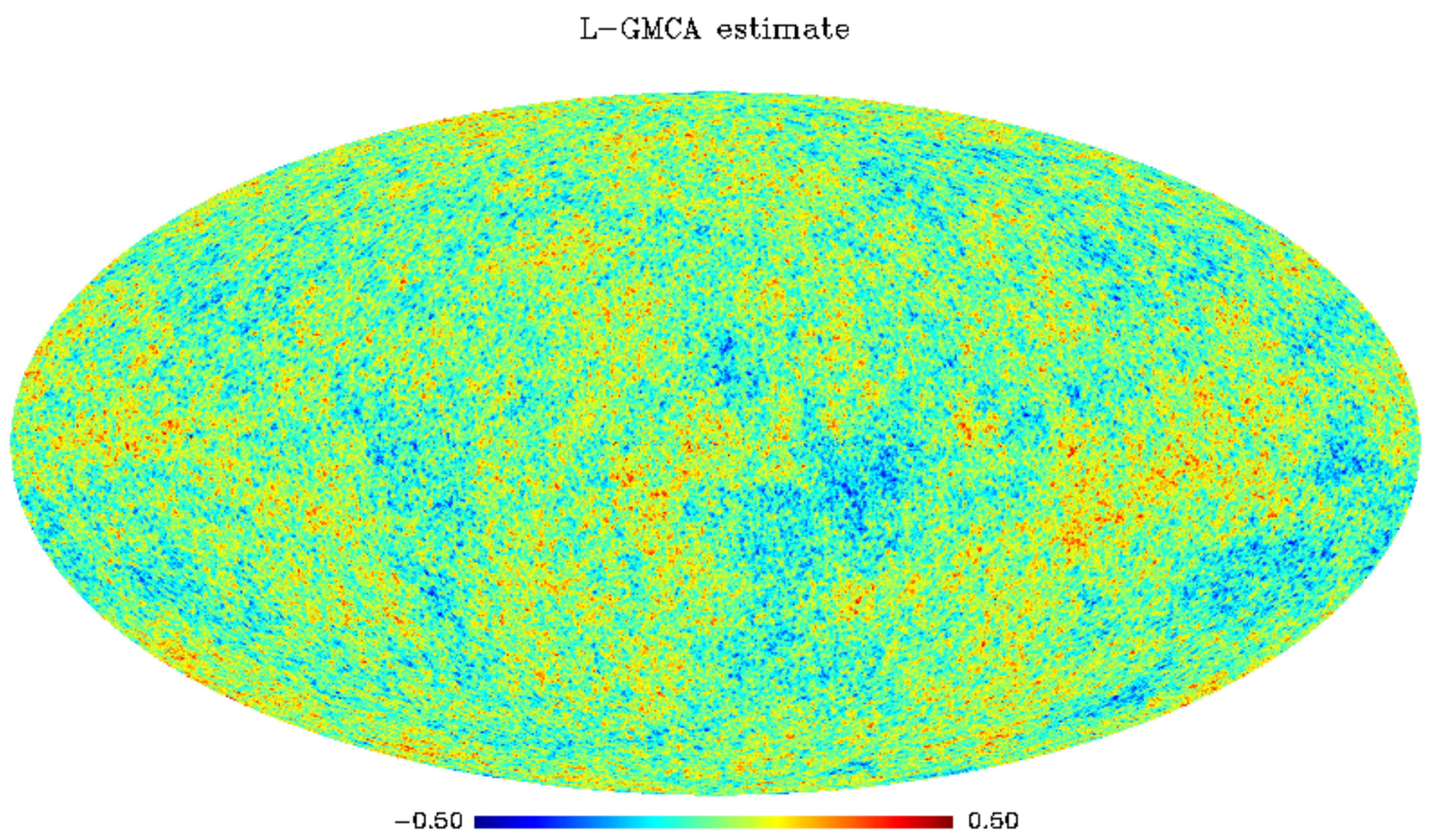
Instrumental modeling: decorrelated but non-stationary gaussian noise, perfect isotropic gaussian beams



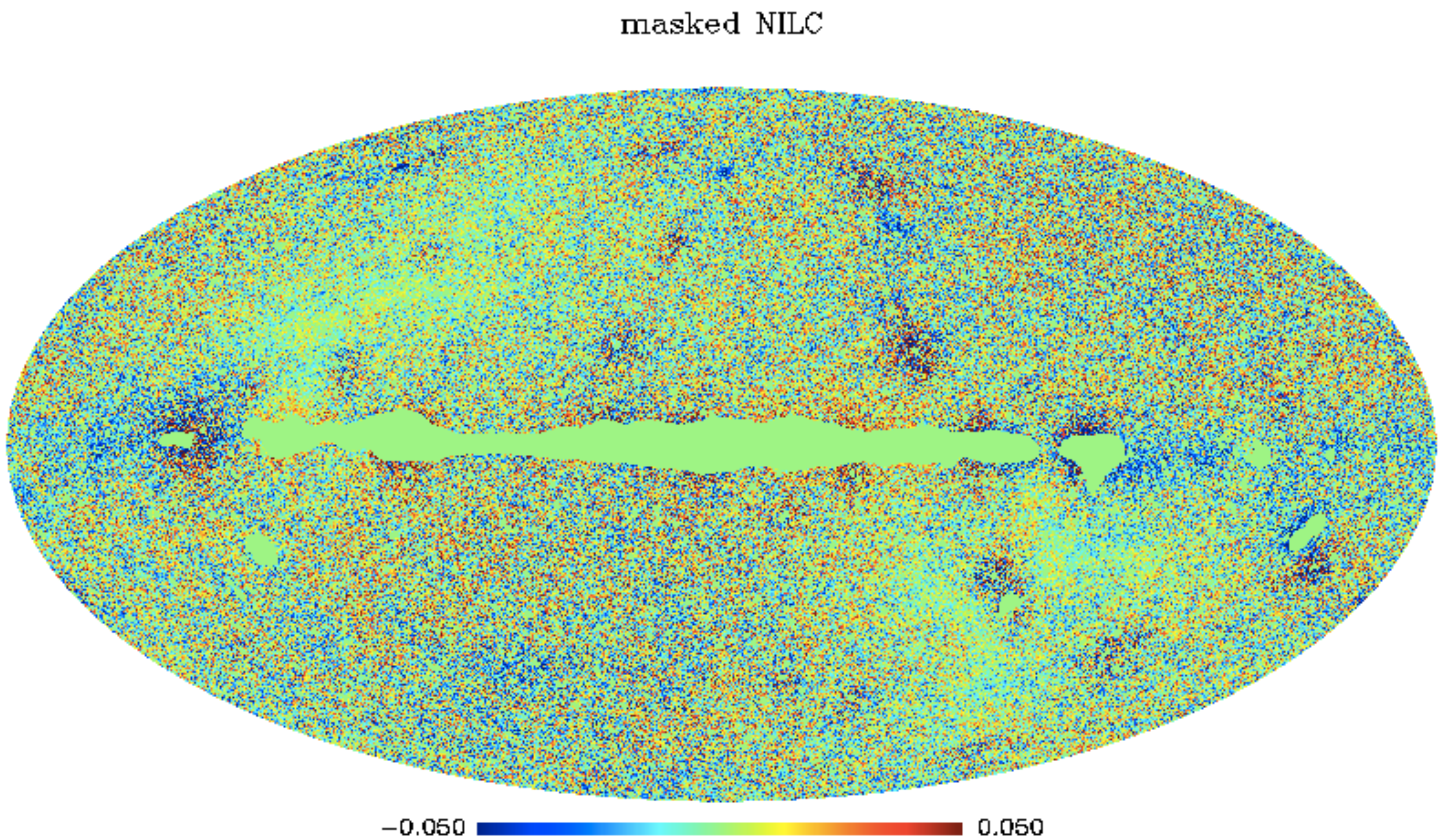
Results



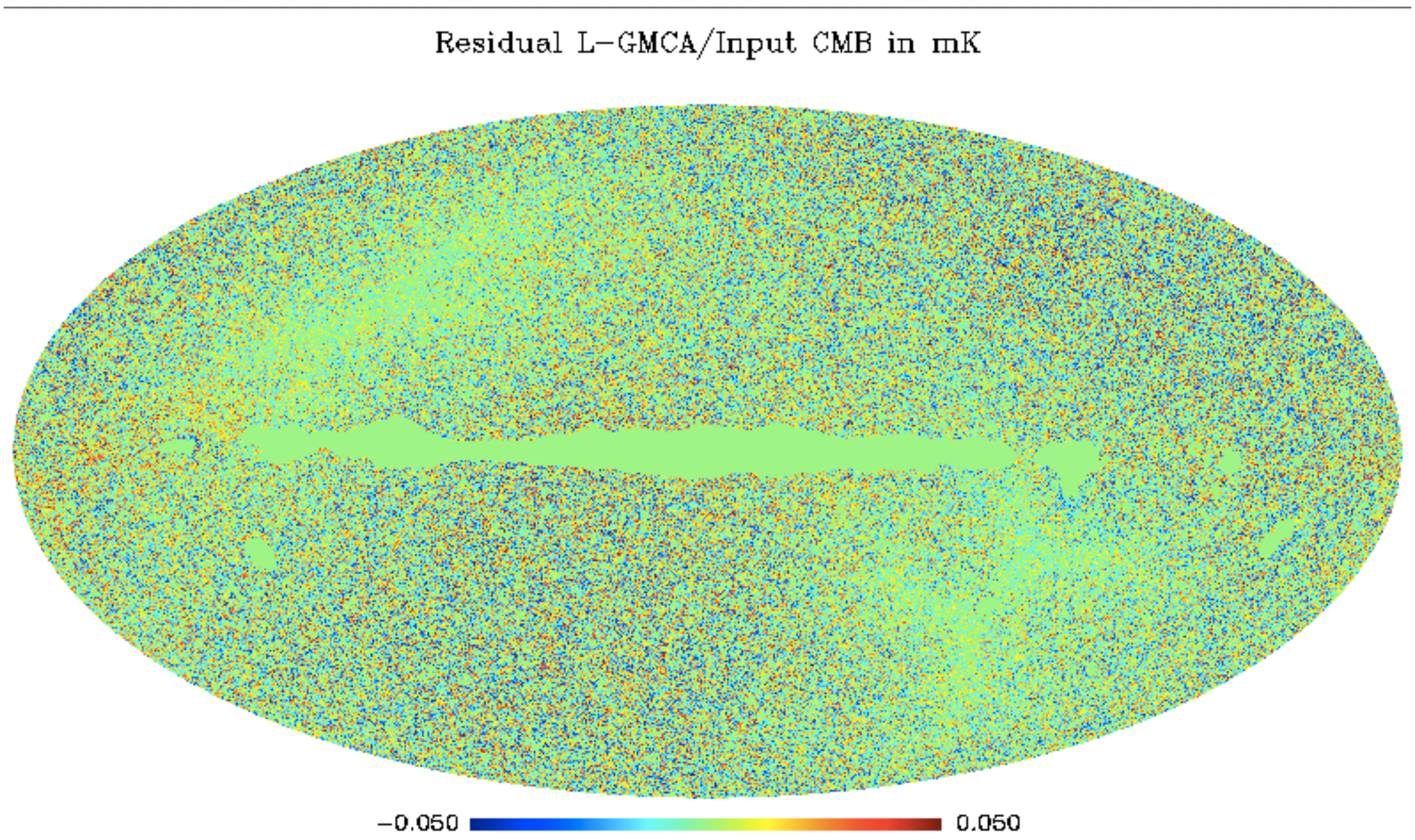
Application to CMB estimation



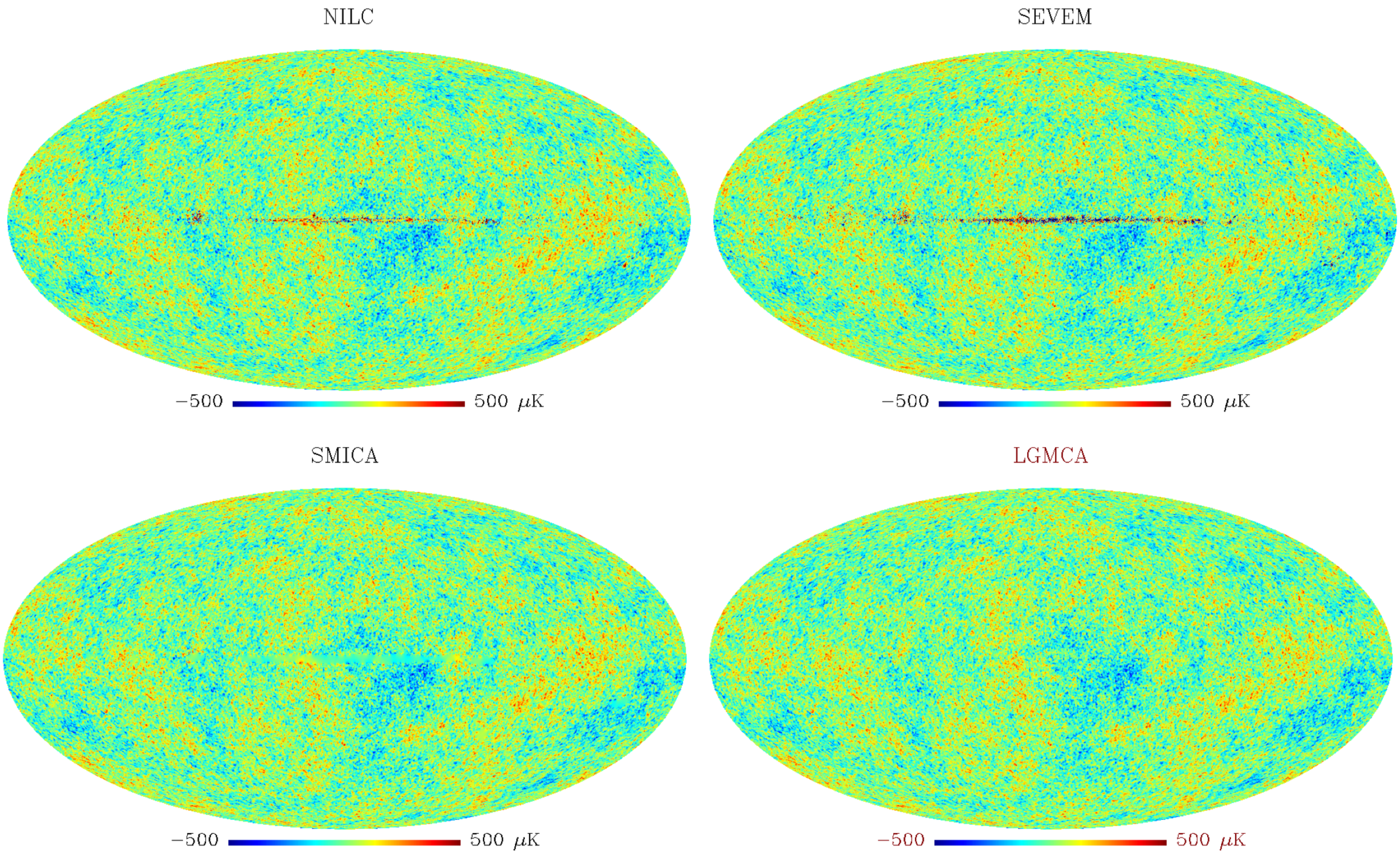
Application to CMB estimation



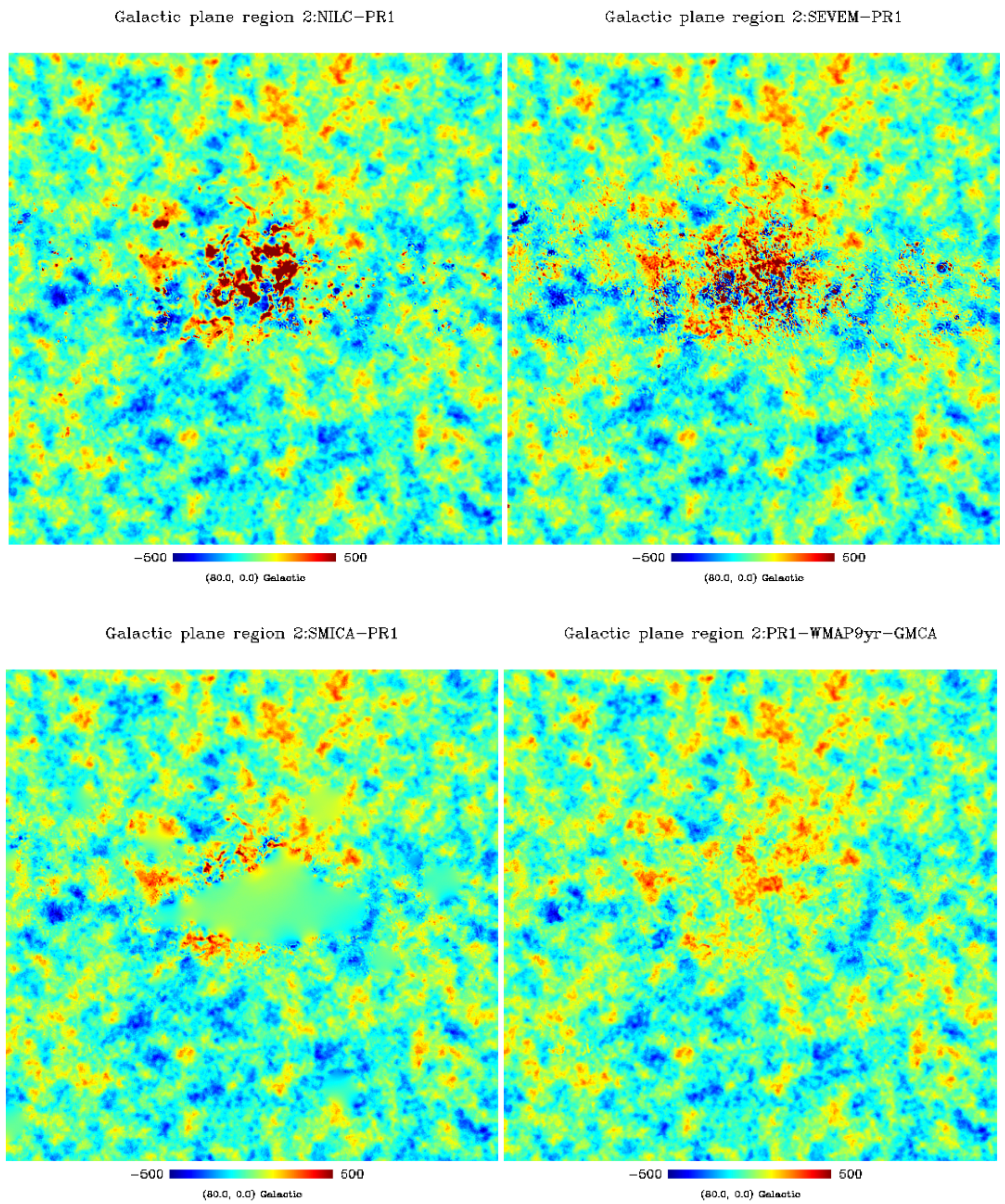
Application to CMB estimation

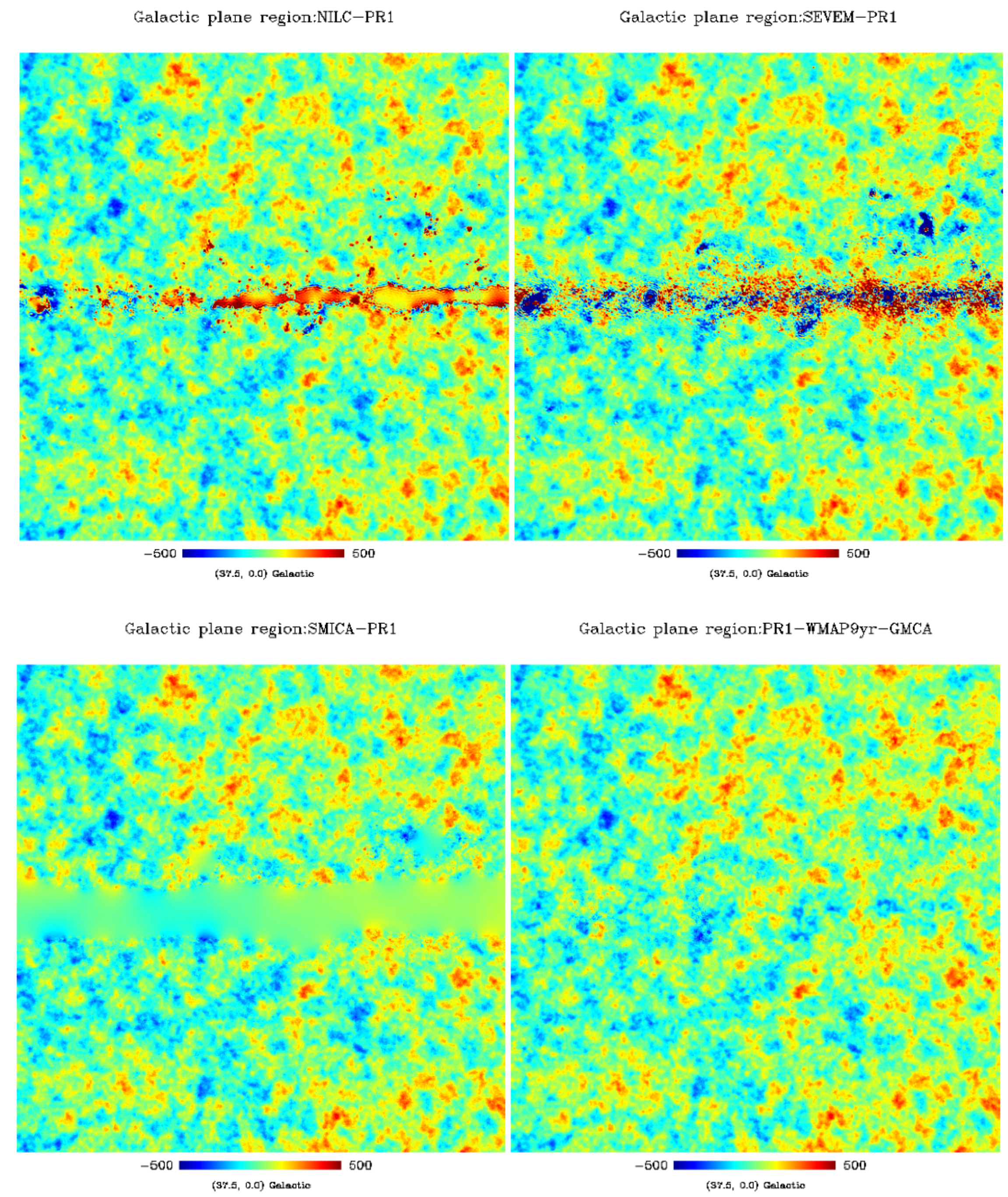


Real data !

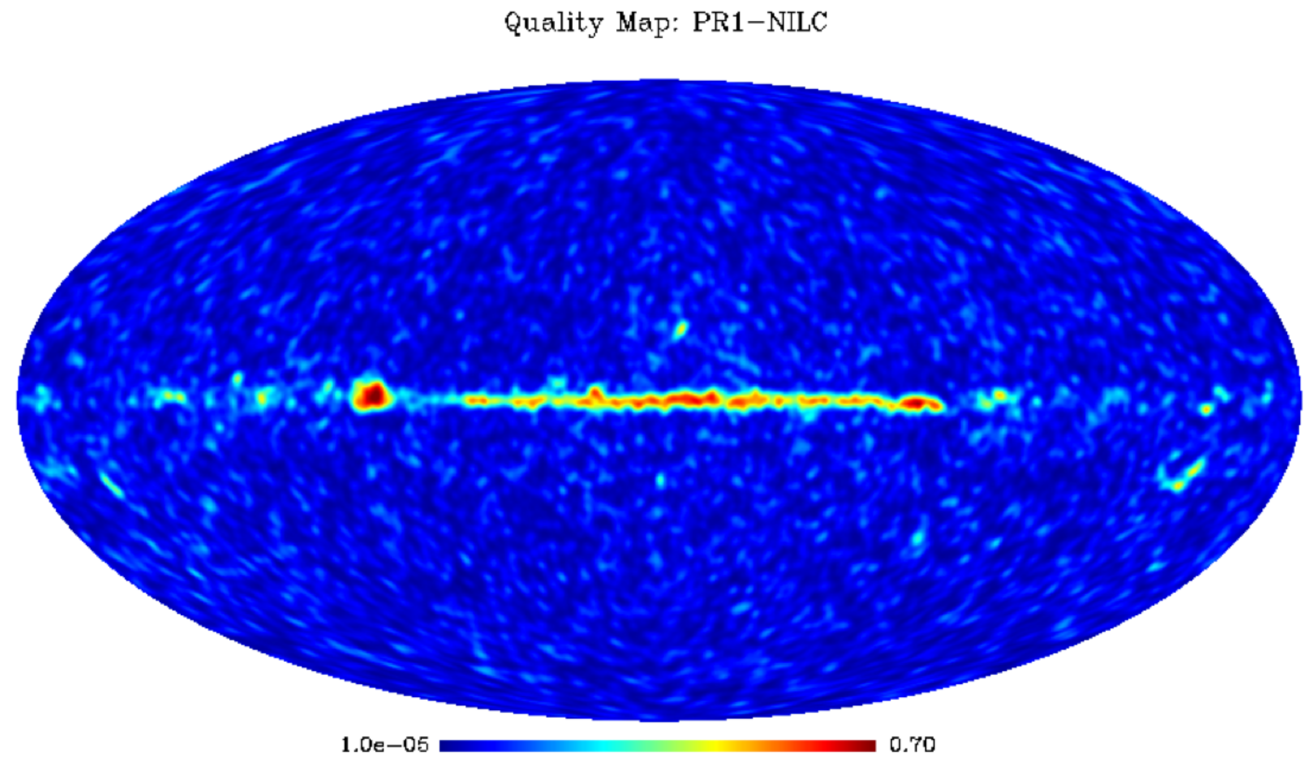

NILC
−500
500 μK
SEVEM
−500
500 μK
SMICA
−500
500 μK
LGMCA
−500
500 μK

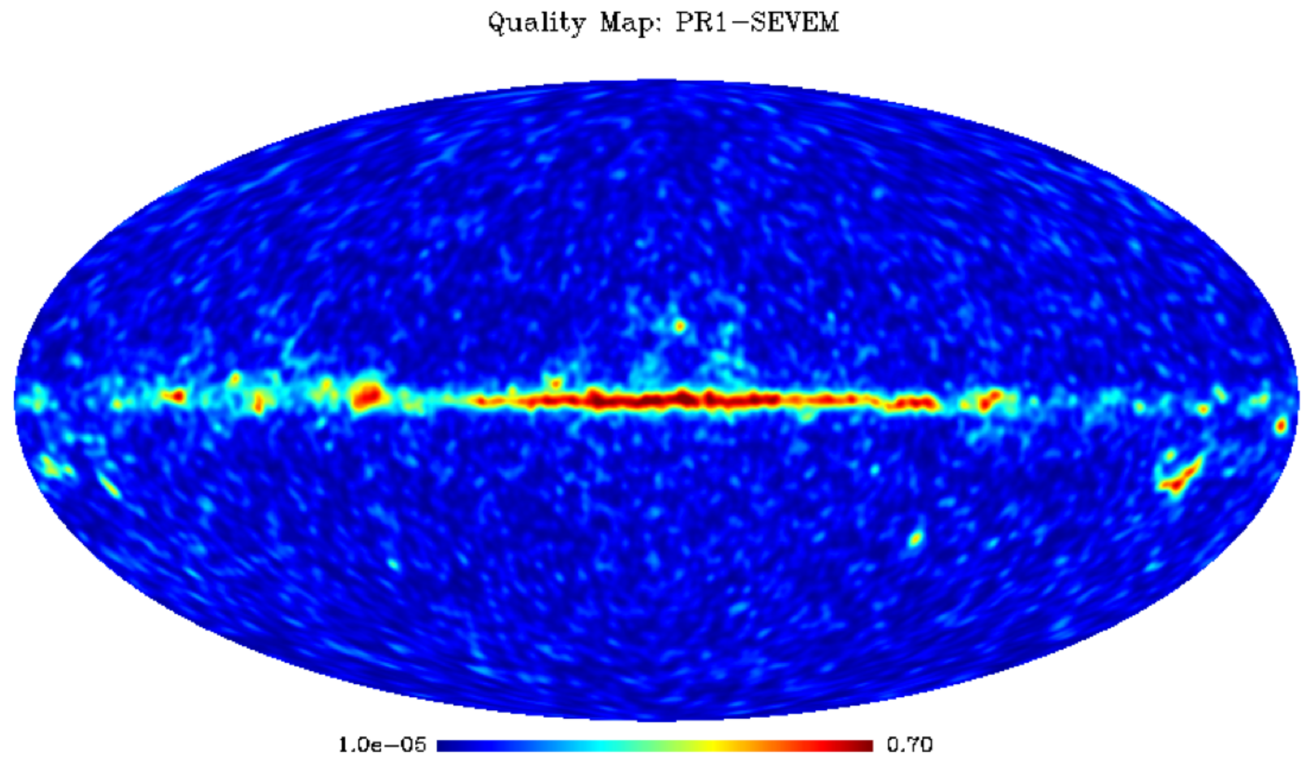
Application to CMB estimation

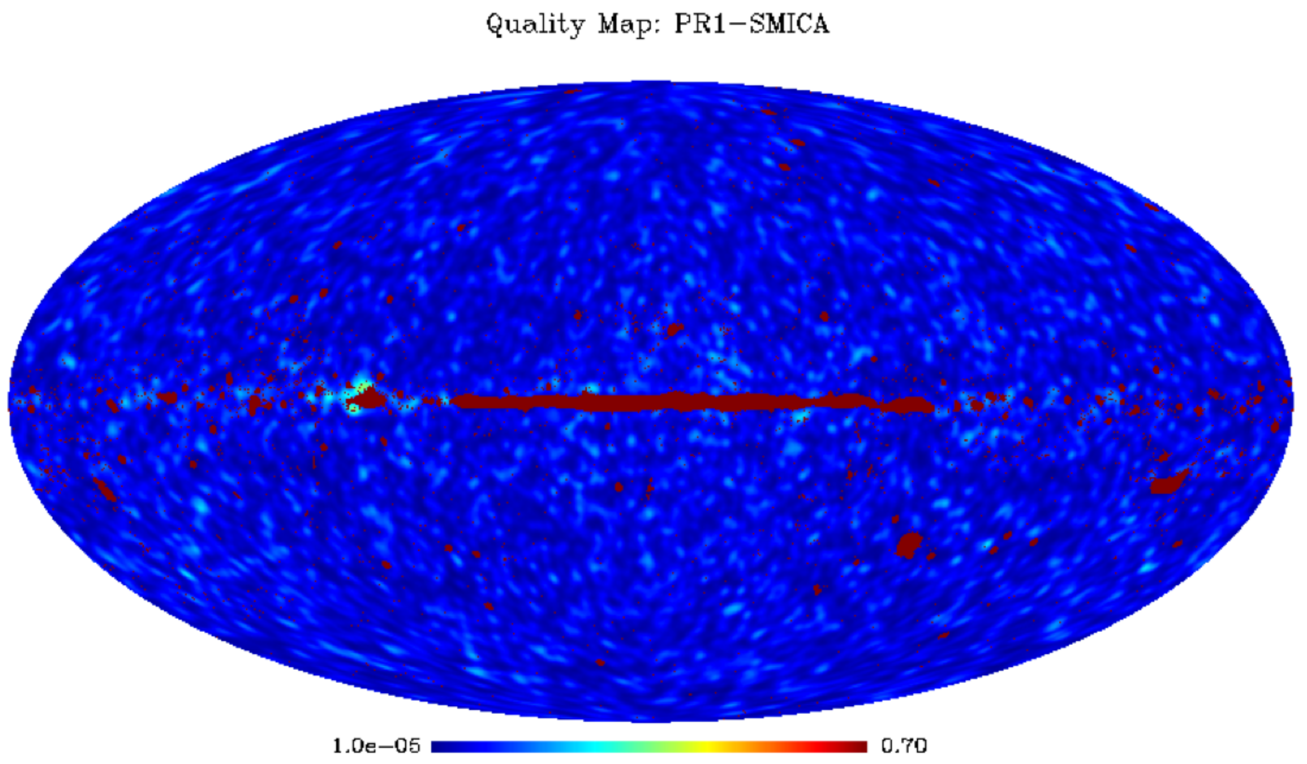


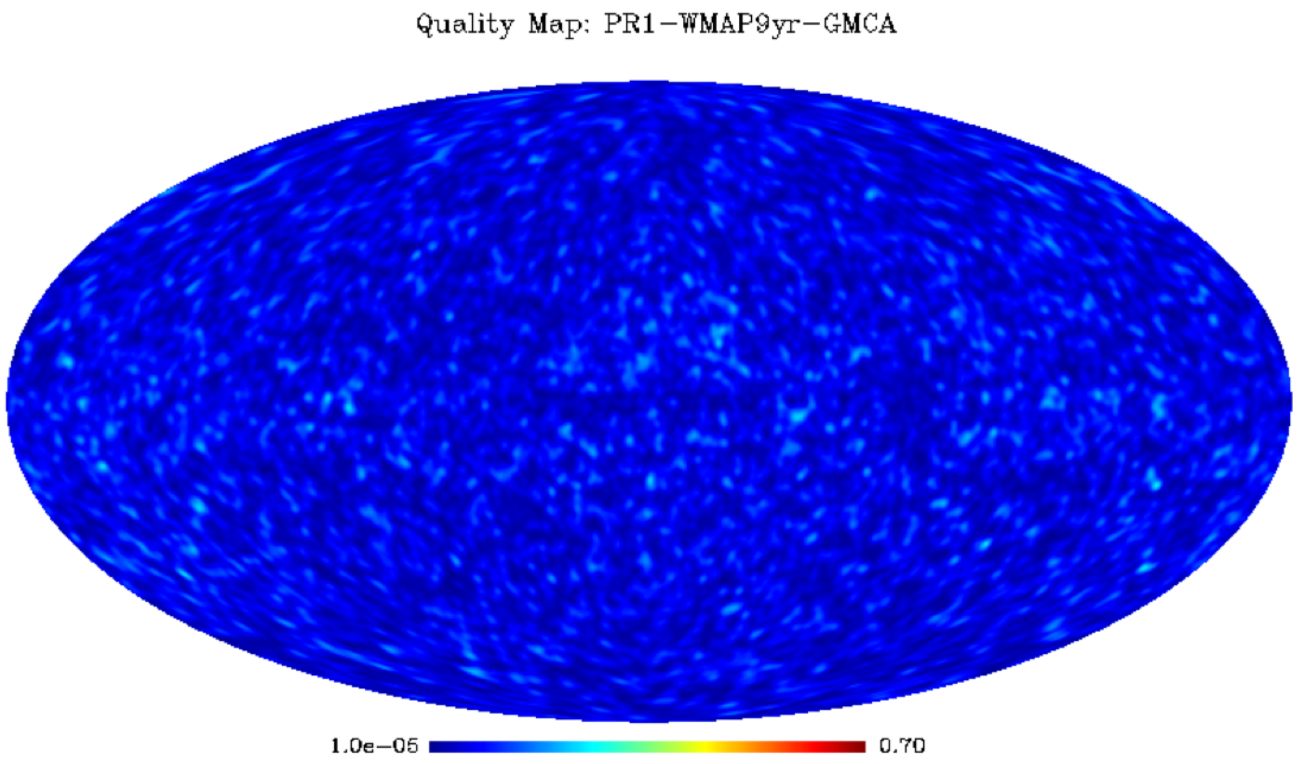


Real data !


Quality Map: PR1−NILC
1.0e−05
0.70


Quality Map: PR1−SEVEM
1.0e−05
0.70

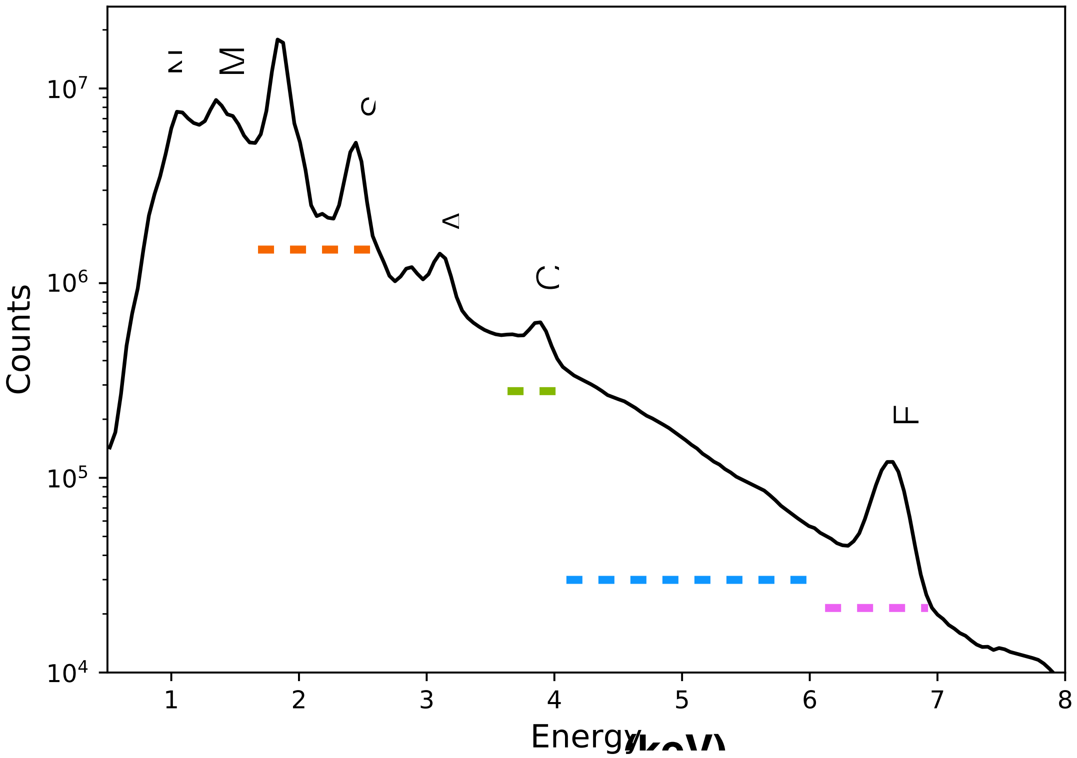

Quality Map: PR1−SMICA
1.0e−05
0.70

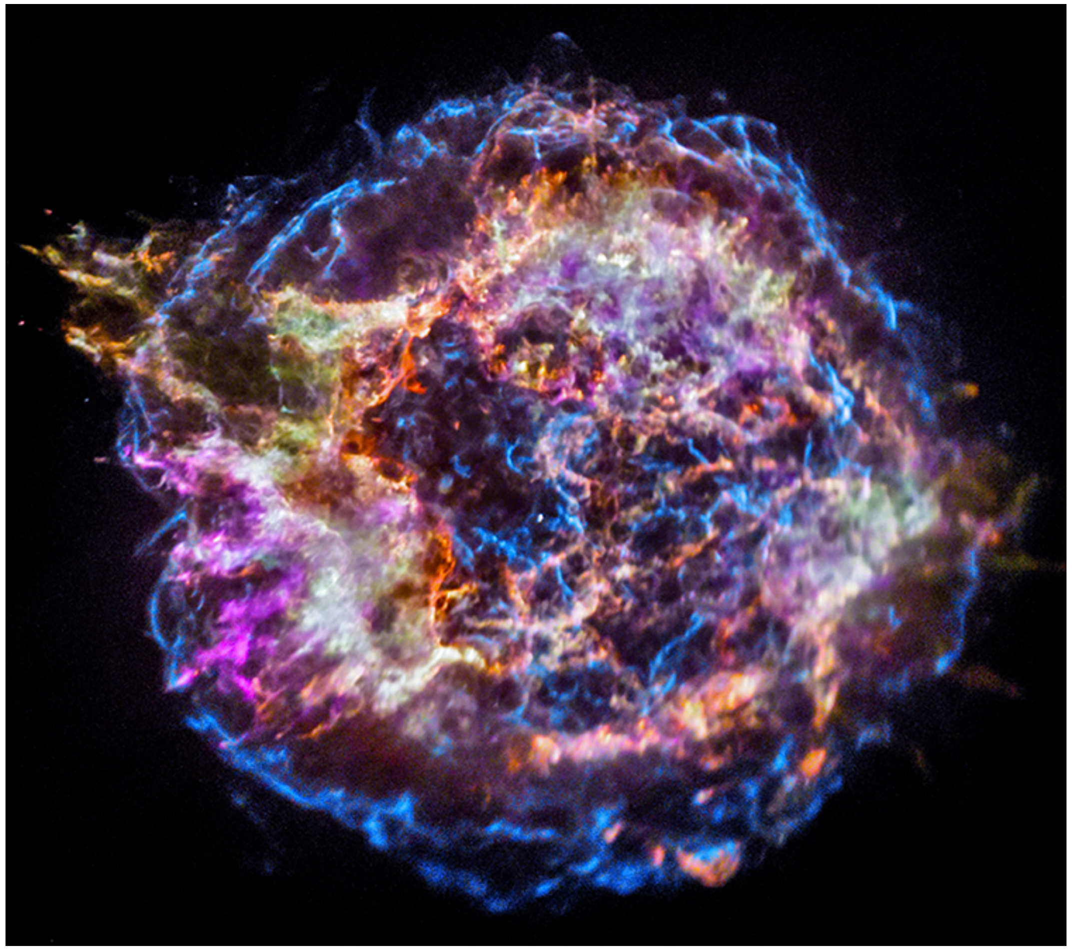

Quality Map: PR1−WMAP9yr−GMCA
1.0e−05
0.70

Application to X-ray imaging


Supernova remnants as seen in X-rays


- **Ejecta thermal emission gives insight on :**
  - **Individual elements distribution**
  - **Morphology, asymmetries**
  - **Velocities**

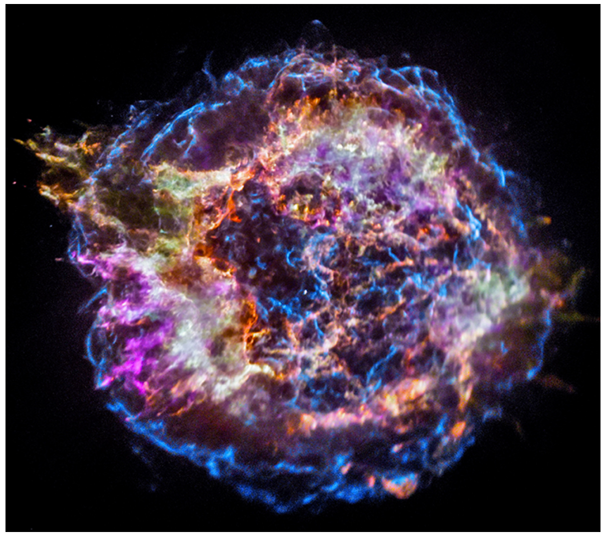


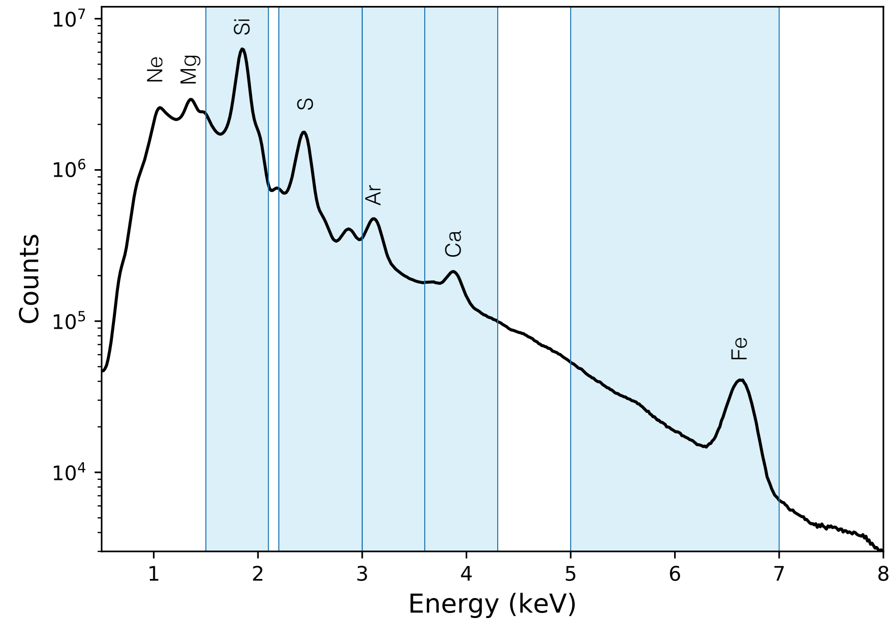


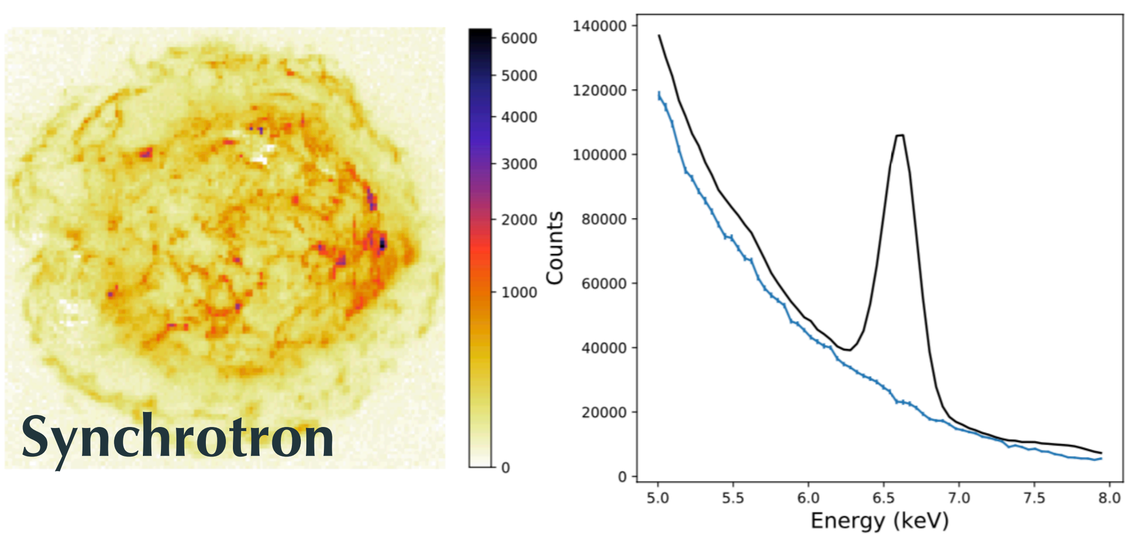
Hard unsupervised component separation problem:

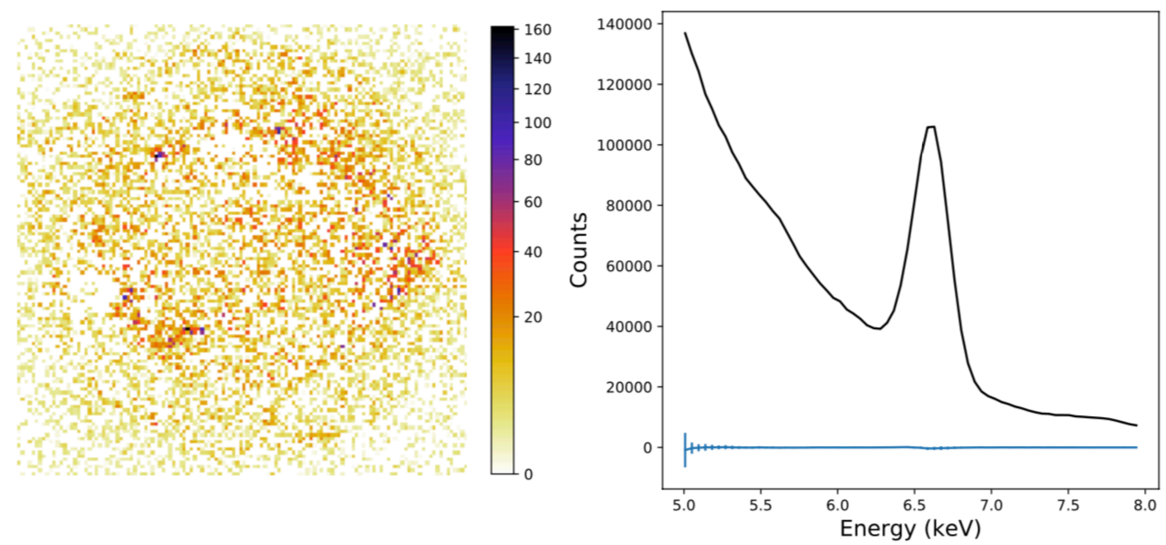
- **High dynamic range between the sources** (> 2 orders of magnitudes)
- **Low signal-to-noise ratio**
- **Strong synchrotron background**
- **Correlated spectra and sources**

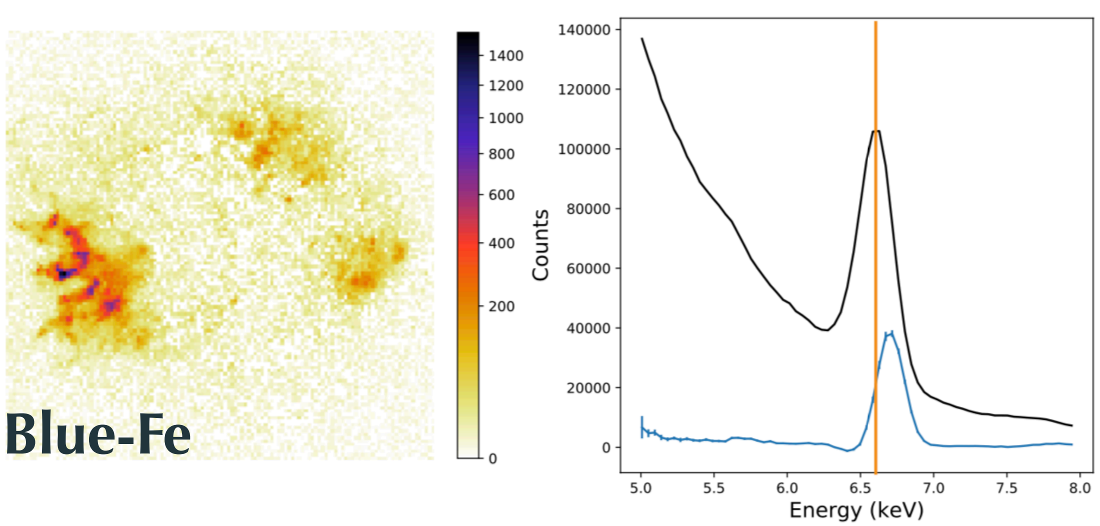
Sparse component analysis

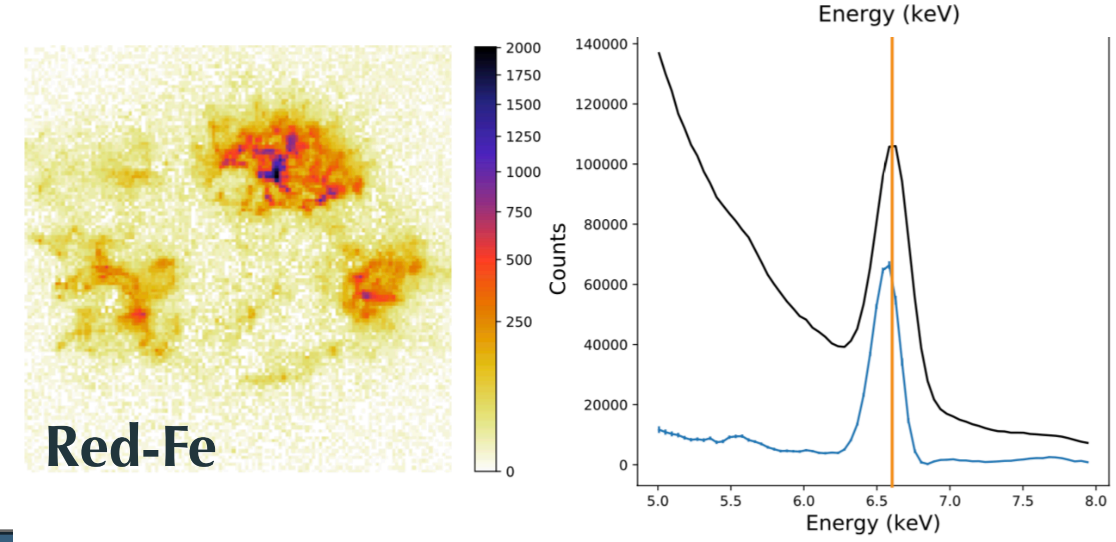




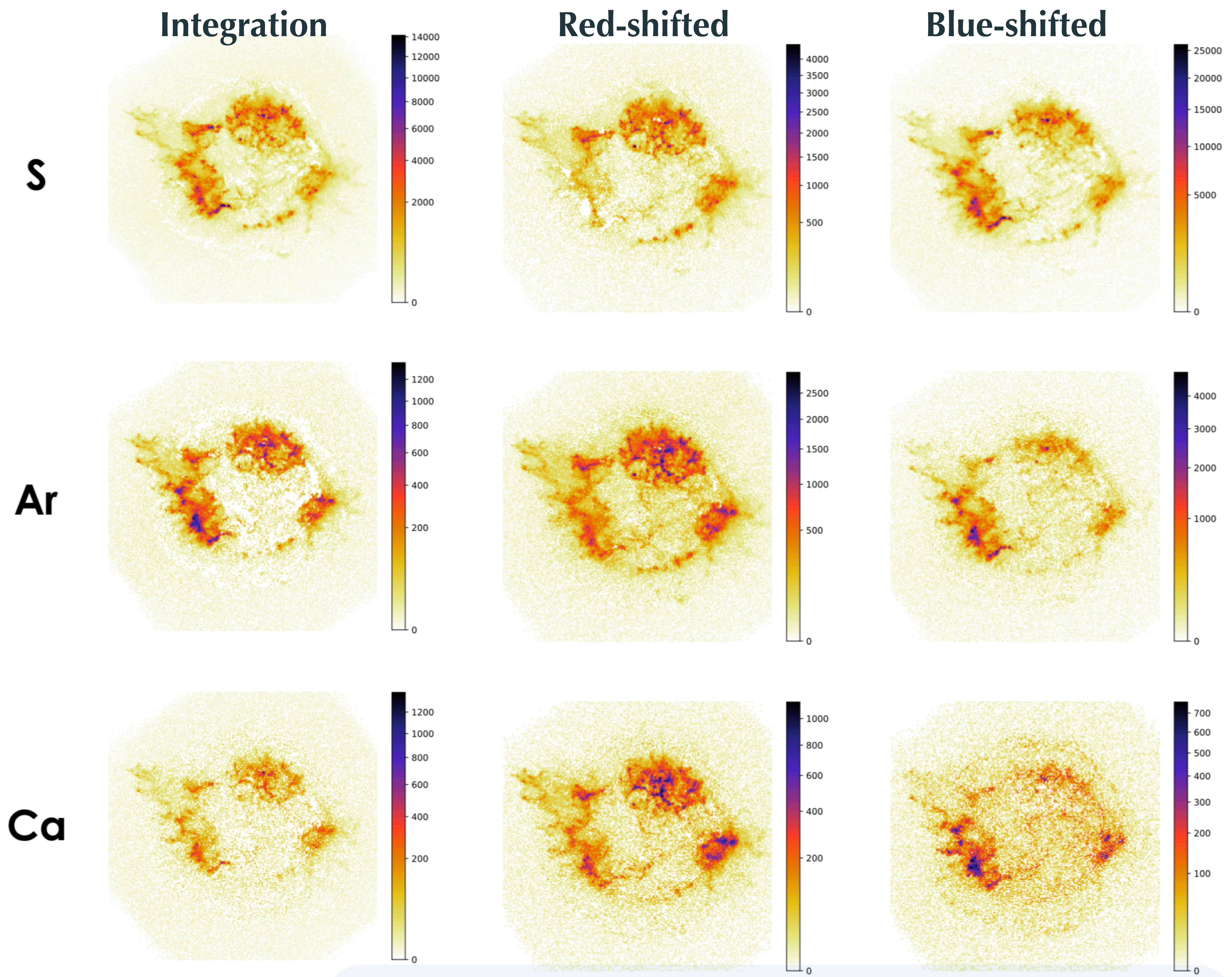








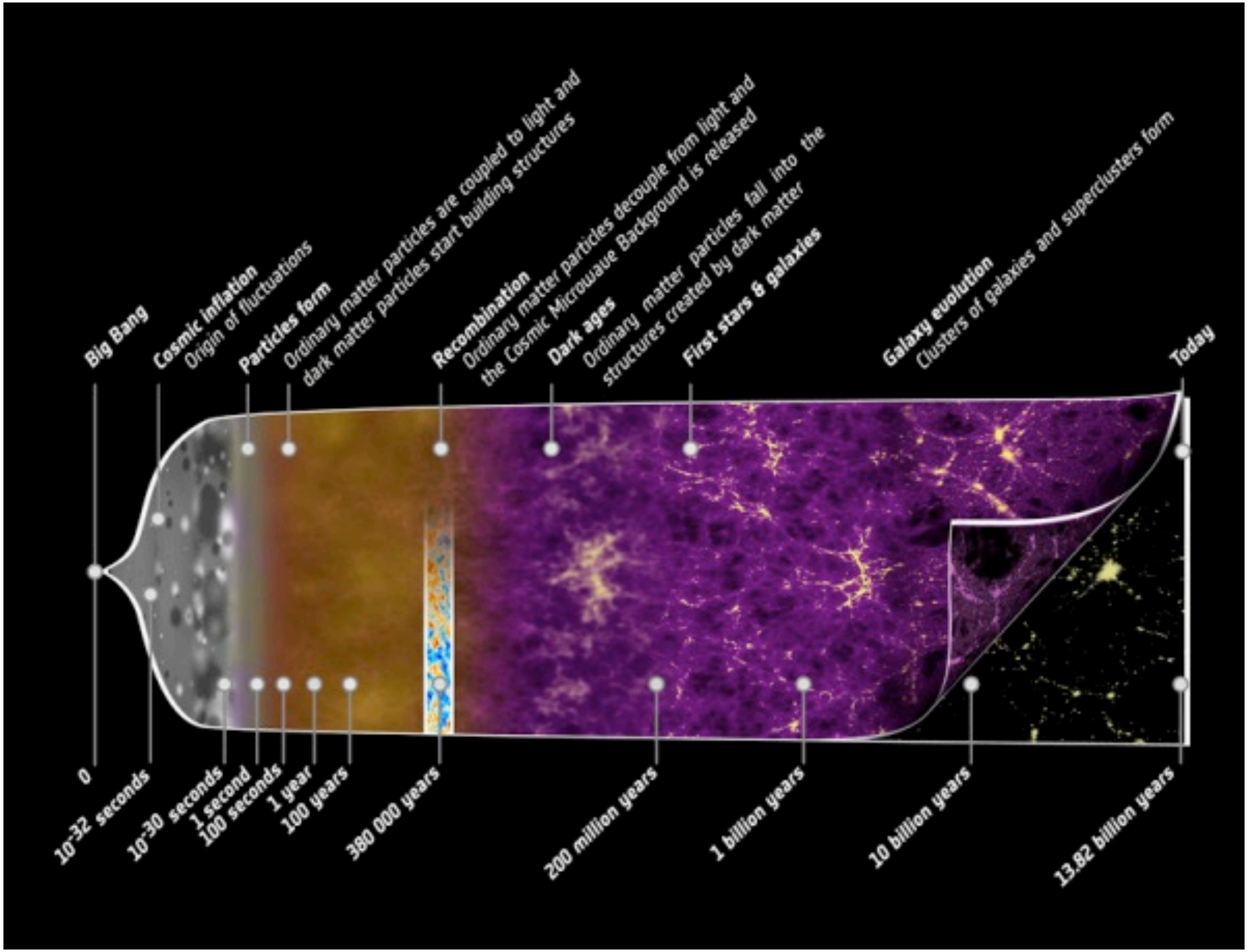
Sparse component analysis



Picquenot et al, 2020.

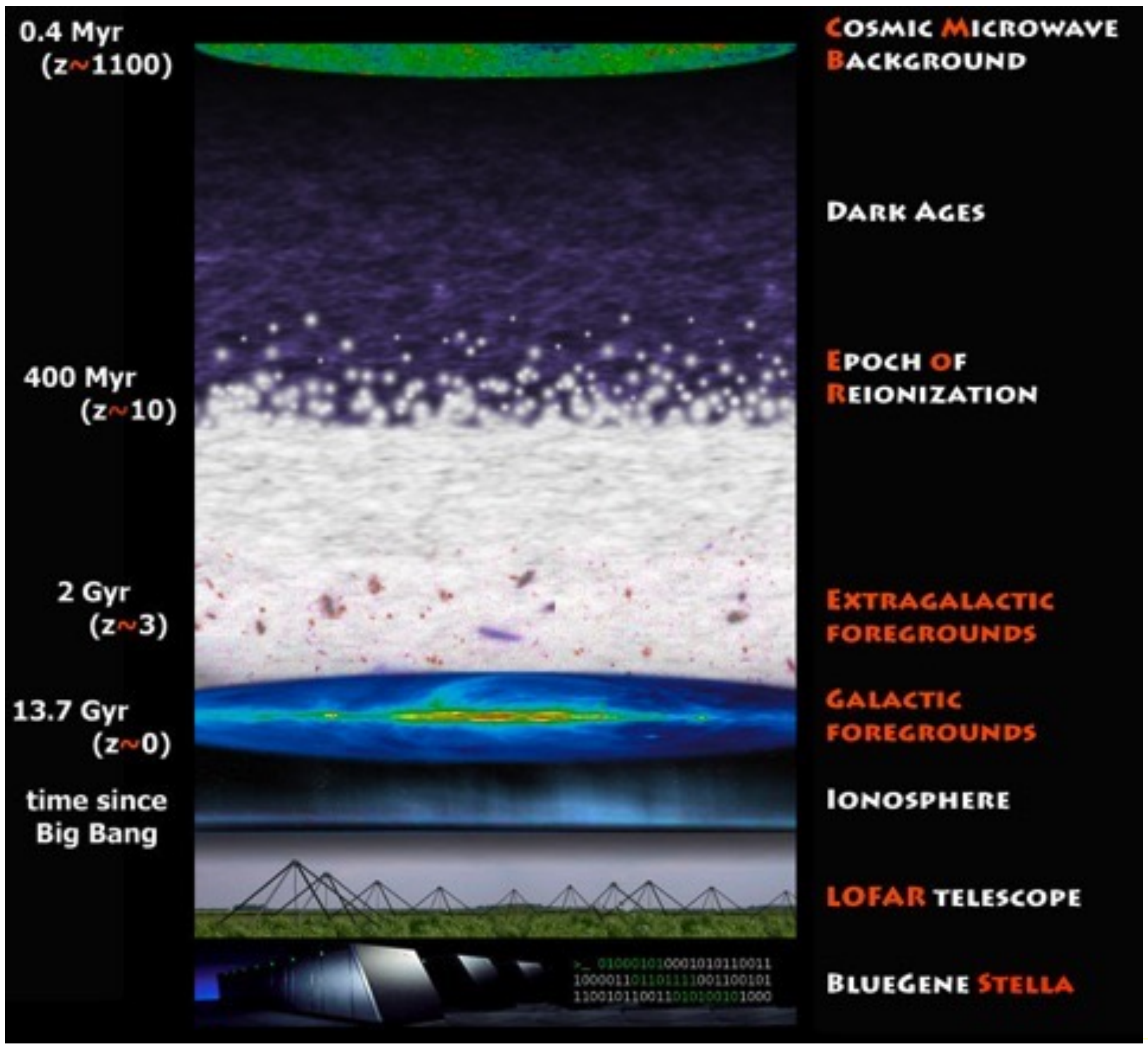
Blindly estimates red/blue-shifted atomic components !

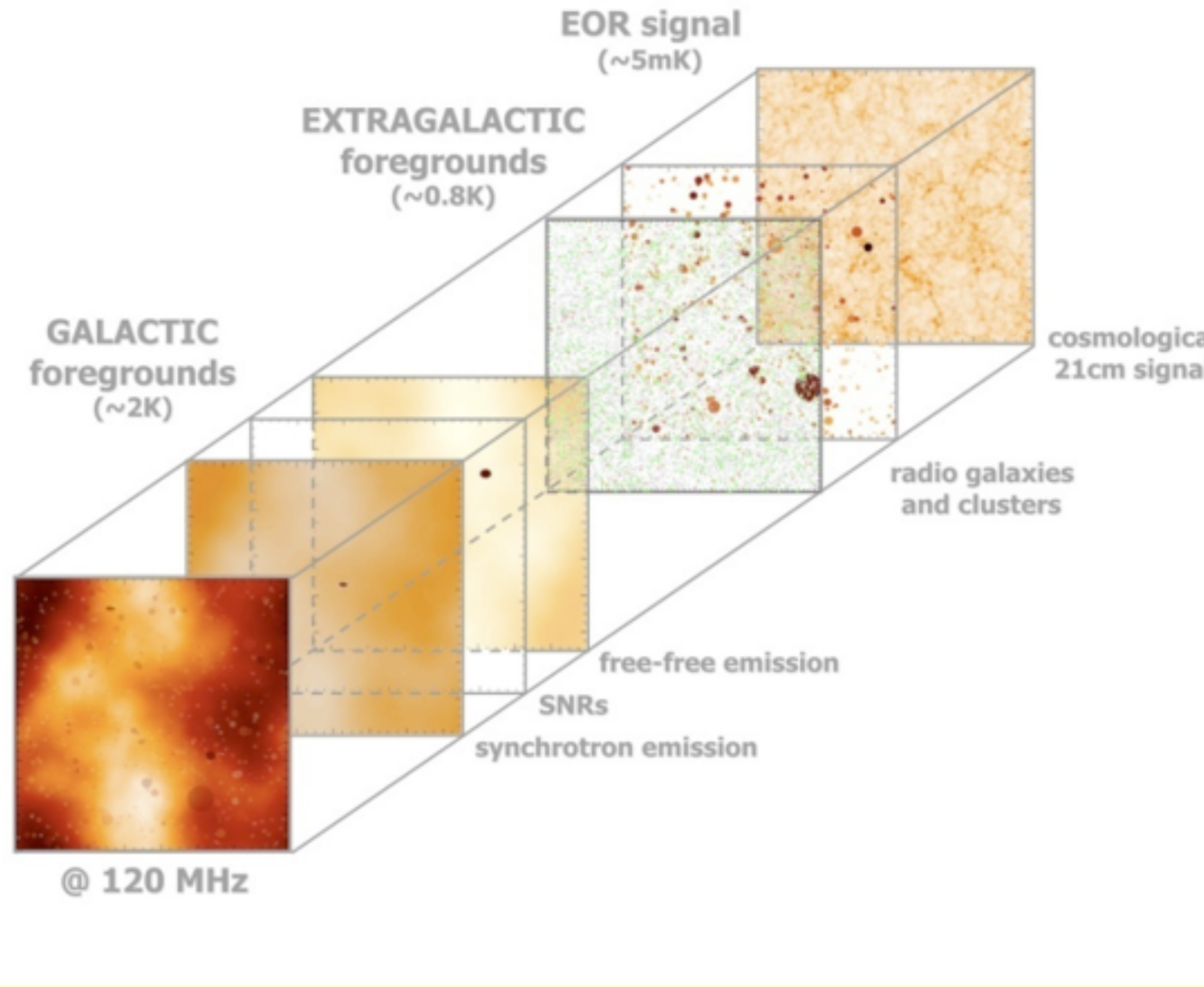
Radio-astronomy



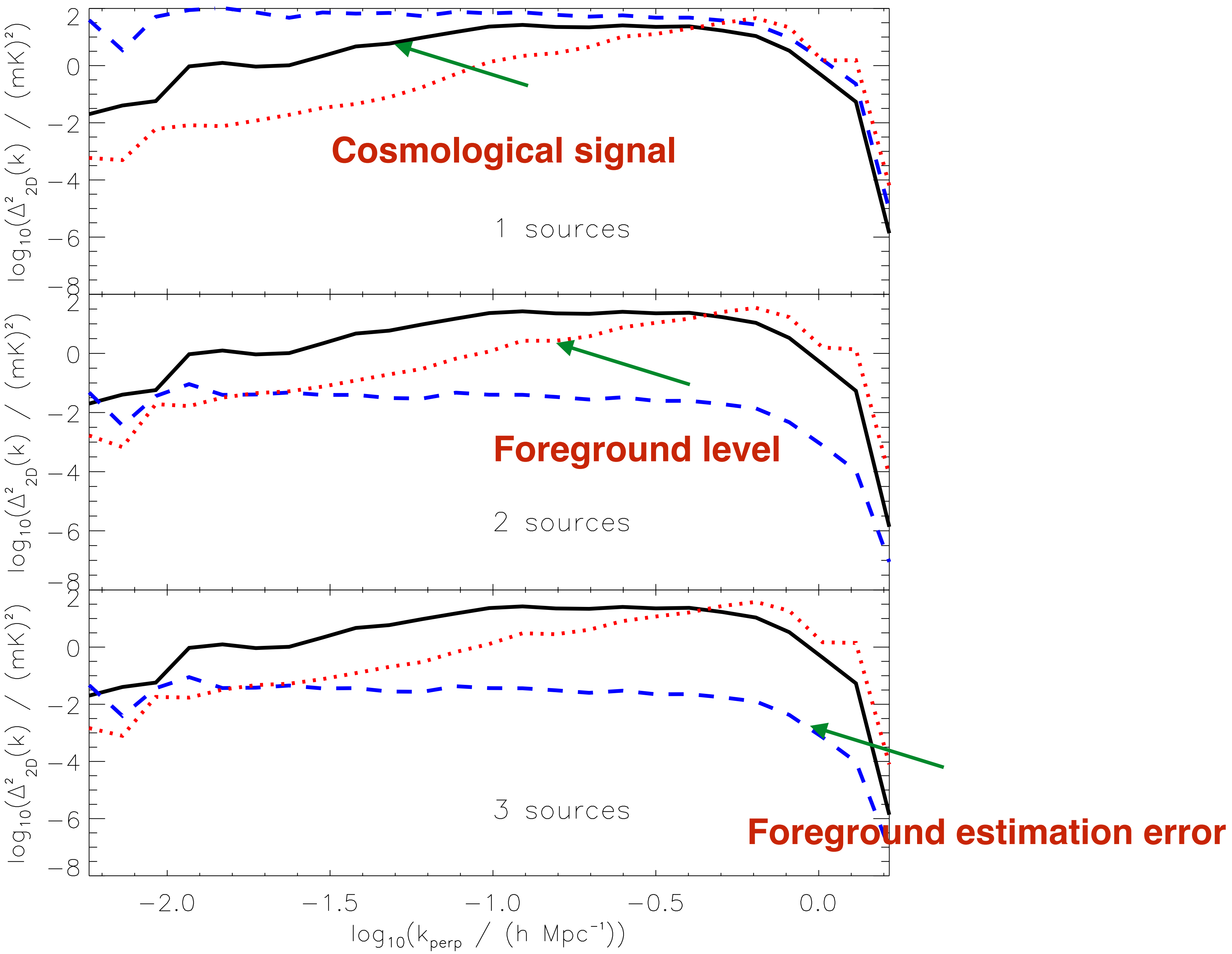
Radio-astronomy

Lofar (Low-frequency array): - an array of radiotelescopes spread all over Europe (interferometry)
- EoR project : observing redshifted 21cm rays lines (in the range 110 to 200MHz) to characterize early epochs of reionization (Dark ages) after the recombination





The scale of the problem: recovering images of reionization with GMCA - Chapman et al., MNRAS, 2013, Vol. 429



Chapman et al., "The Scale of the Problem : Recovering Images of Reionization with GMCA", 2013