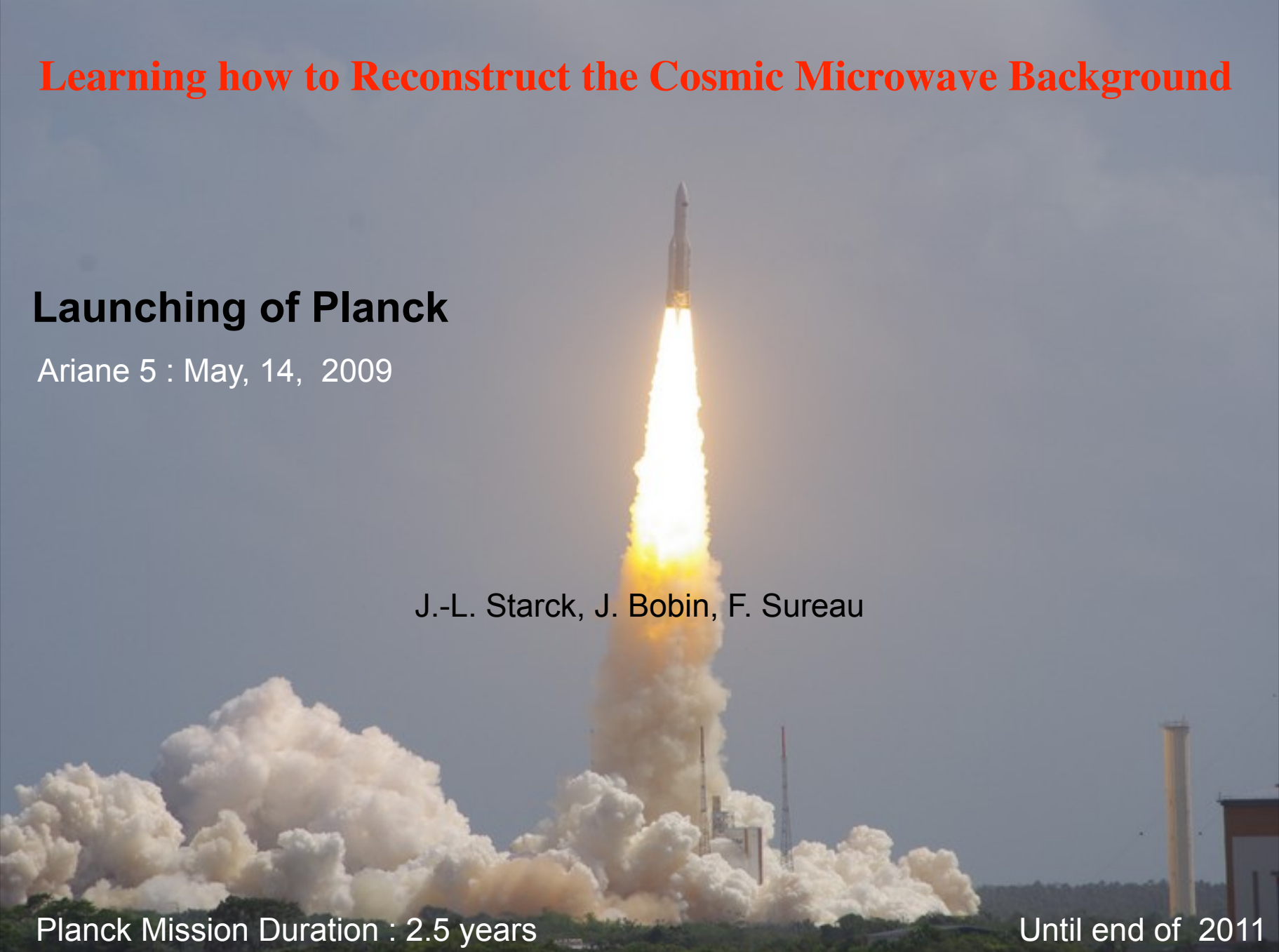
Learning how to Reconstruct the Cosmic Microwave Background



Launching of Planck

Ariane 5 : May, 14, 2009

J.-L. Starck, J. Bobin, F. Sureau

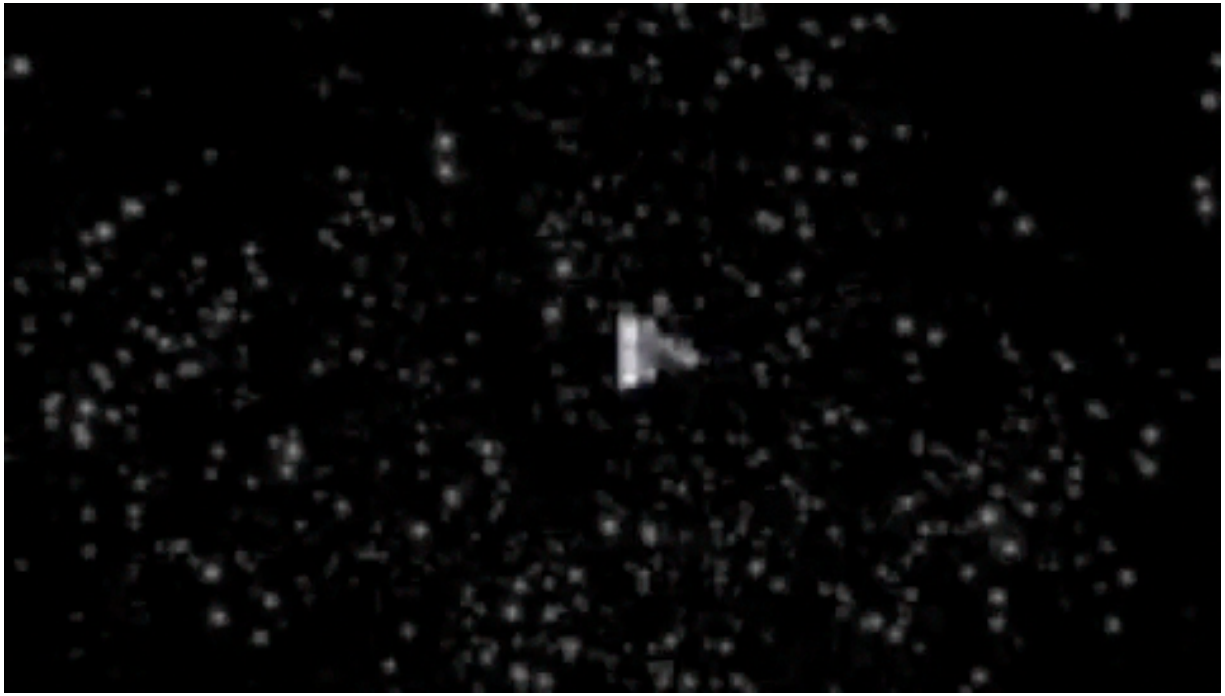
Planck Mission Duration : 2.5 years

Until end of 2011

The Cosmic Microwave Background

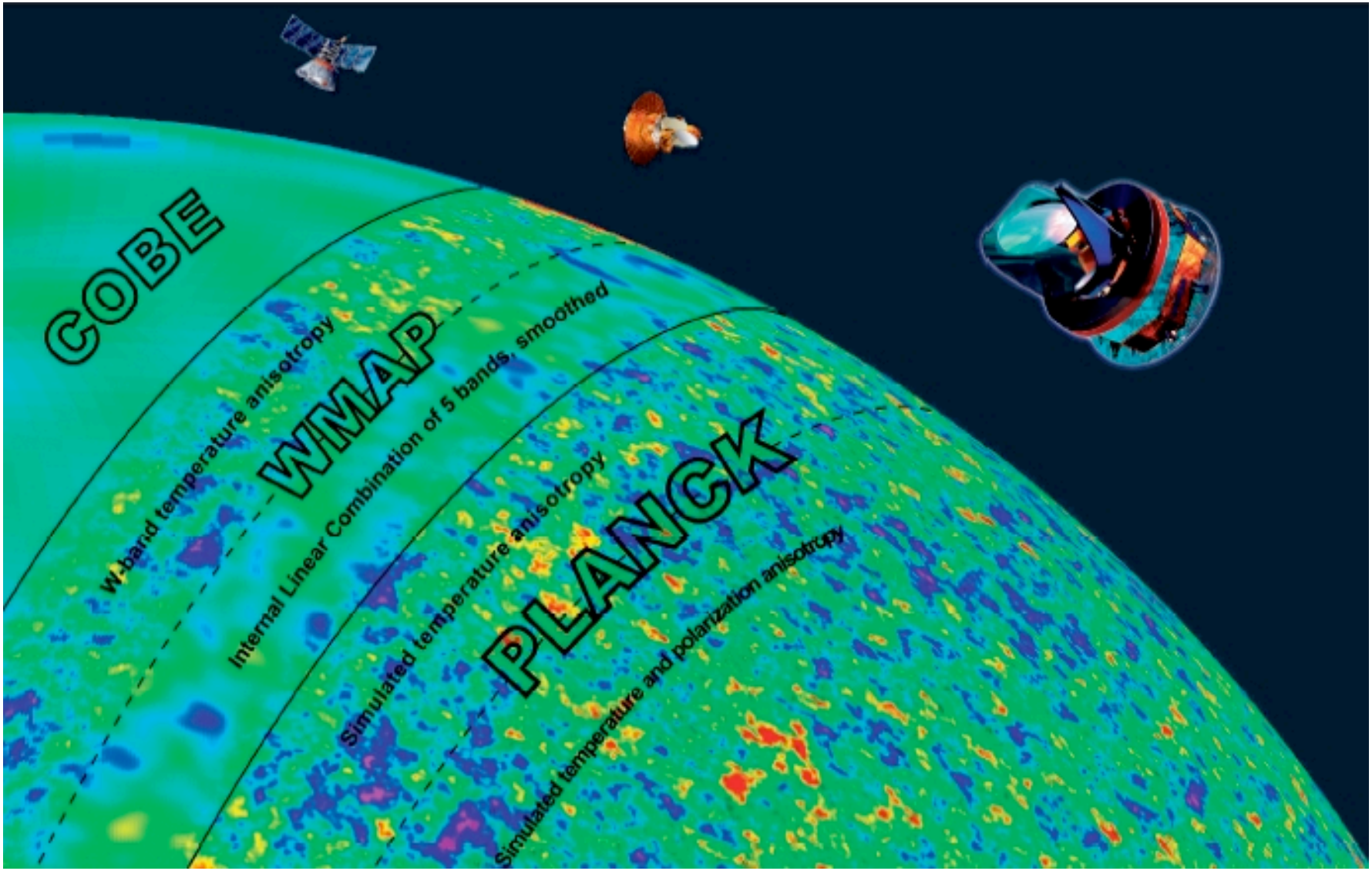
- The Universe is filled with a blackbody radiation field at a temperature of 3K.
- Predicted by G. Gamow in 1948
- Observed for the first time by Penzias and Wilson (1965)
- Confirmed by COBE (1990)
- Spectacular measurement of anisotropies by WMAP
- WMAP observed the CMB since 2002. Fifth and Last data release in August 2011.
- PLANCK first cosmological results in January 2013.

PLANCK and Sparsity

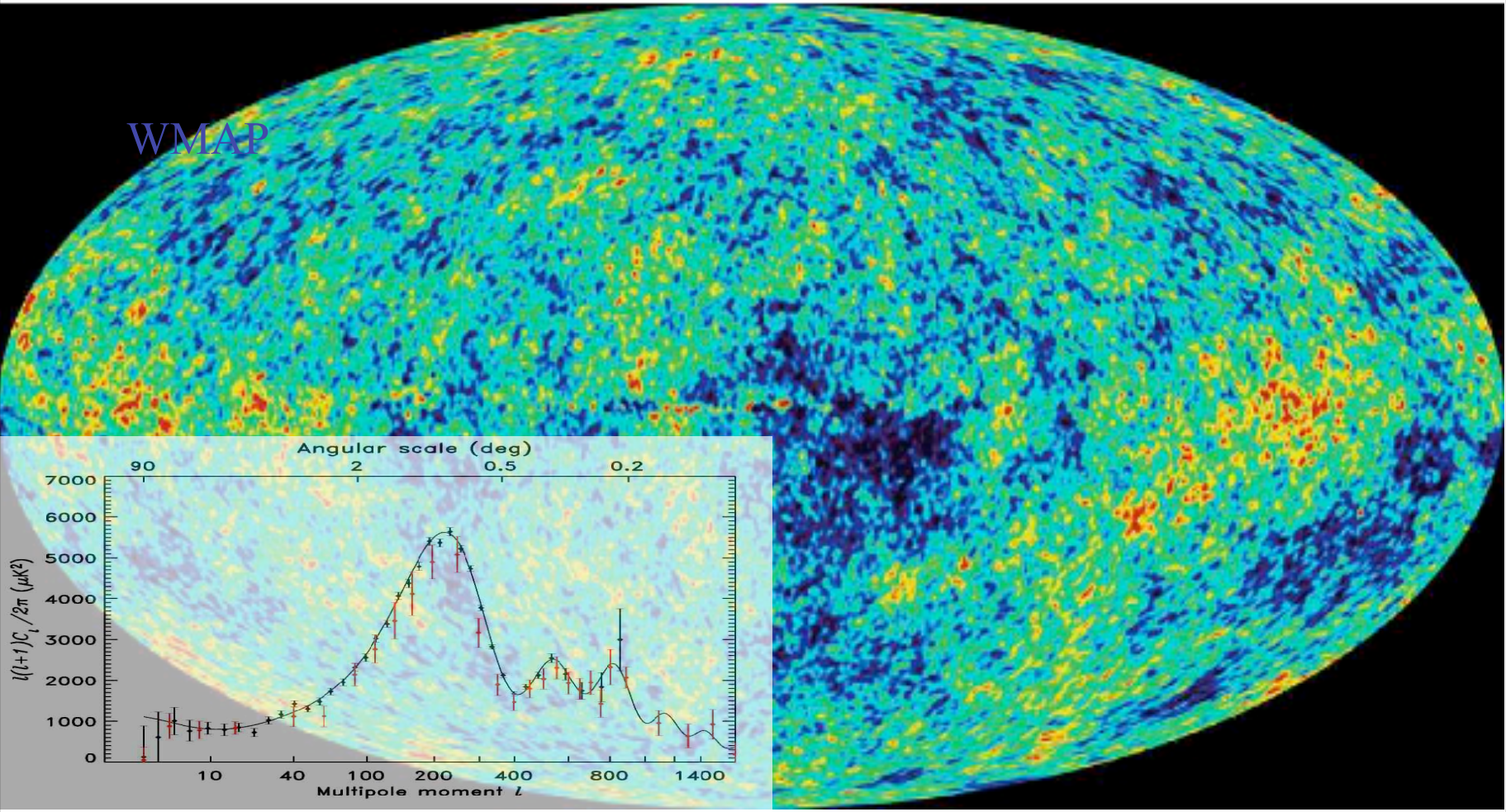


- **Successor of WMAP (better resolution, better sensitivity, more channels)**
- **Launched on May 14, 2009**
- **Two instruments LFI and HFI**
- **Nine Temperature maps at 30,44,70,100,143,217,353,545,857 GHz + Polarization**
- **Angular resolutions: 33', 24', 14', 10', 7.1', 5', 5', 5', 5'**

==> DATA Released in 2013


COBE
WMAP
PLANCK
W-band temperature anisotropy
Internal Linear Combination of 5 bands, smoothed
Simulated temperature anisotropy
Simulated temperature and polarization anisotropy

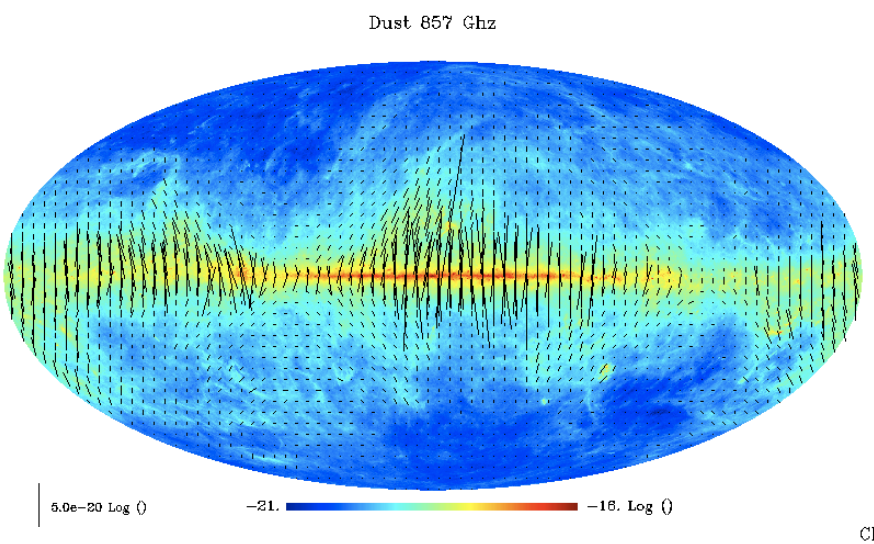
The CMB exhibits Fluctuations

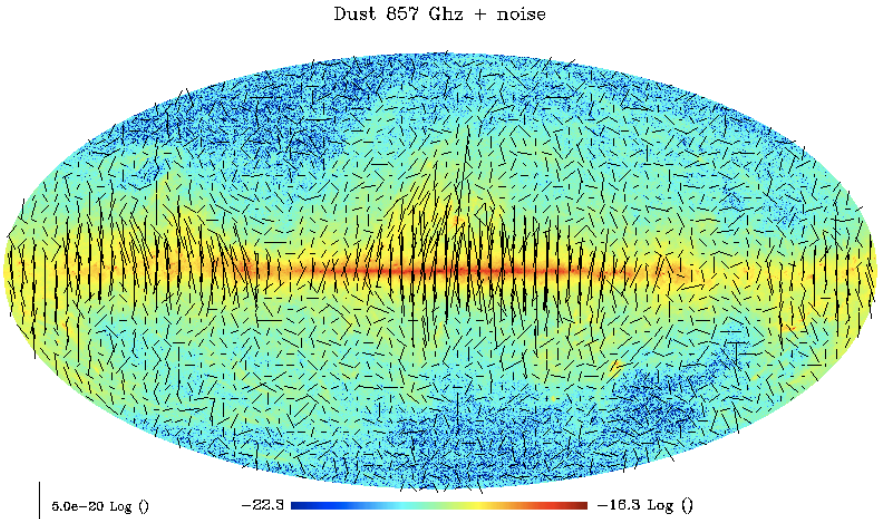


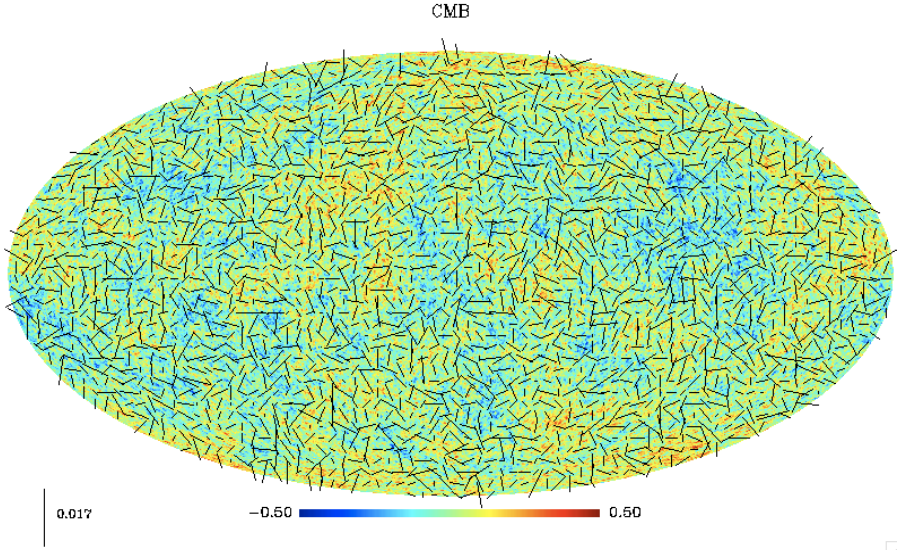
The Cosmic Microwave Background (CMB) is a relic radiation (with a temperature equals to 2.726 Kelvin) emitted 13 billion years ago when the Universe was about 370000 years old.

PLANCK SIMULATED POLARIZED DATA:

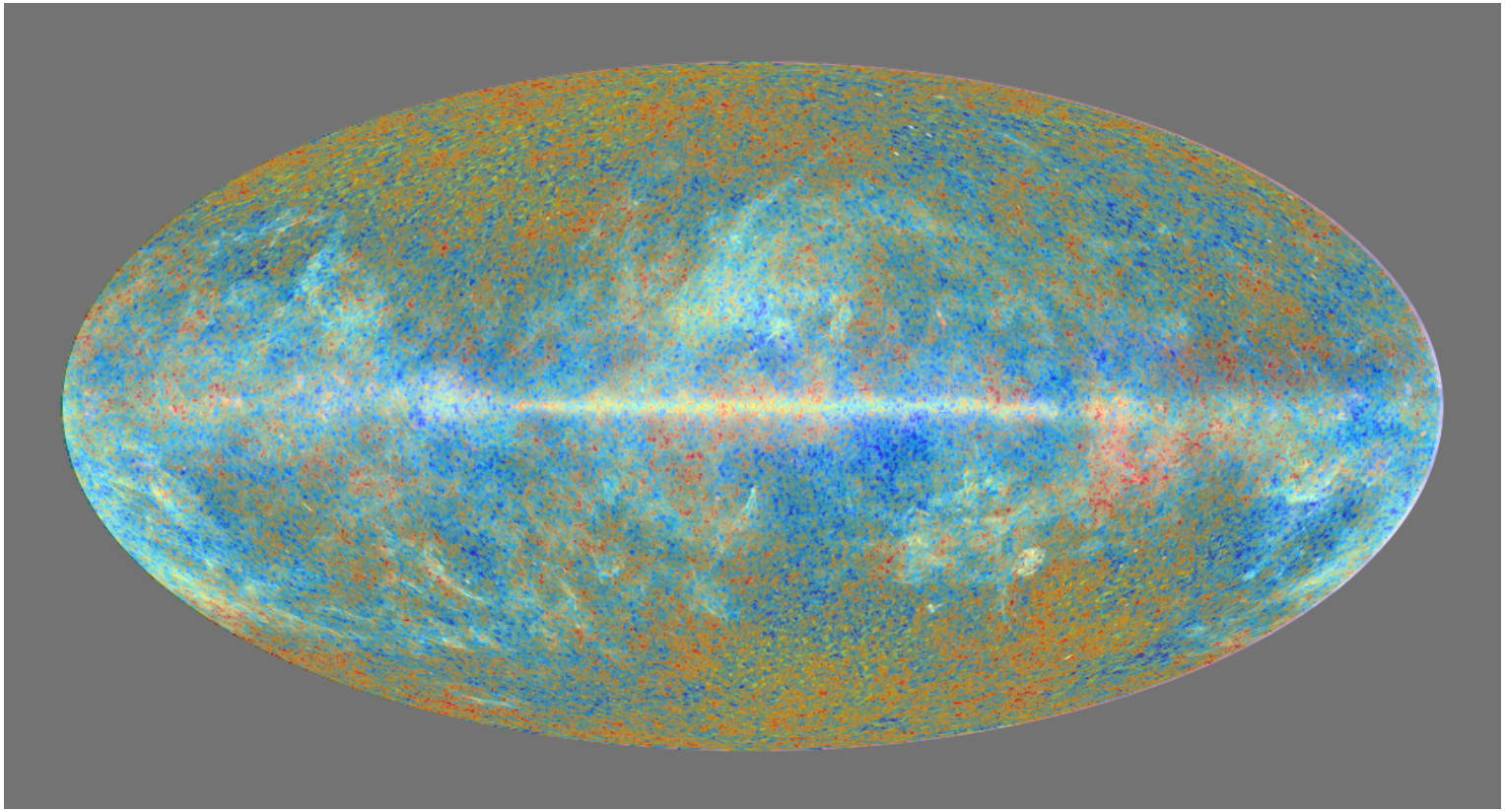
Magnitude $P = \sqrt{Q^2 + V^2}$ Orientation $\alpha = \arctan(U/Q)$







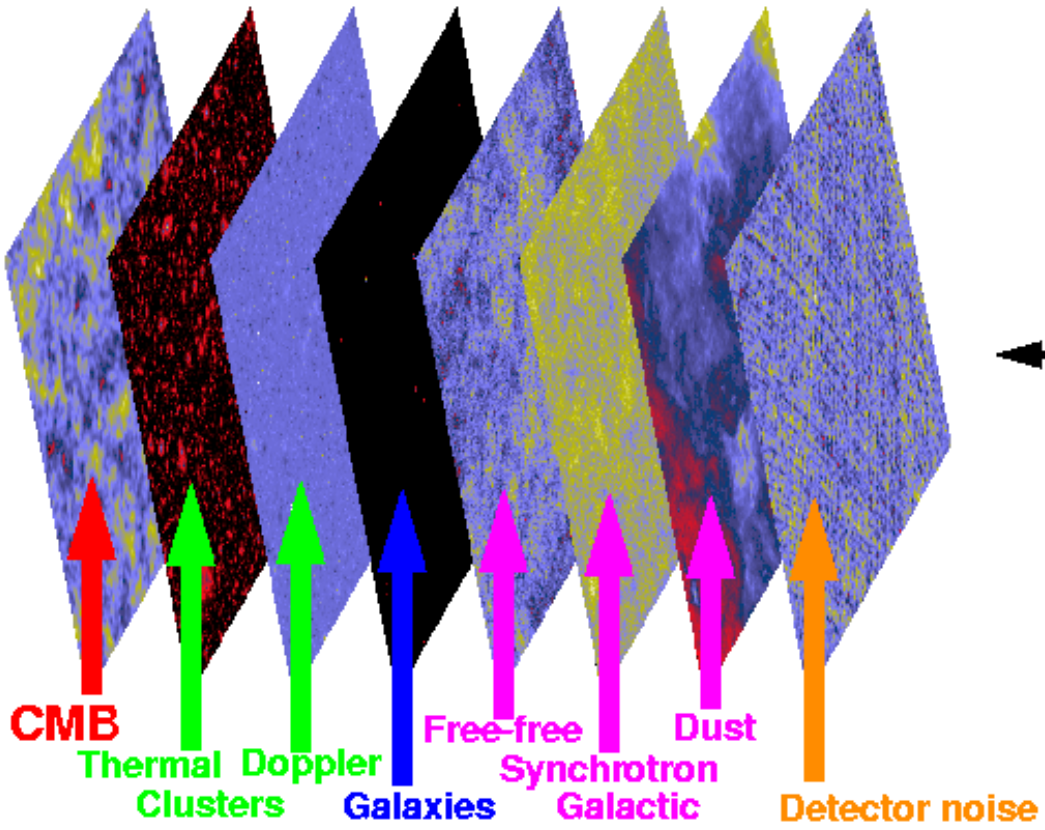
CMB simulated map

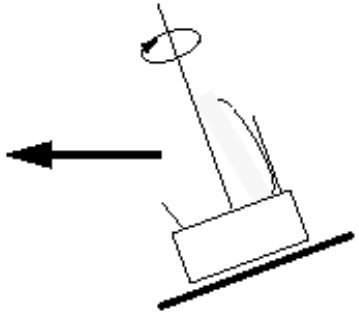


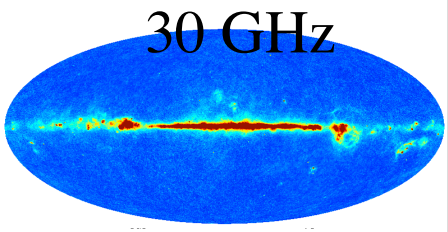
The importance of Source Separation

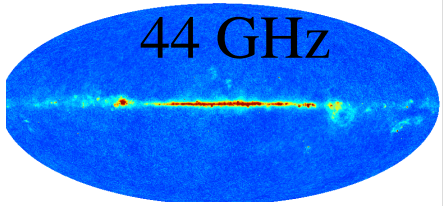
Extra foregrounds are superimposed with the CMB !!!

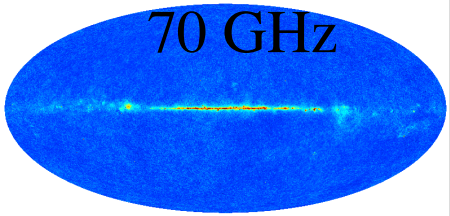
Point sources, galactic foregrounds, ... etc

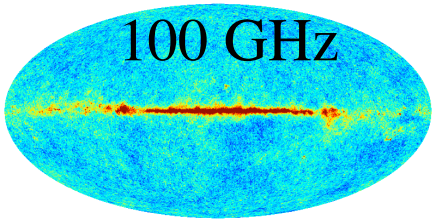

CMB
Thermal
Clusters
Doppler
Galaxies
Free-free
Synchrotron
Galactic
Dust
Detector noise

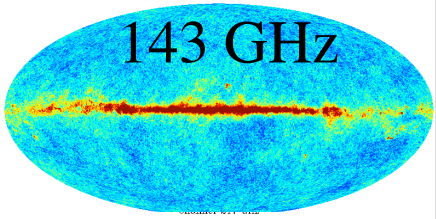


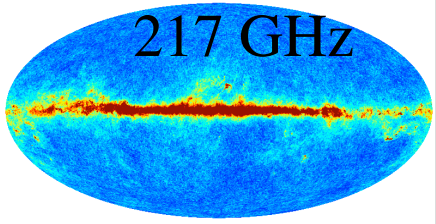

30 GHz

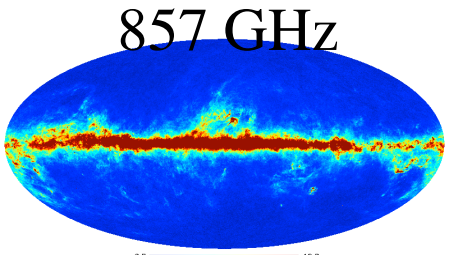

44 GHz

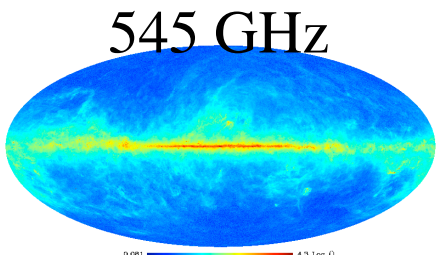

70 GHz

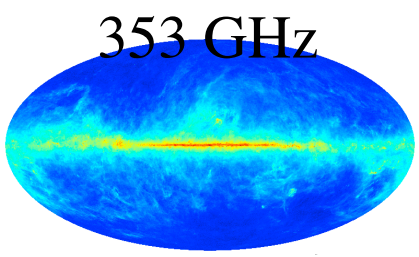

100 GHz

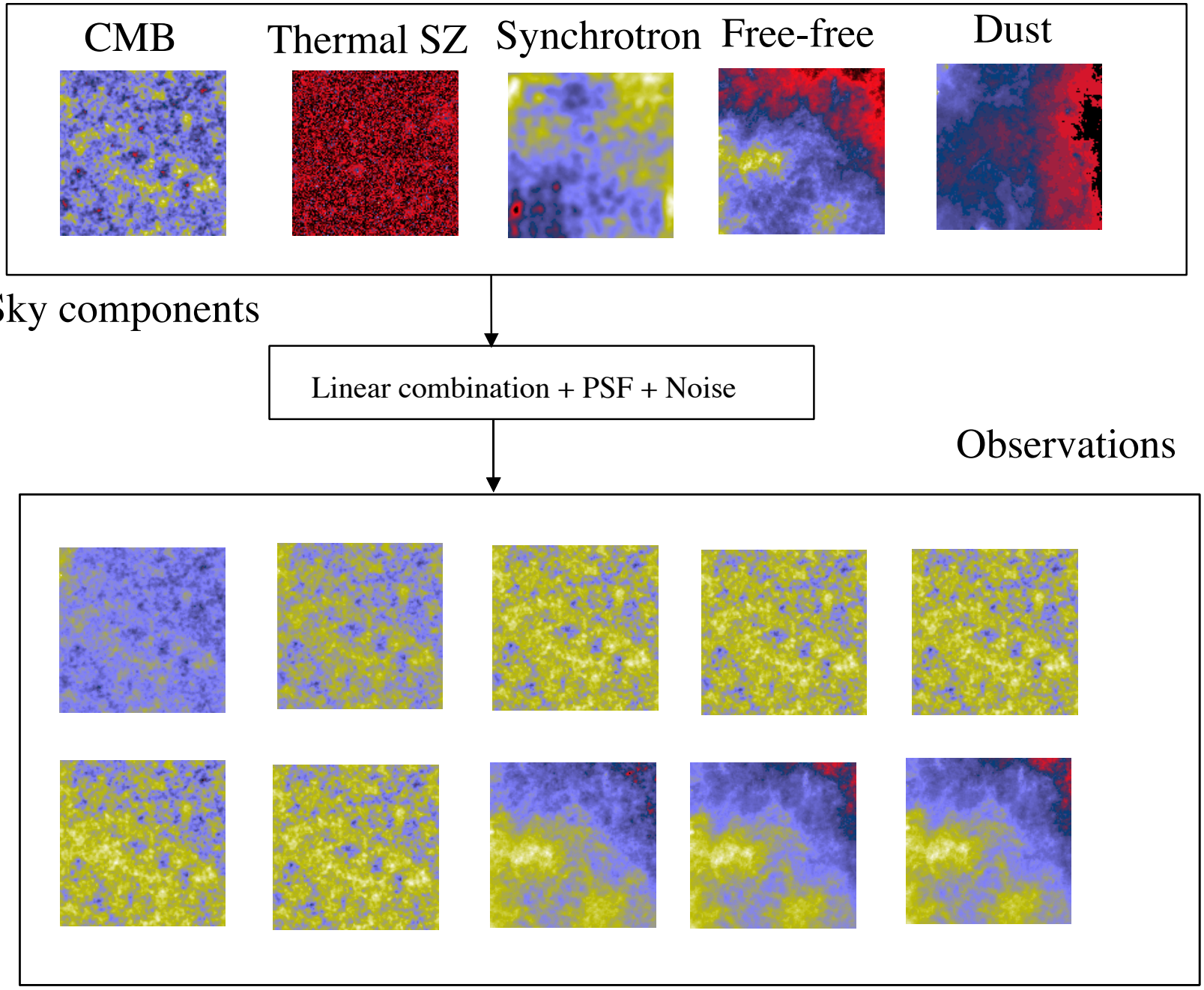

143 GHz


217 GHz


857 GHz


545 GHz


353 GHz


CMB
Thermal SZ
Synchrotron
Free-free
Dust
Sky components
Linear combination + PSF + Noise
Observations

METHODS

Template Fitting

Spectral Index Fitting

Independent Component Analysis: JADE, FastICA, …

$$\mathrm{Y} = \mathrm{A}\ \mathrm{X}$$

$$\mathrm{Y} = \mathrm{A}\ \mathrm{X} + \mathrm{N}$$

==> Another approach: Sparsity

What is a good sparse representation for data?

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

Dictionary
(basis, frame)

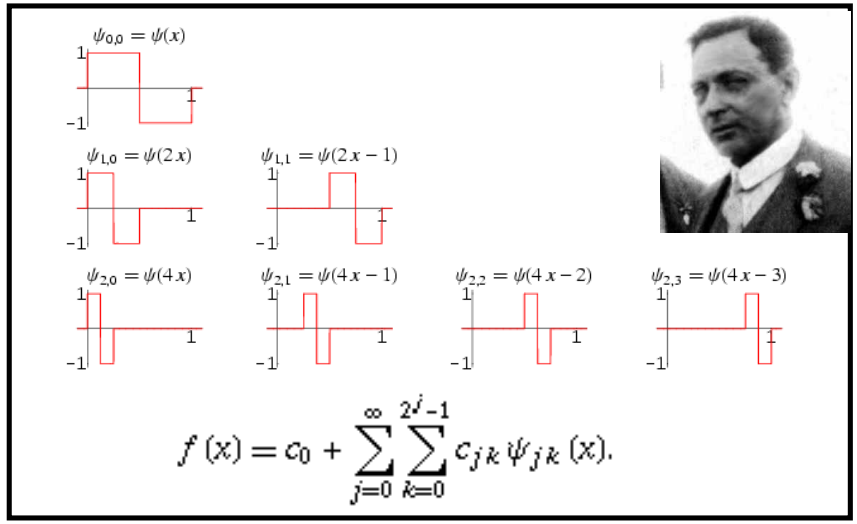
$$\Phi = \{\phi_1, \ldots, \phi_K\}$$

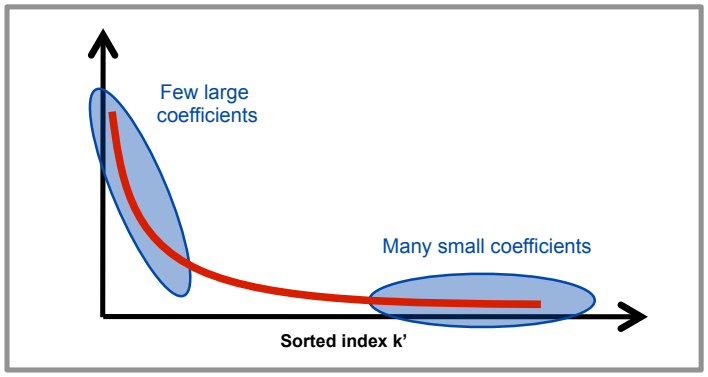
Atoms

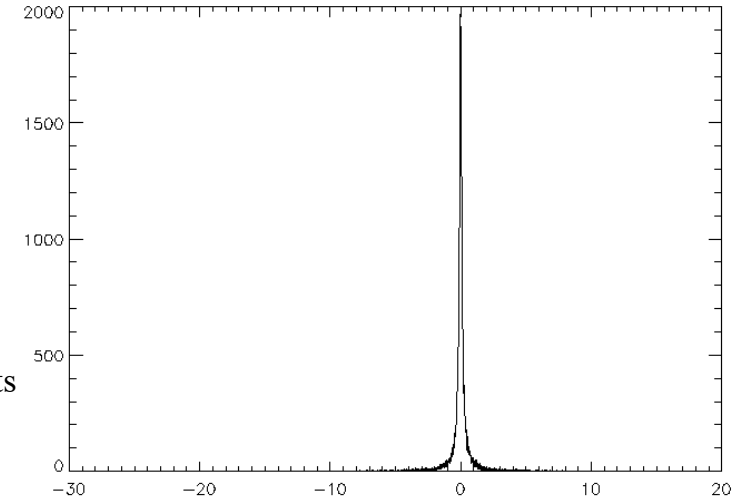
Ex: Haar wavelet

coefficients

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



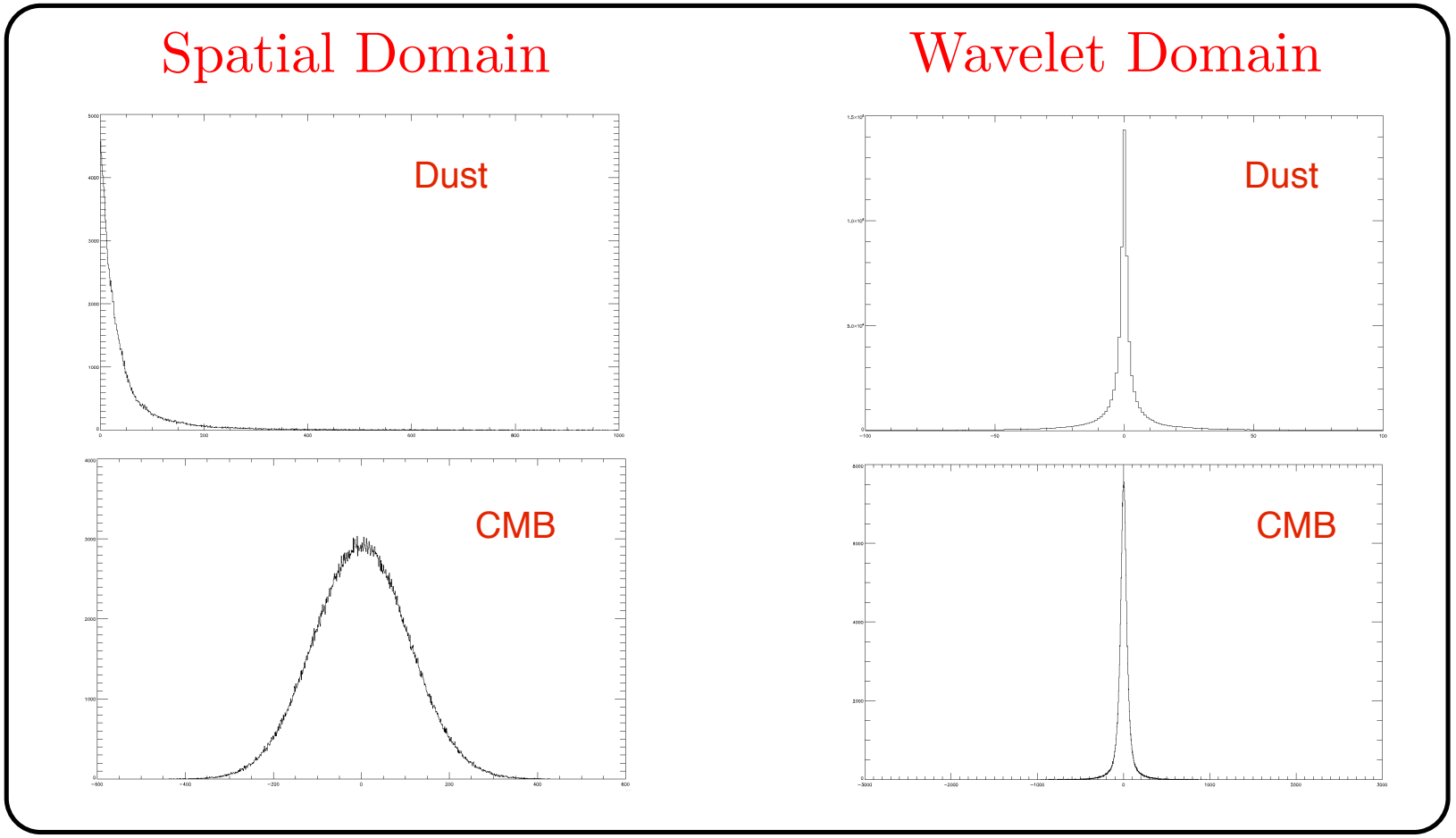




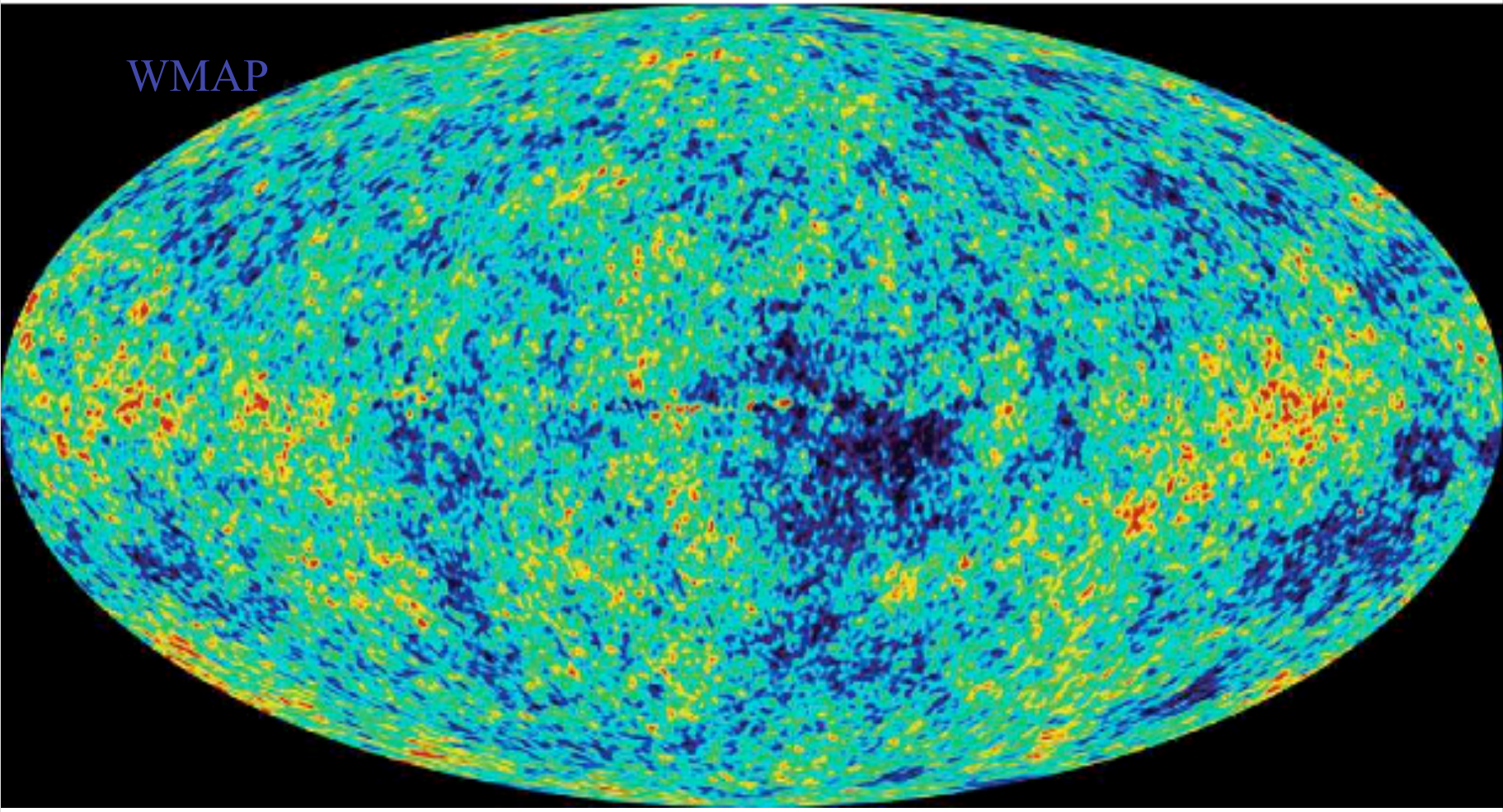
- Fast calculation of the coefficients
- Analyze the signal through the statistical properties of the coefficients
- Approximation theory uses the sparsity of the coefficients

Is Sparse models adapted to PLANCK Data ?

- Components are sparse in a wavelet dictionary



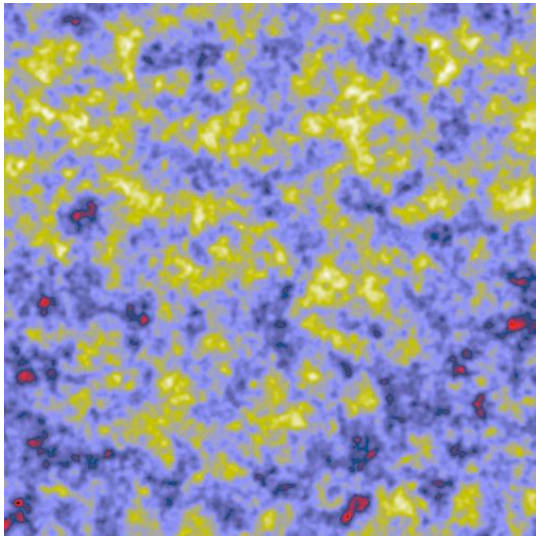
The CMB exhibits Fluctuations

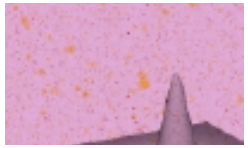


The Cosmic Microwave Background (CMB) is a relic radiation (with a temperature equals to 2.726 Kelvin) emitted 13 billion years ago when the Universe was about 370000 years old.

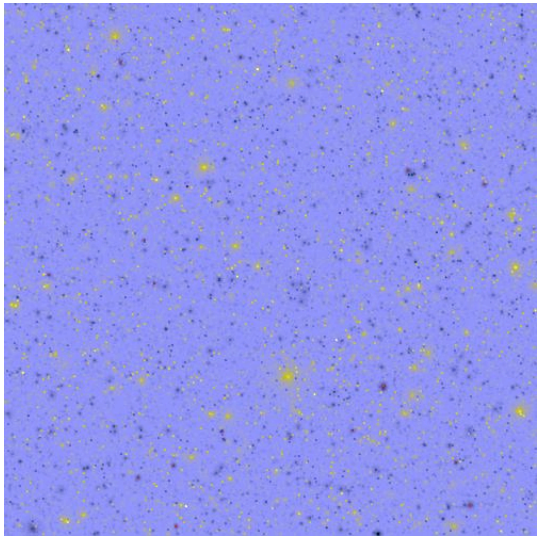
Cosmic Strings

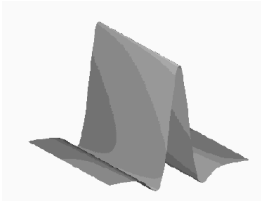
CMB



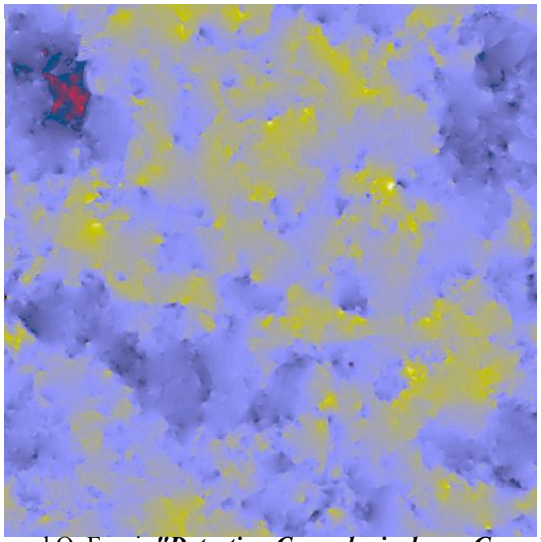


SZ

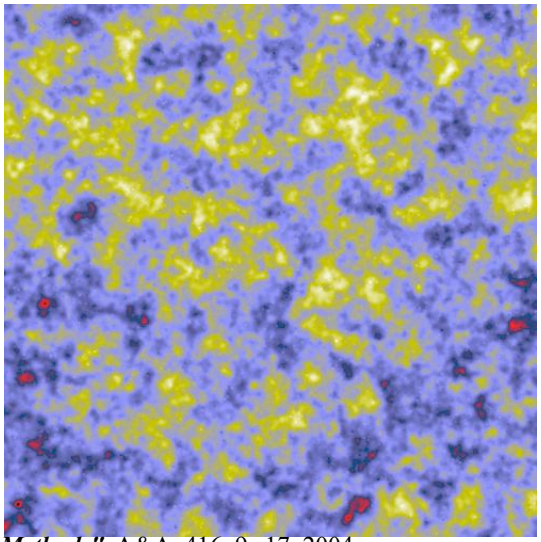




CS



Total



•J.-L. Starck, N. Aghanim and O. Forni, ***"Detecting Cosmological non-Gaussian Signatures by Multi-scale Methods"***, A&A, 416, 9--17, 2004.

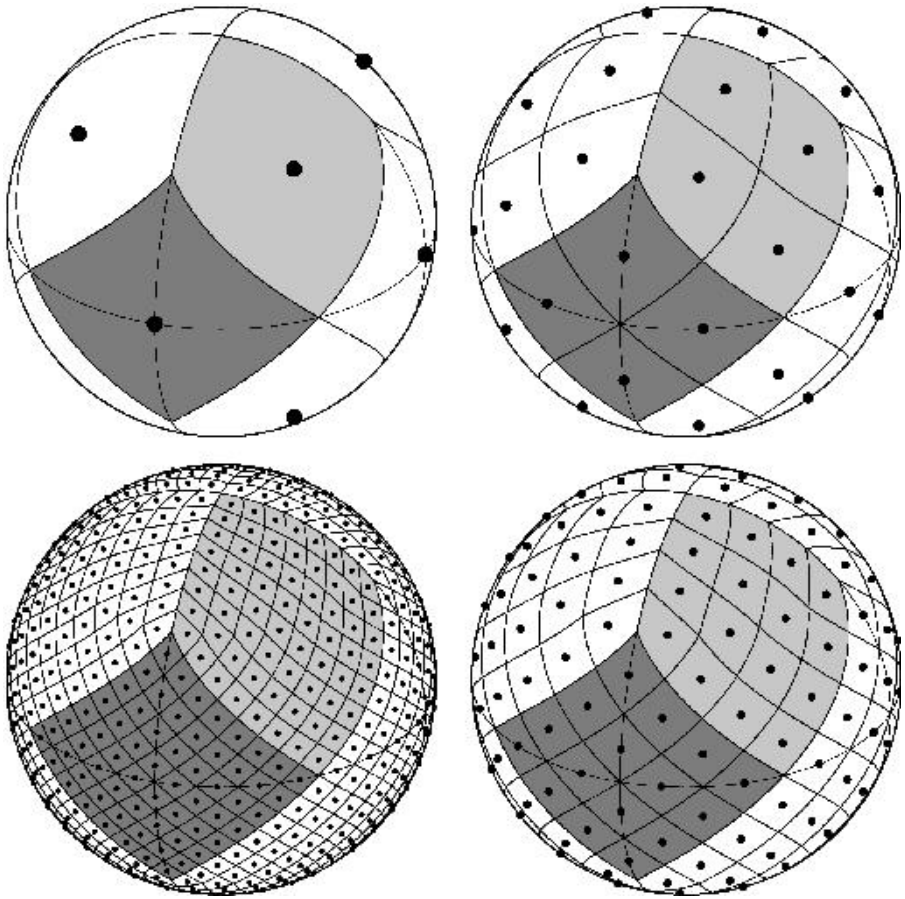
•J. Jin, J.-L. Starck, D.L. Donoho, N. Aghanim and O. Forni, ***"Cosmological Non-Gaussian Signatures Detection: Comparison of Statistical Tests"***, Eurasip Journal on Applied Signal Processing, 15 pp 2470-2485, 2005.

Healpix

K.M. Gorski et al., 1999, astro-ph/9812350,
http://www.eso.org/science/healpix

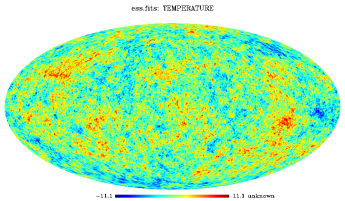
- Pixel = Rhombus
- Same Surfaces
- For a given latitude : regularly spaced
- Number of pixels:

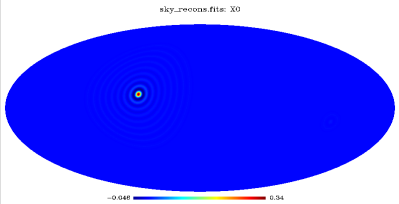
  $12 \times (N_{sides})^2$
- Included in the software:
  - Anafast
  - Synfast



Isotropic Undecimated Wavelet on the Sphere

Wavelets, Ridgelets and Curvelets on the Sphere, Astronomy & Astrophysics, 446, 1191-1204, 2006.





$$\hat{\psi}_{\frac{l_c}{2^j}}(l,m) = \hat{\phi}_{\frac{l_c}{2^{j-1}}}(l,m) - \hat{\phi}_{\frac{l_c}{2^j}}(l,m)$$

$$\hat{H}_j(l,m) = \begin{cases} \dfrac{\hat{\phi}_{\frac{l_c}{2^{j+1}}}(l,m)}{\hat{\phi}_{\frac{l_c}{2^j}}(l,m)} & \text{if } l < \frac{l_c}{2^{j+1}} \quad \text{and} \quad m = 0 \\ 0 & \text{otherwise} \end{cases}$$

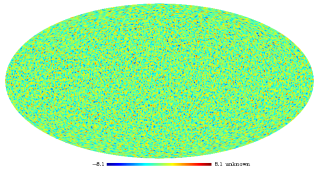
$$\hat{G}_j(l,m) = \begin{cases} \dfrac{\hat{\psi}_{\frac{l_c}{2^{j+1}}}(l,m)}{\hat{\phi}_{\frac{l_c}{2^j}}(l,m)} & \text{if } l < \frac{l_c}{2^{j+1}} \quad \text{and} \quad m = 0 \\ 1 & \text{if } l \geq \frac{l_c}{2^{j+1}} \quad \text{and} \quad m = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\hat{c}_{j+1}(l,m) = \hat{H}_j(l,m)\hat{c}_j(l,m)$$
$$\hat{w}_{j+1}(l,m) = \hat{G}_j(l,m)\hat{c}_j(l,m)$$

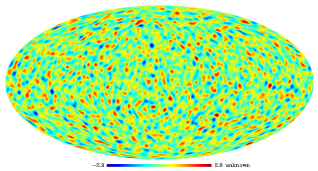
$$c_0(\vartheta,\varphi) = c_J(\vartheta,\varphi) + \sum_{j=1}^{J} w_j(\vartheta,\varphi)$$

Undecimated Wavelet Transform

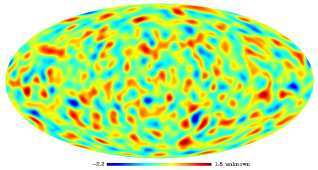
j=1



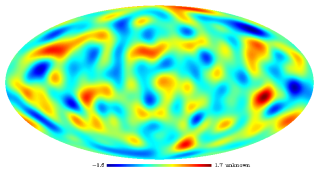
j=2

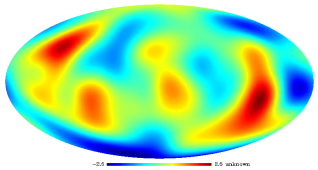


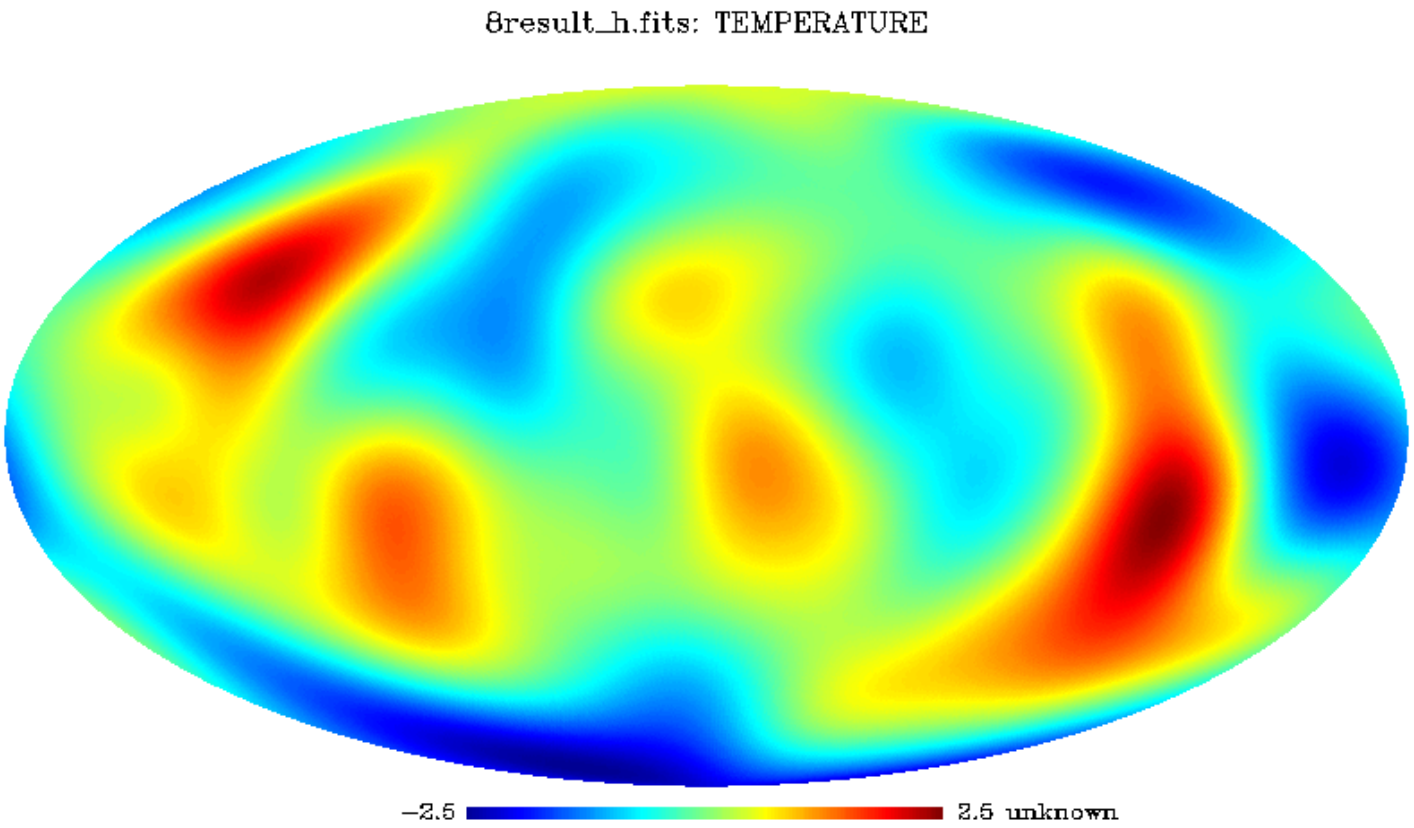
j=3



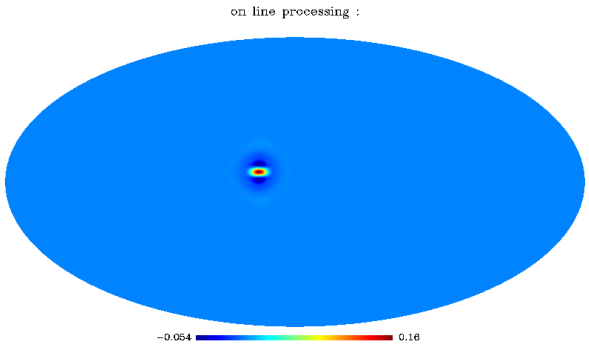
j=4

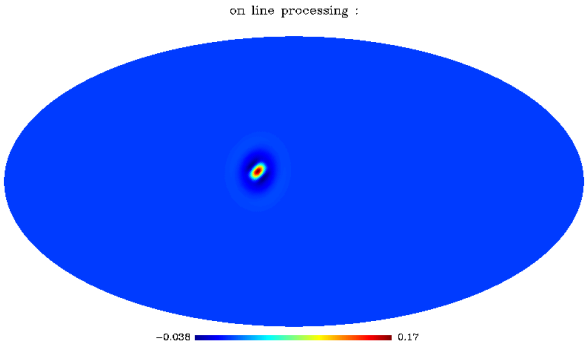


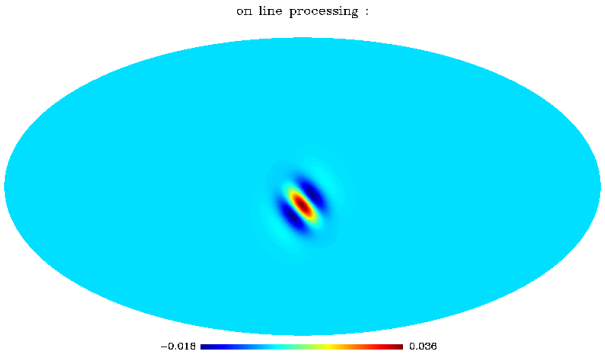


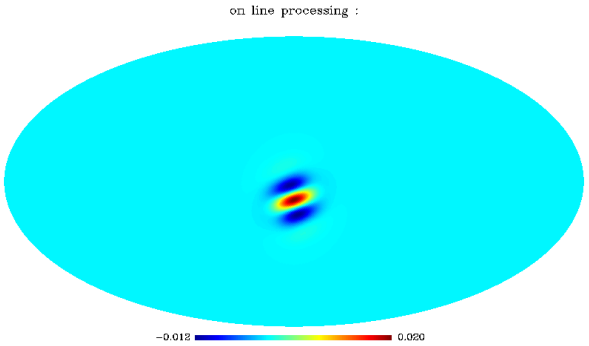

8result_h.fits: TEMPERATURE
−2.5
2.5 unknown

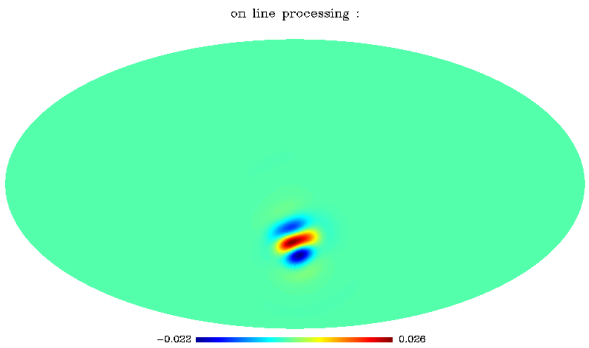
Curvelet functions on the sphere

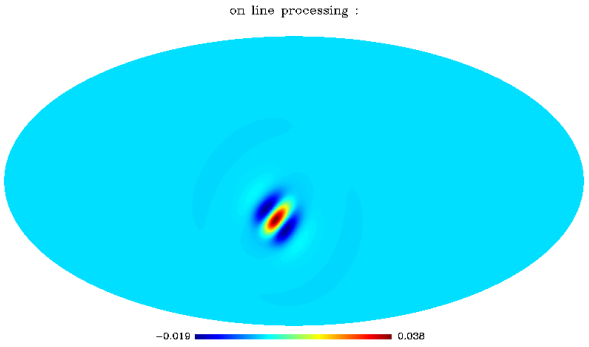






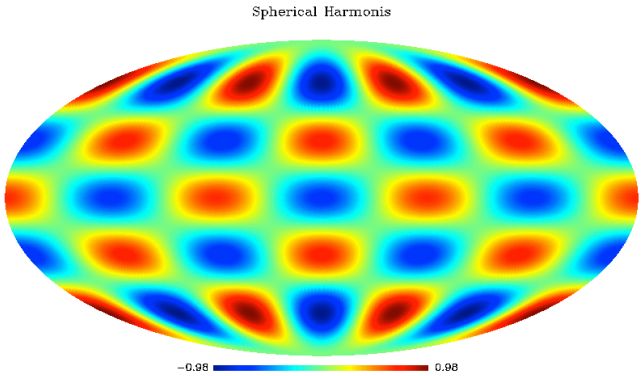




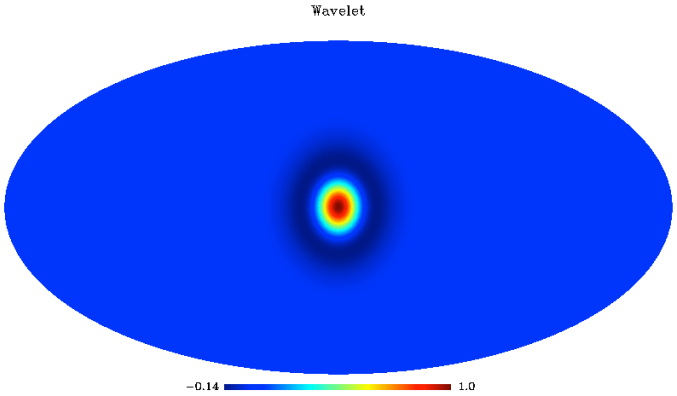


Dictionaries

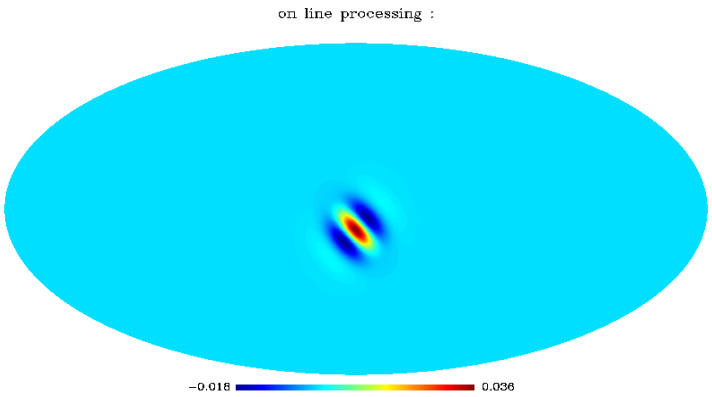
Spherical Harmonics



Wavelets

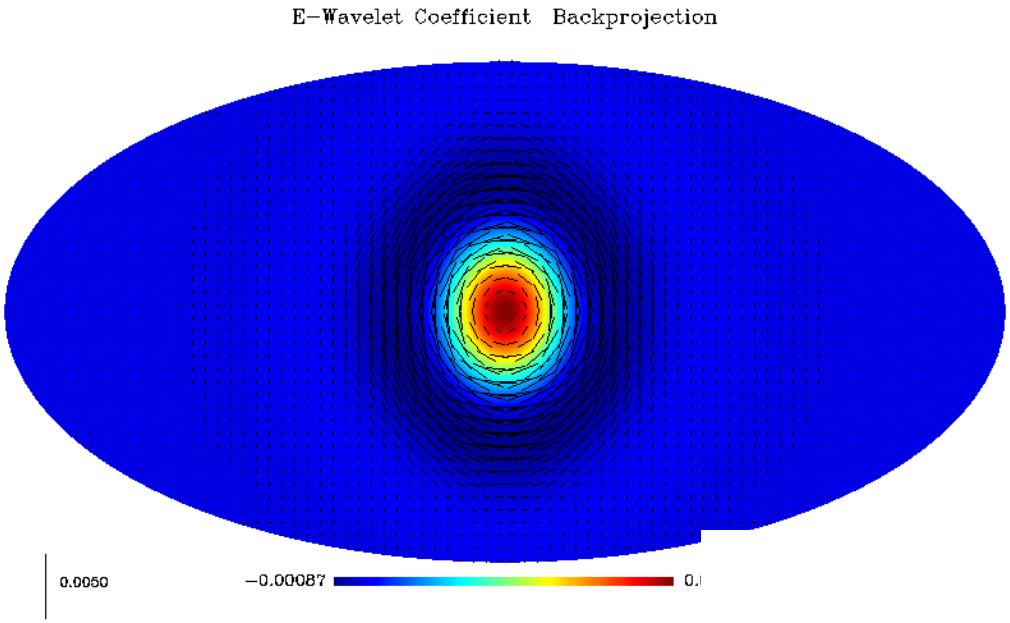


Curvelets

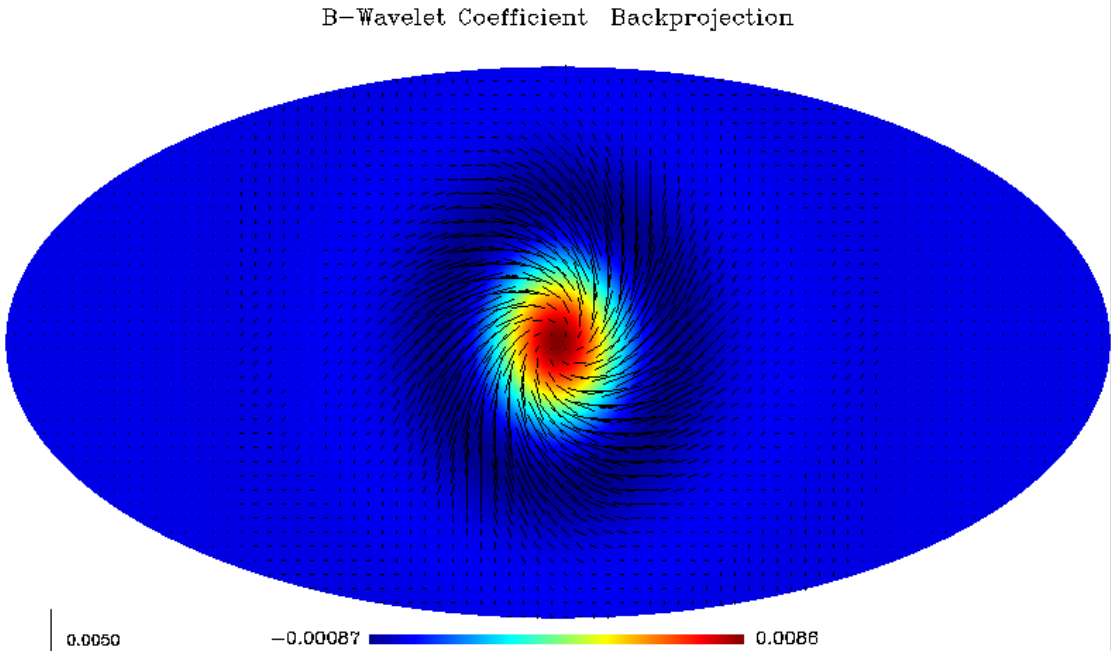


Polarized Dictionary: E/B Polarized Wavelet

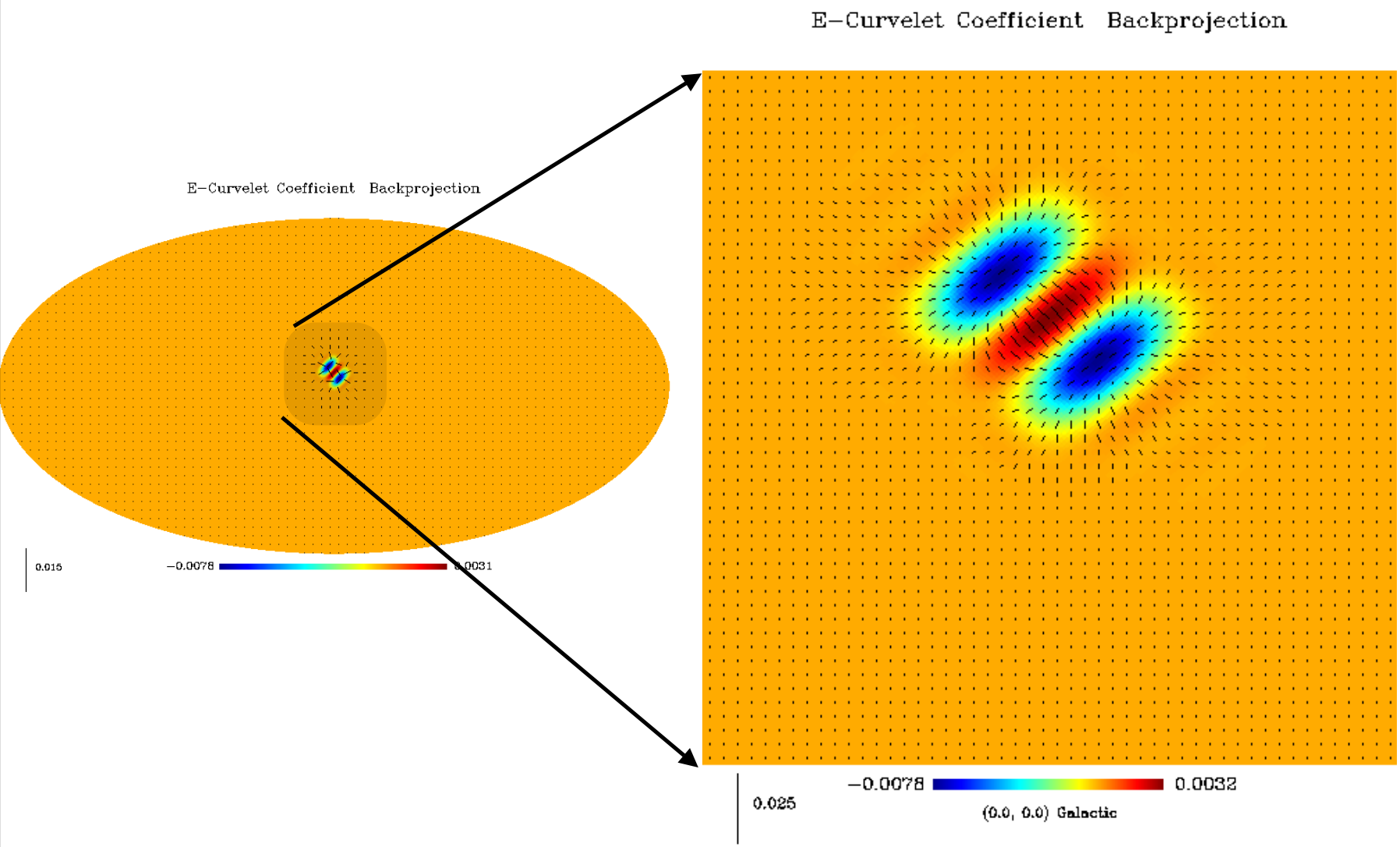
J.-L. Starck, Y. Moudden and J. Bobin, "Polarized Wavelets and Curvelets on the Sphere", Astronomy and Astrophysics, 497, 3, pp 931--943, 2009.



j=4



Curvelet and E/B Mode Decomposition



Polarized Data Denoising

$$Q(\theta,\phi) = \sum_{l,m} c^E_{J,l,m} Z^+_{l,m} + i c^B_{J,l,m} Z^-_{l,m} + \sum_j \sum_{l,m} \tilde{w}^E_{j,l,m} Z^+_{l,m} + i \tilde{w}^B_{j,l,m} Z^-_{l,m}$$

$$U(\theta,\phi) = \sum_{l,m} c^B_{J,l,m} Z^+_{l,m} - i c^E_{J,l,m} Z^-_{l,m} + \sum_j \sum_{l,m} \tilde{w}^B_{j,l,m} Z^+_{l,m} - i \tilde{w}^E_{j,l,m} Z^-_{l,m}$$

Where

$$\tilde{w}^E_{j,k} = \delta(w^E_{j,k})$$

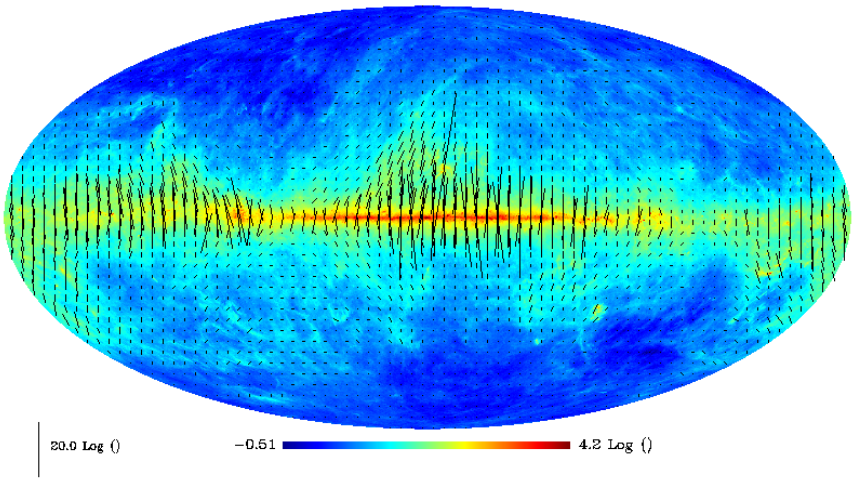
$$\tilde{w}^B_{j,k} = \delta(w^B_{j,k})$$

Hard thresholding corresponds to the following non linear operation:

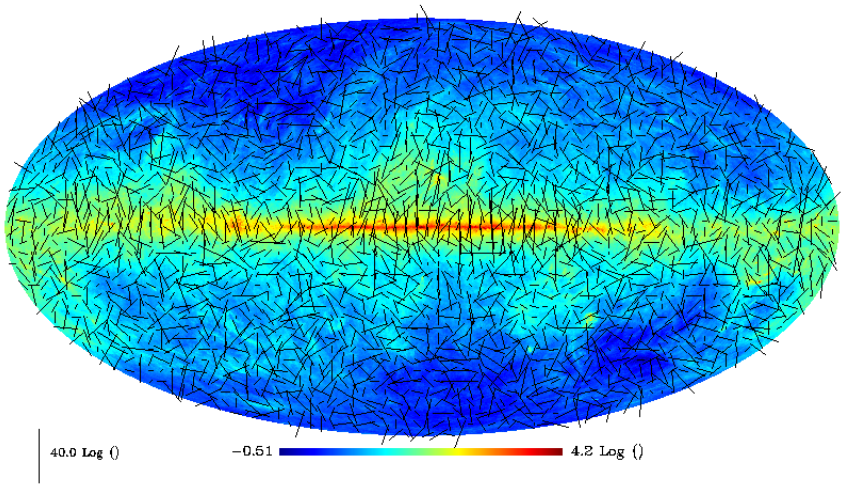
$$\tilde{w}_{j,k} = \begin{cases} w_{j,k} & \text{if } \mid w_{j,k} \mid \geq T_j \\ 0 & \text{otherwise} \end{cases}$$

Polarized Data Denoising

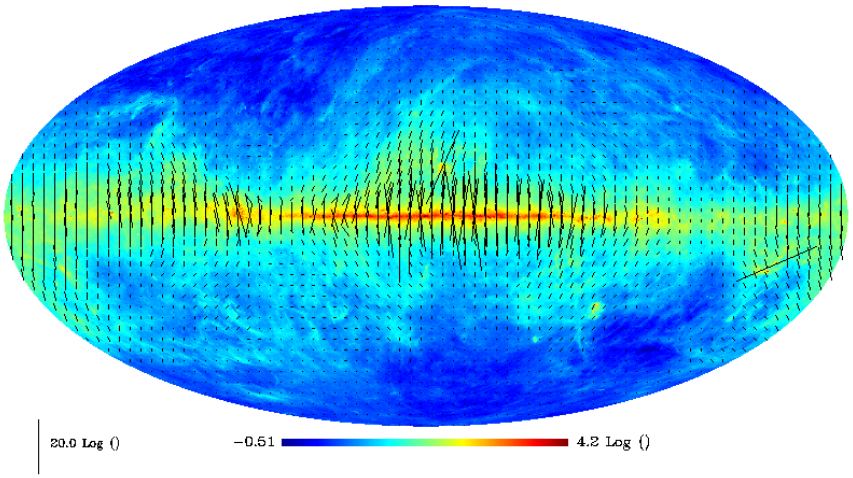
Dust 857 Ghz



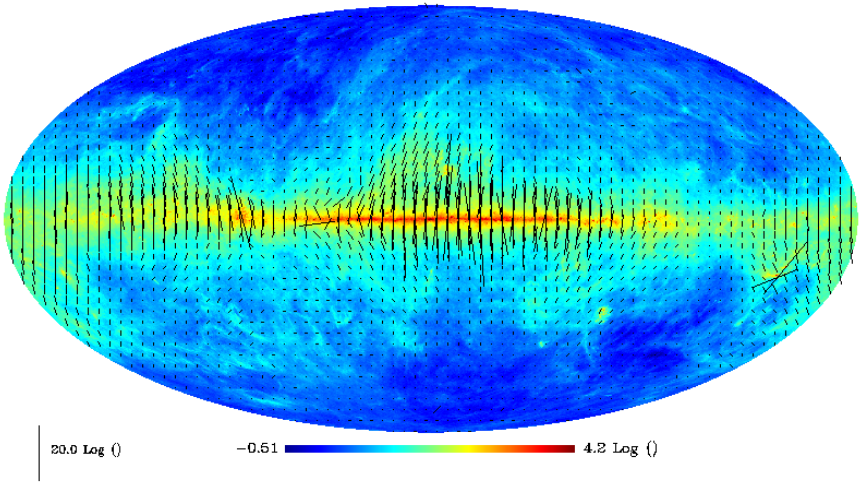
Dust 857 Ghz + noise



E−B Wavelet filtering



E−B Curvelet Denoising



Morpho-Spectral Diversity

Data: $X = [x_1, \ldots, x_m]$

Source: $S = [s_1, \ldots, s_n]$

$$X = [x_1, \ldots, x_m] = AS$$

$$x_l = \sum_{i=1}^{n} a_{i,l} s_i$$

$$\min_{\alpha} \|\alpha\|_p \text{ s.t } \mathbf{X} = \sum_{\gamma \in \Gamma} \alpha_\gamma \psi_\gamma$$

$\mathbf{\Phi_A} = [\mathbf{\Phi_{A,1}}, \mathbf{\Phi_{A,2}}]$ Spatial Dictionary with

$\mathbf{\Phi_S}$ Spectral Dictionary

$$\mathbf{\Psi} = [\mathbf{\Phi_{A,1}} \otimes \mathbf{\Phi_S}, \mathbf{\Phi_{A,2}} \otimes \mathbf{\Phi_S}]$$

Sparse Component Separation: the GMCA Method

•J. Bobin, J.-L. Starck, M.J. Fadili, and Y. Moudden, "Sparsity, Morphological Diversity and Blind Source Separation", IEEE Trans. on 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.

•P. Paykari and J.-L. Starck, "Cosmic Microwave Background Data Analysis",, in Advances in Machine Learning and Data Mining for Astronomy, Chapman and Hall, Edtitors M. Way, J. Scargle, K. Ali, A. Srivastava, in press, 2011.

Source: $S = [s_1, s_n]$

We define a dictionary ϕ

Data: $X = [x_1, ..., x_n] = AS + N$

==> GMCA searches a sparse solution S in the dictionary ϕ subject to the constraint that the norm $\| X - AS \|^2$ is minimal.

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

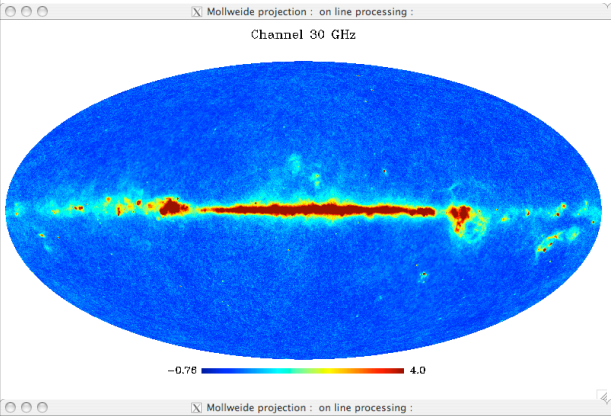
1) Estimate S assuming A is fixed (iterative thresholding) :

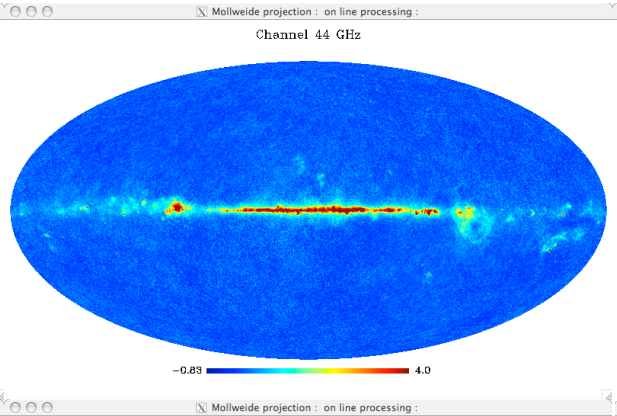
$$\{S\} = \mathrm{Argmin}_S \sum_j \lambda_j \|s_j \mathbf{W}\|_1 + \|\mathbf{X} - \mathbf{AS}\|^2_{F,\mathbf{\Sigma}}$$

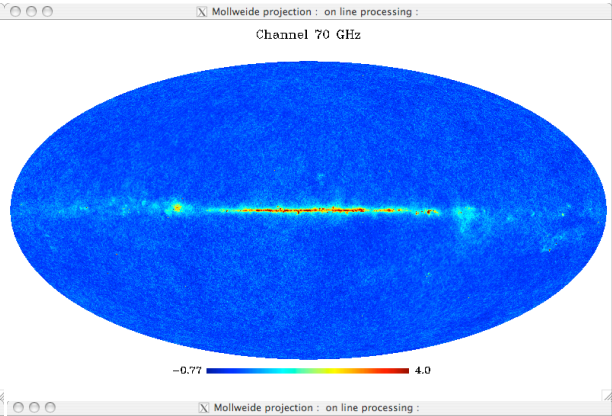
2) Estimate A assuming S is fixed (a simple least square problem) :

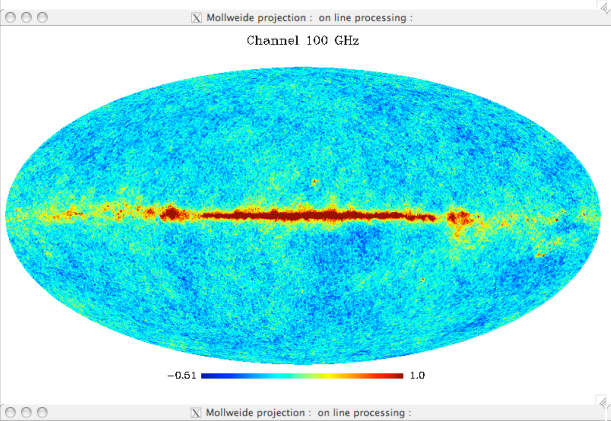
$$\{A\} = \mathrm{Argmin}_A \|\mathbf{X} - \mathbf{AS}\|^2_{F,\mathbf{\Sigma}}$$

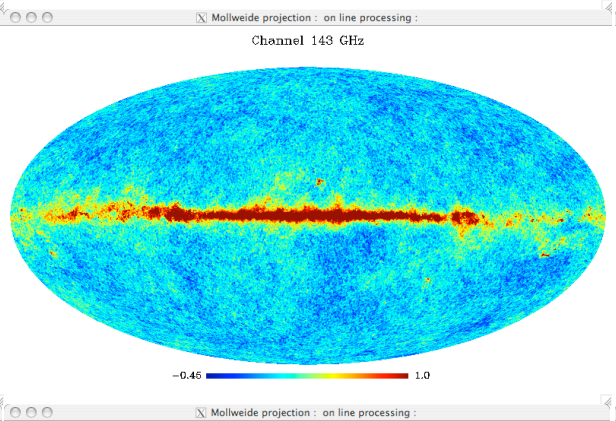
PLANCK Simulated Data

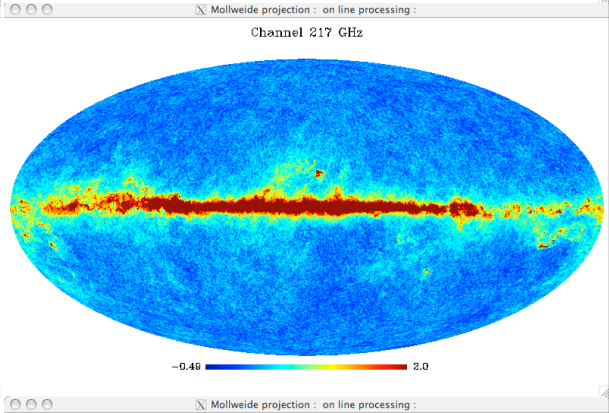


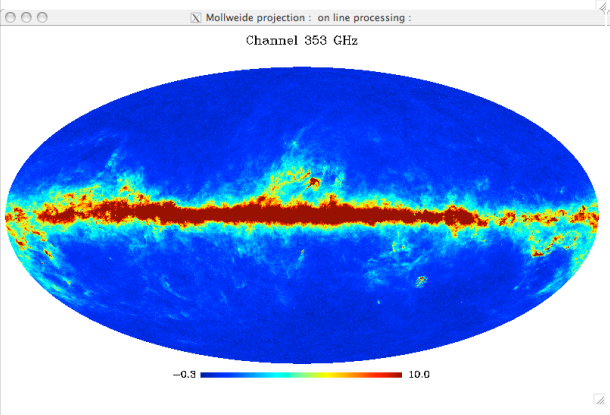


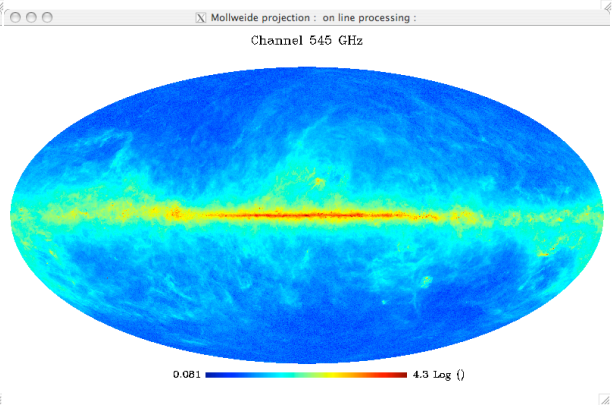


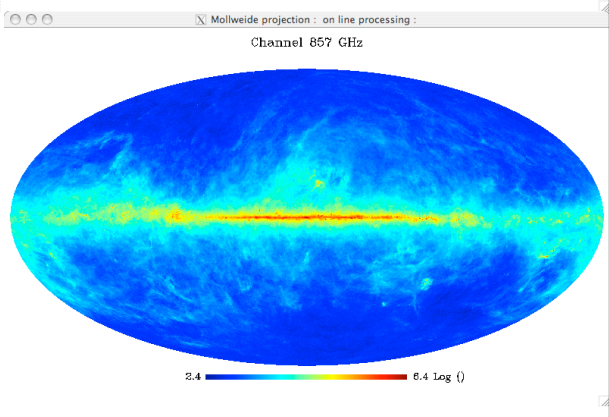




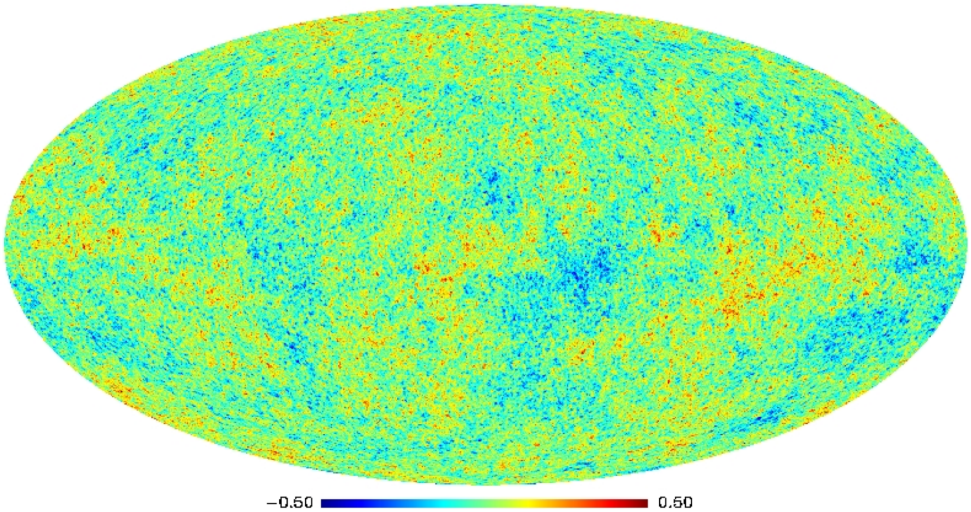




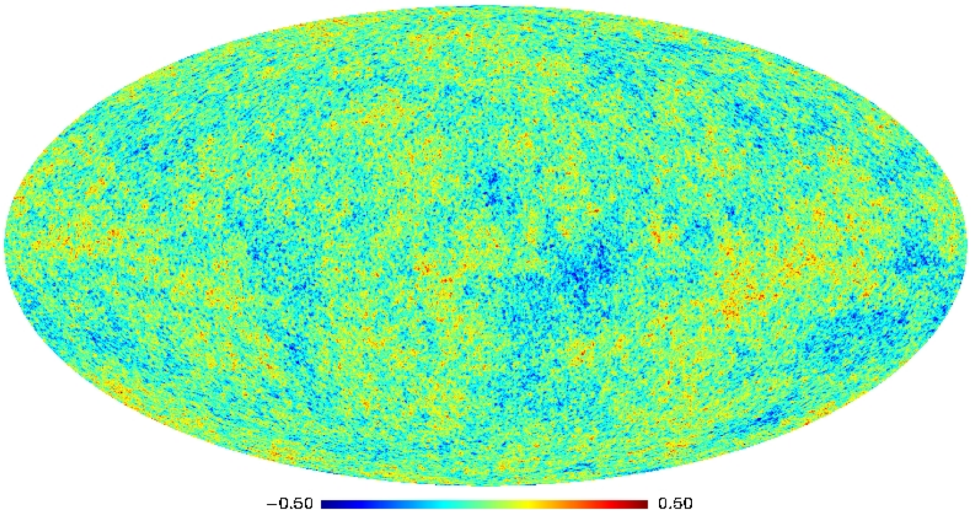




Planck - WG2 - Challenge 2



Input simulated CMB map (mK)



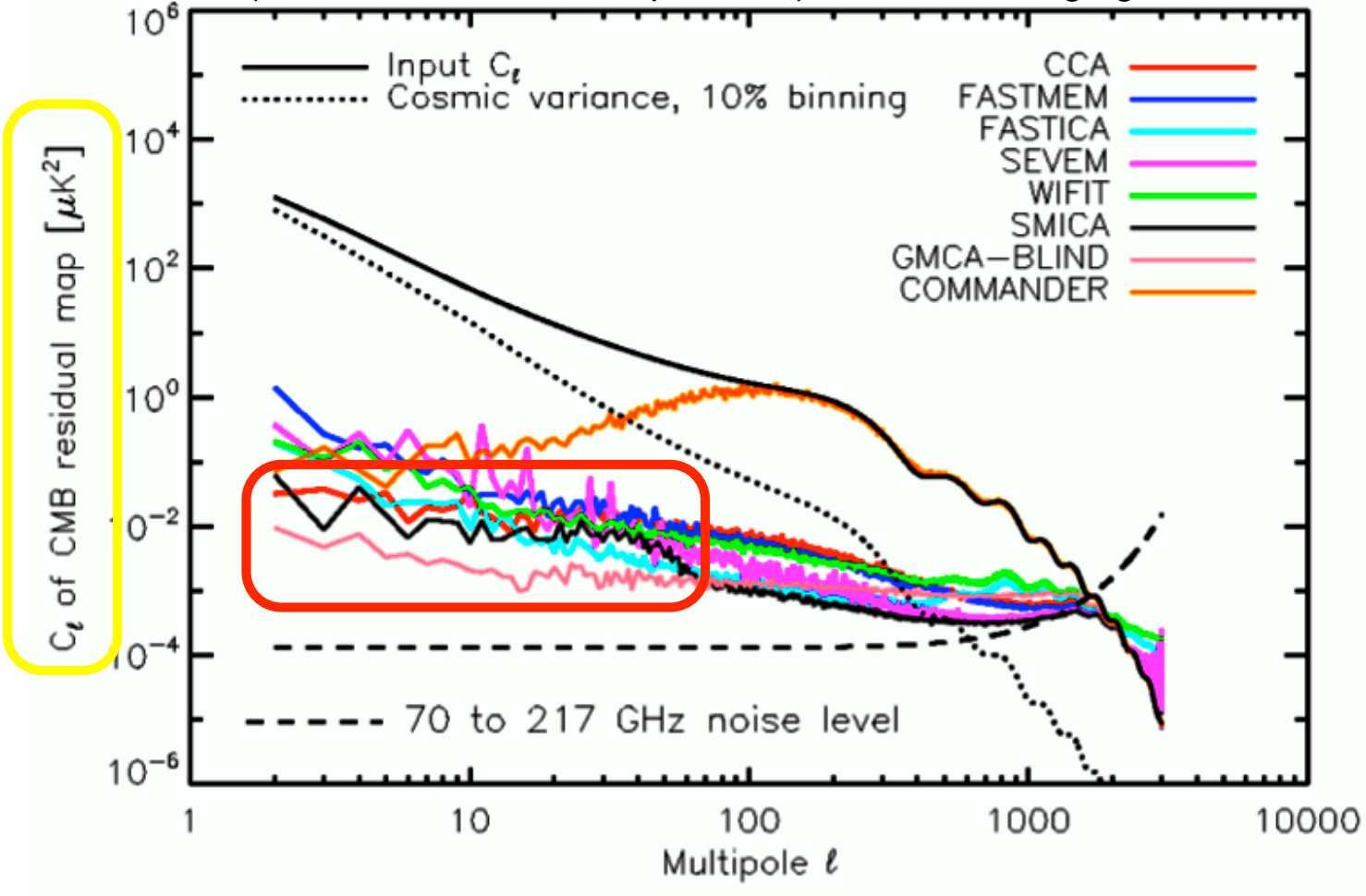
Estimated map with GMCA

Planck - WG2 - Challenger 2

comparaisons - power spectrum

GMCA has the lowest spectral residuals at low l.

Plot shows C_l of (reconstructed CMB – input CMB) evaluated at high galactic latitudes.



- **Sam Leach** (SISSA), June 19, 2008, WG2 meeting, Munich

Limitations

GMCA Model: $Y = A X + N$

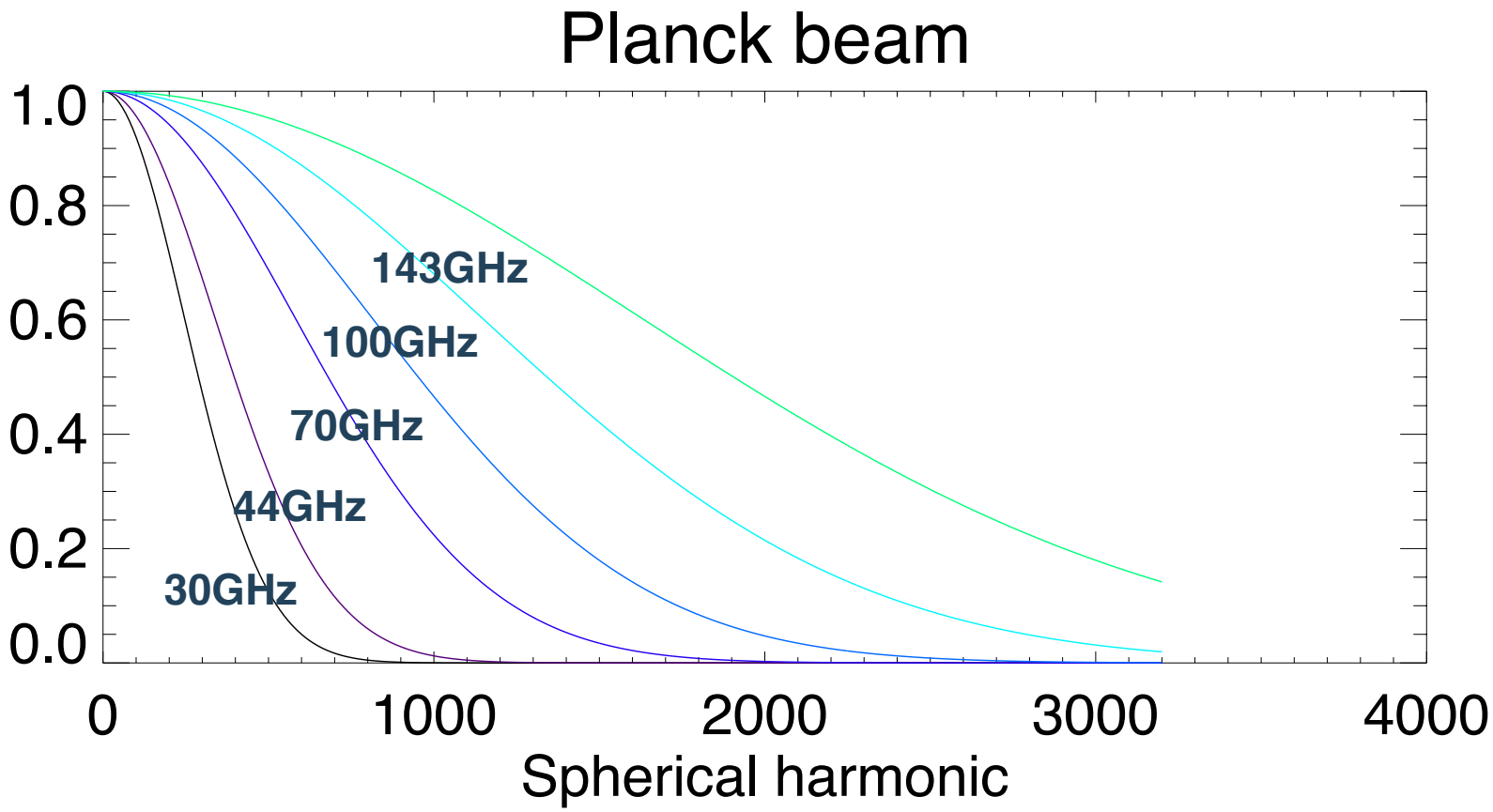
But three main problems:
i) A is spatially variant.
ii) This model does not take into account the beam.
iii) Noise is not homogeneous.

- Limitation of GMCA:

- One matrix to describe the whole sky (i.e. the simplest model !)
- PSF were not taken into account properly
- Non stationary noise.

Component Separation

Planck provides data that do not share the same resolution :



Component Separation: more problems

More formally:

$$\forall i;\ x_i = b_i \star \left(\sum_j a_{ij} s_j \right) + n_i$$

Globally: $\mathbf{X} = \mathcal{H}(\mathbf{AS}) + \mathbf{N}$

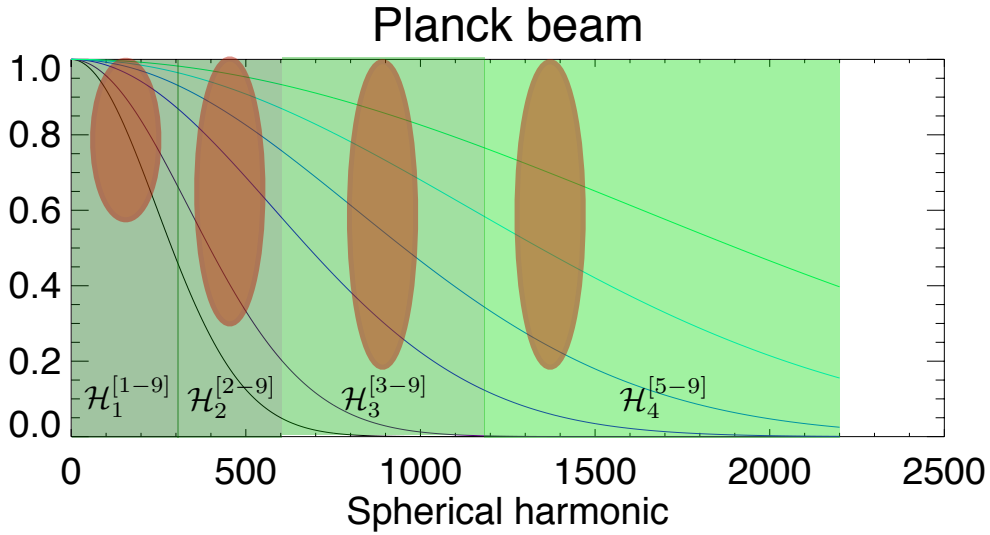
where $\mathcal{H}$ is the multichannel convolution operator

The mixture model no more holds ! **$\mathcal{H}$ is singular !**

Spectral behavior varies spatially for some components (dust, synchroton):

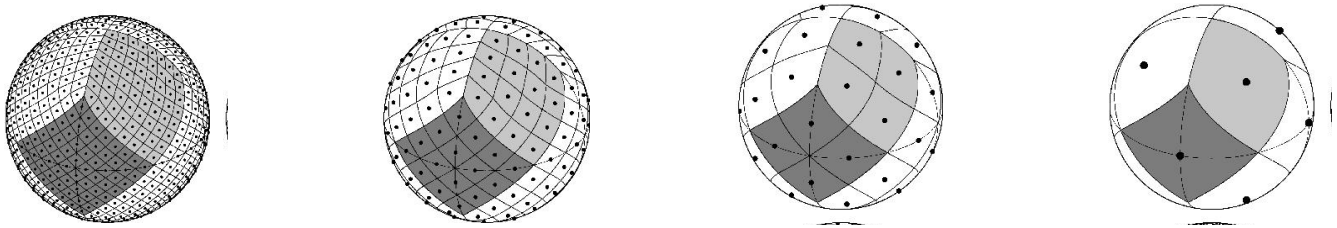
Component Separation: the beam problem

=> Use Wavelets to work at different resolutions:



=> Assume the mixing matrix varies smoothly

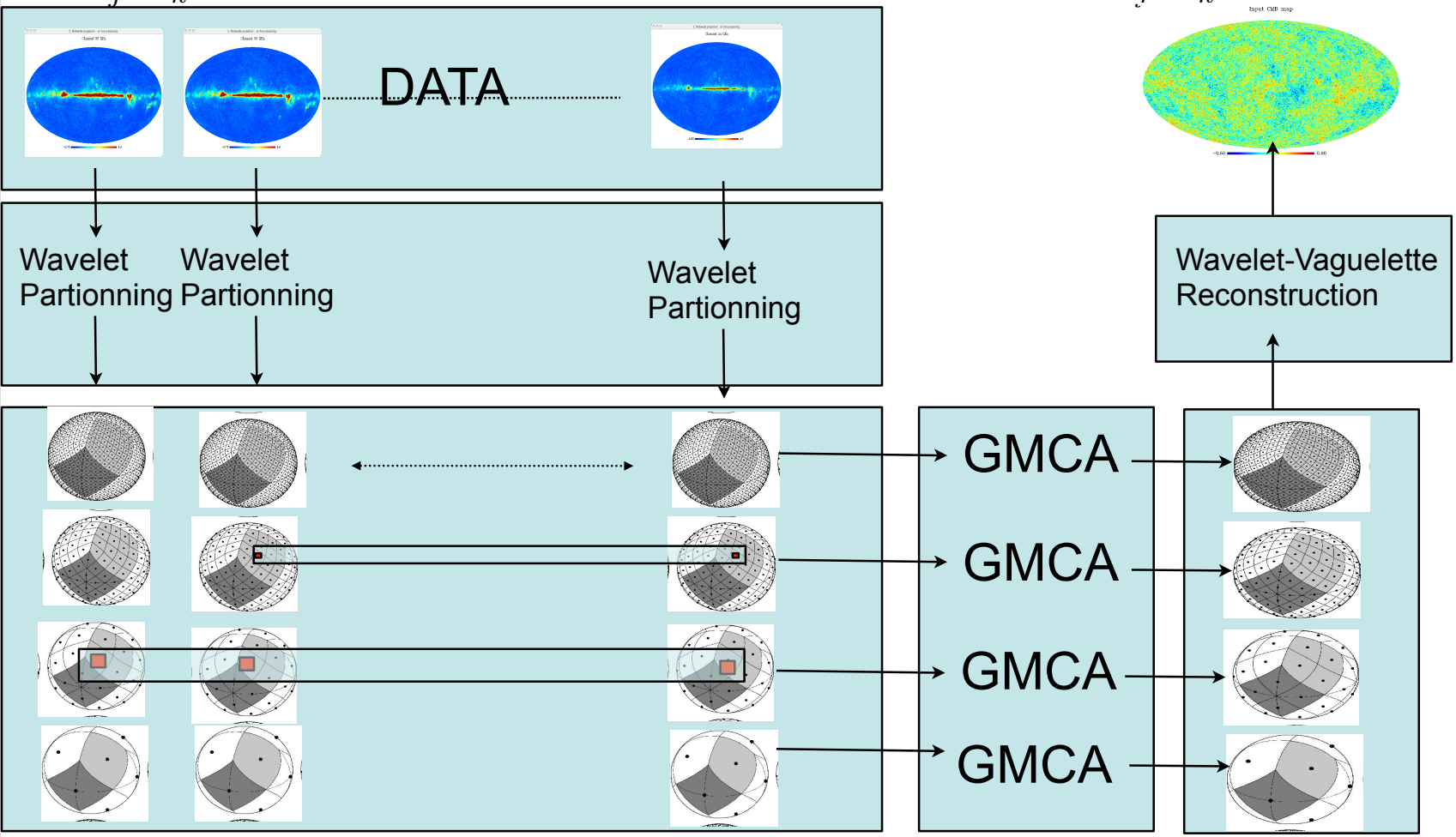
Partitionning of the Wavelet Scales



Wavelet-Vaguelette GMCA Decomposition

$$f = \sum_j \sum_k \langle Kf, \Psi_{j,k} \rangle \psi_j, k \quad \text{with } K^* \Psi_{j,k} = \psi_{j,k}$$

$$\tilde{f} = \sum_j \sum_k \Delta(\langle y, \Psi_{j,k} \rangle) \psi_j, k$$



Wavelet-Vaguelette GMCA Decomposition

Inverse problem

$$y = Kf + n \xrightarrow{\text{WVD}} f = \sum_j \sum_k \langle Kf, \Psi_{j,k} \rangle \psi_j, k \qquad \text{with } K^* \Psi_{j,k} = \psi_{j,k}$$

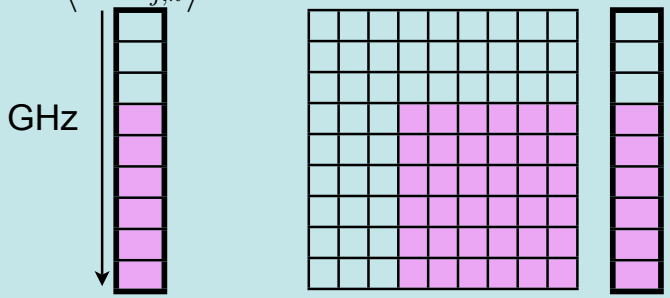
$$\tilde{f} = \sum_j \sum_k \Delta(\langle y, \Psi_{j,k} \rangle) \psi_j, k$$

Multi-channel WVD

$$Y_i = \sum_j K_i (A_j X)_i + N_i \longrightarrow \tilde{X}_s = \left[\sum_j \sum_k \tilde{A}_j^+ \left\langle Y_i, \Psi_{j,k}^{(i)} \right\rangle \psi_{j,k} \right]_s$$

with $K_i^* \Psi_{j,k}^{(i)} = \psi_{j,k}$

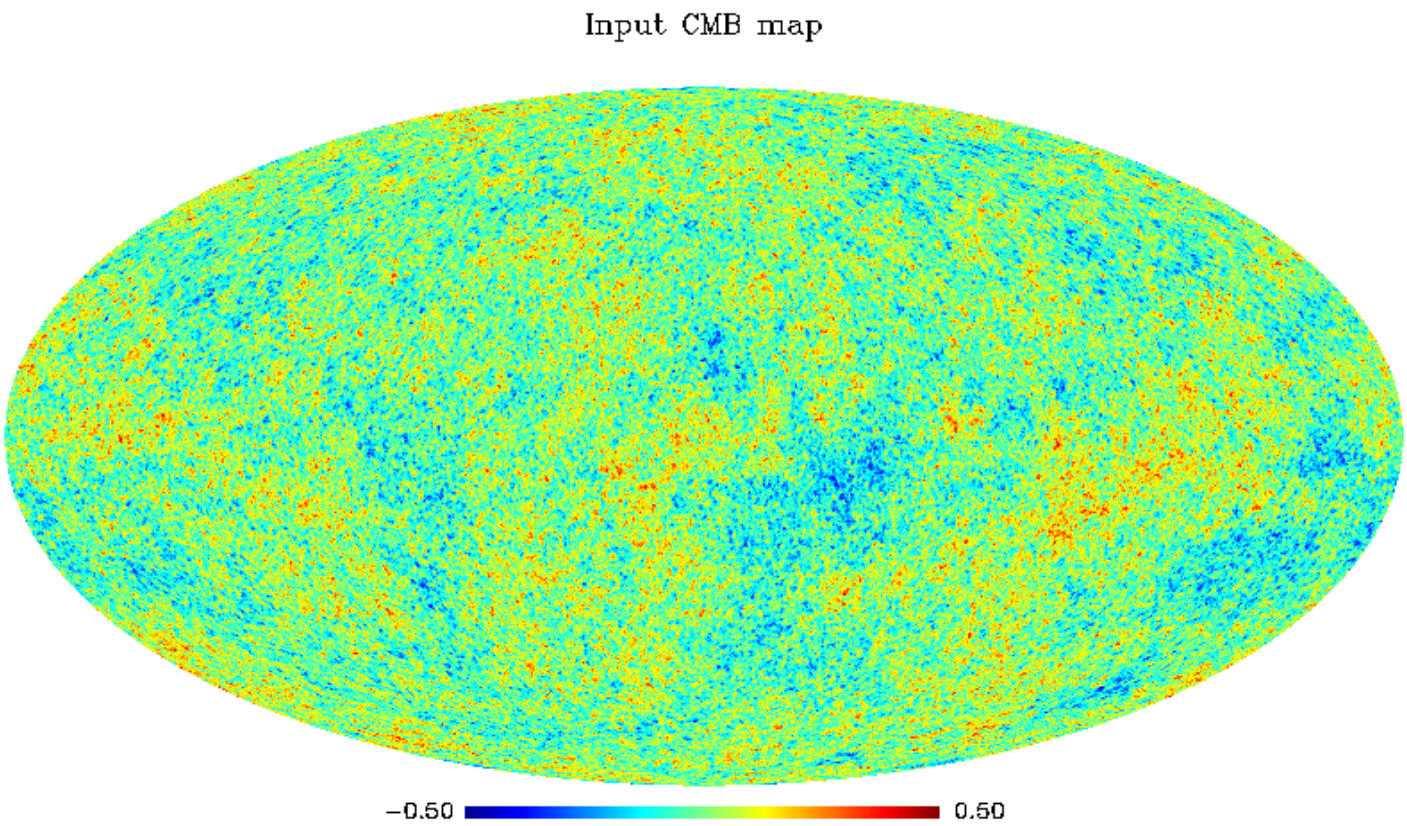
$\beta_{j,k} = \left\langle Y_i, \Psi_{j,k}^{(i)} \right\rangle$ = A_j $\alpha_{j,k} = \langle X_s, \psi_{j,k} \rangle$



The sparse GMCA solution is obtained by minimizing:

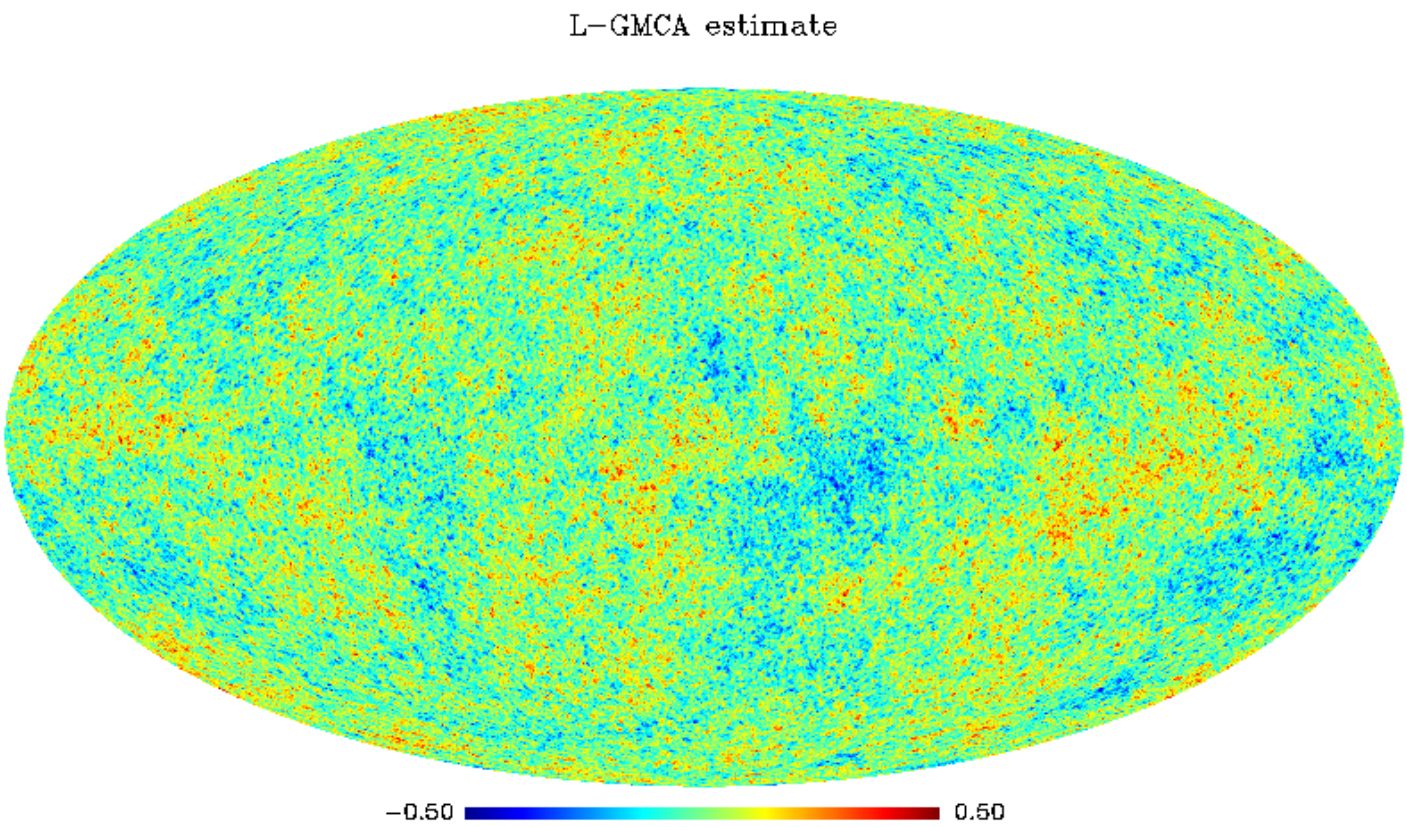
$$\min_{\alpha_j, A_j} \sum_j \frac{1}{2\sigma^2} \|\beta_j - A_j \alpha_j\|^2 \quad \text{s.t.} \quad \alpha \text{ is sparse}$$

CMB map estimation



Input CMB Map

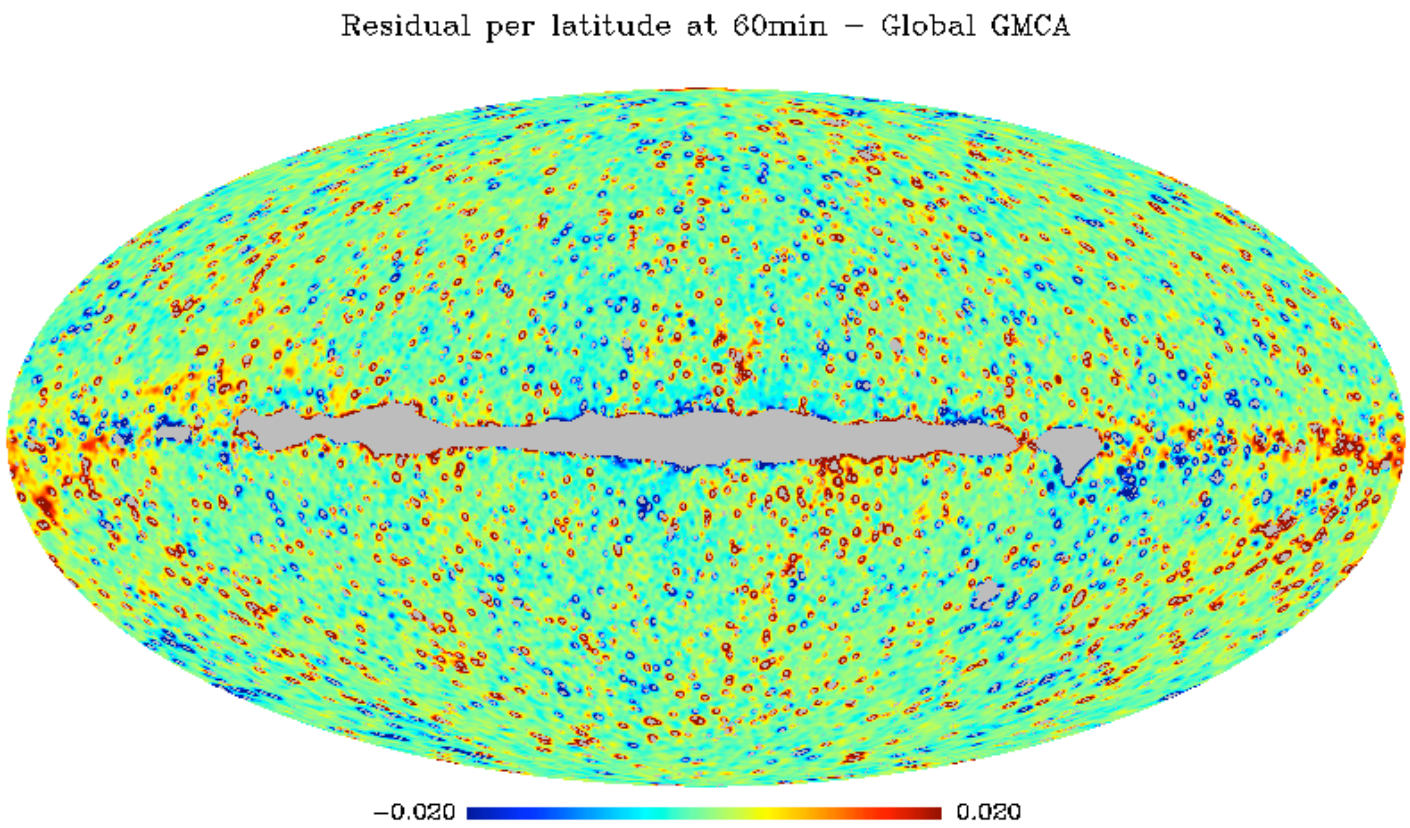
CMB map estimation



Estimate CMB Map

CMB map estimation

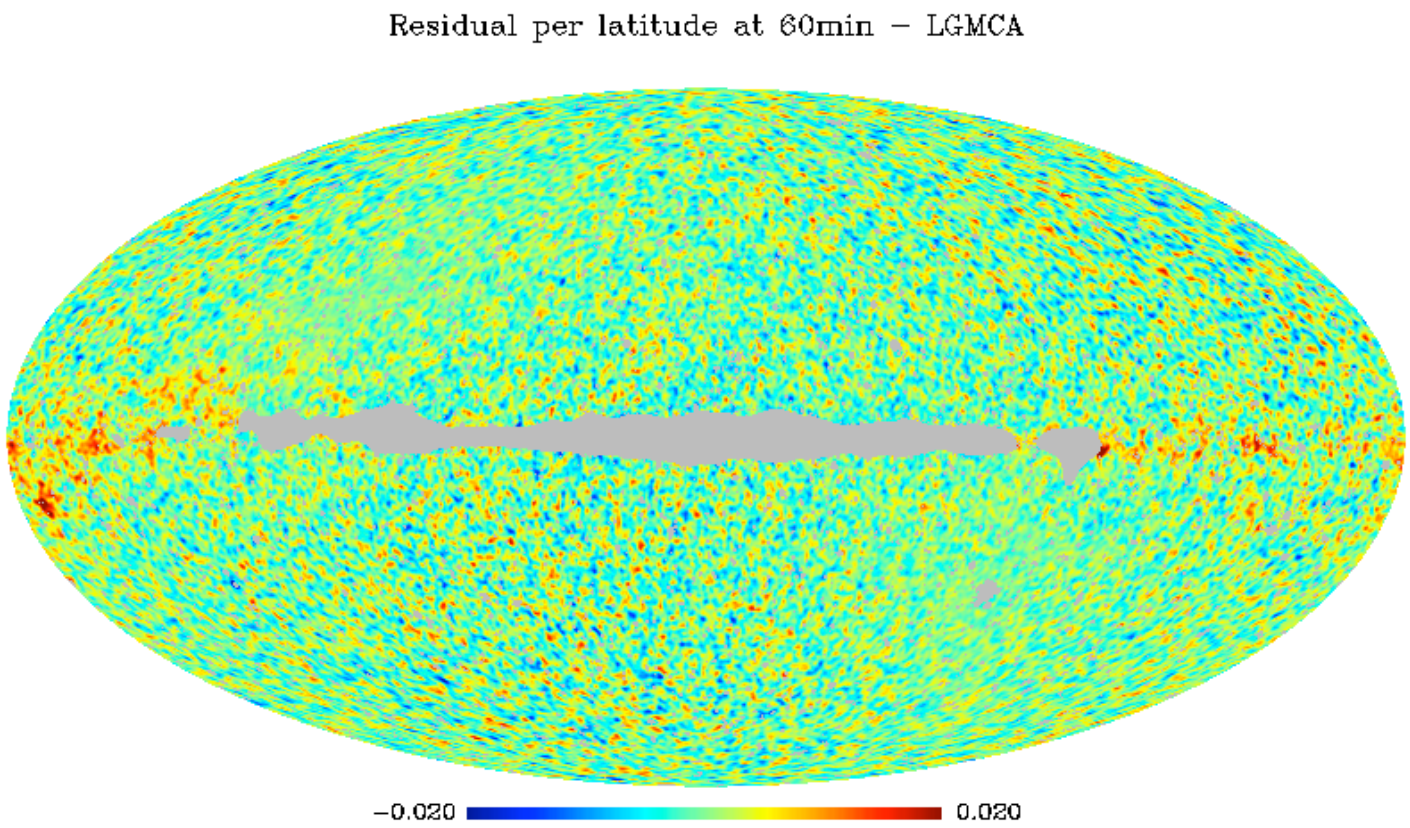
Residual at 60min



Residual map for Global GMCA

CMB map estimation

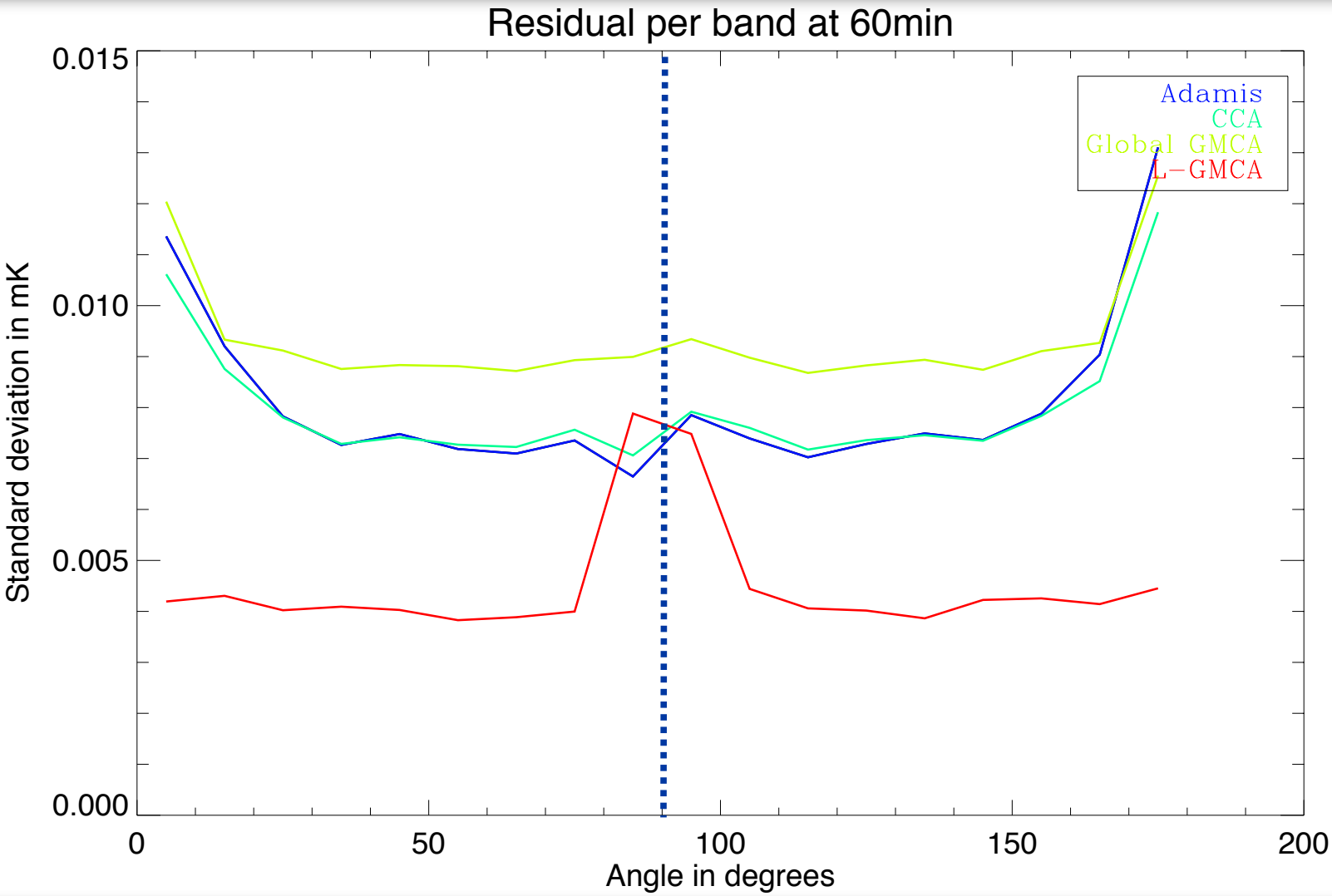
Residual at 60min



Residual map for L-GMCA

CMB map estimation

Residual per latitude



Wiener Filtering

Noise have an additive contribution to the CMB map : $y = x + n$

Standard approach - Wiener in spherical harmonics: $\forall \ell, m > 0; \quad a^{\hat{x}}_{\ell m} = \frac{C_\ell}{C_\ell + C^n_\ell} a^y_{\ell m}$

But does not account for the non-stationarity of the noise !!

More rigorously,

$$\min_x \quad x^T \mathcal{F}^H \mathbf{C}^{-1} \mathcal{F} x + (y - x)^T \mathbf{\Sigma}^{-1} (y - x)$$

CMB power spectrum

Sph. Harmonics

Noise covariance matrix

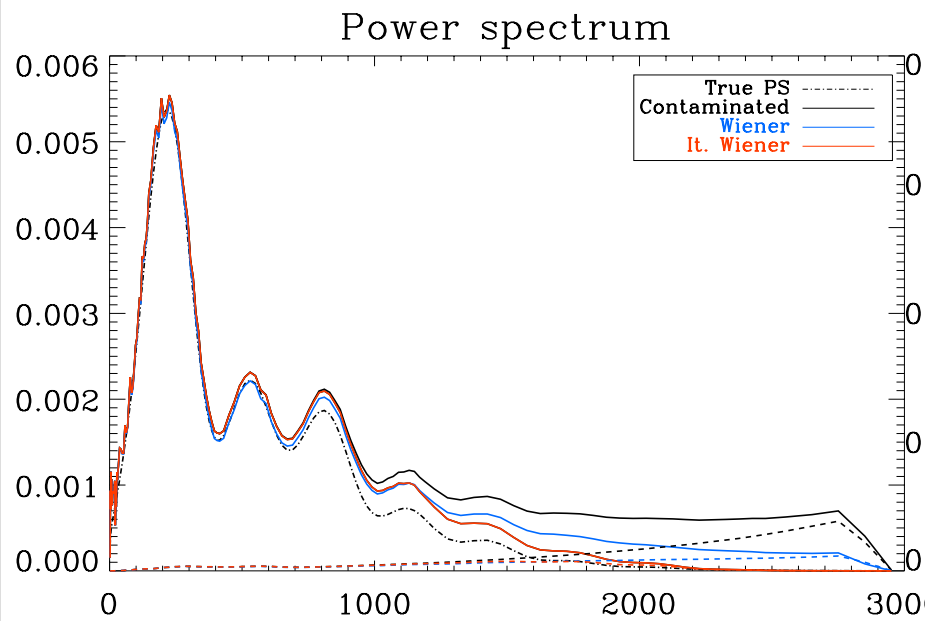
But wavelets are well known to decorrelate 1/f noise ==>

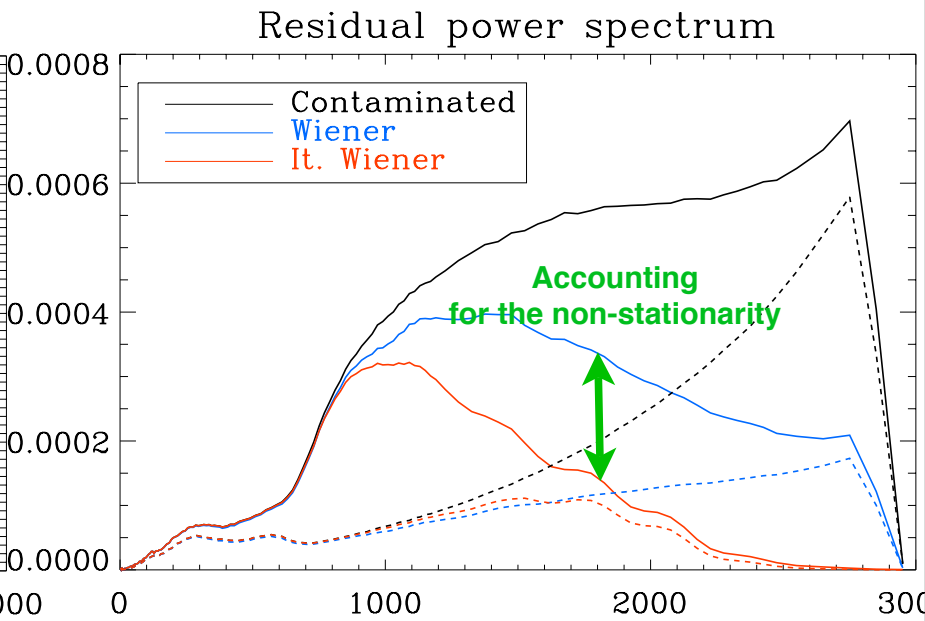
LInear Wavelet-based Filtering (Iterative Wiener Algorithm)

- J. Bobin, J.-L. Starck, F. Sureau, J. Fadili, "***CMB Map Restoration***, **http://arxiv.org/abs/1111.3149**, submitted.

• Wavelet-based Gaussian modeling of the noise covariance matrix, either from the JackKnife map or noise simulations.
• Solve the equation using an iterative Forward-Backward Algorithm.

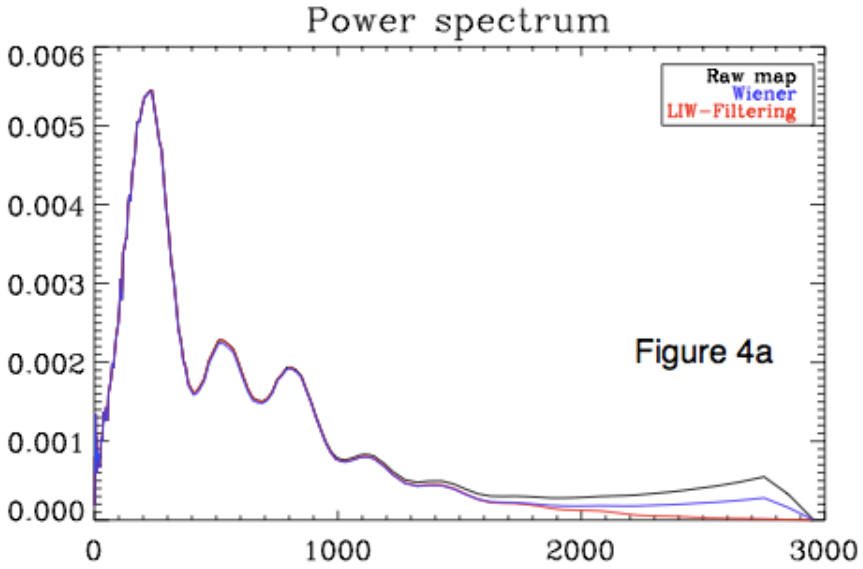
Result on WG2 Simulations

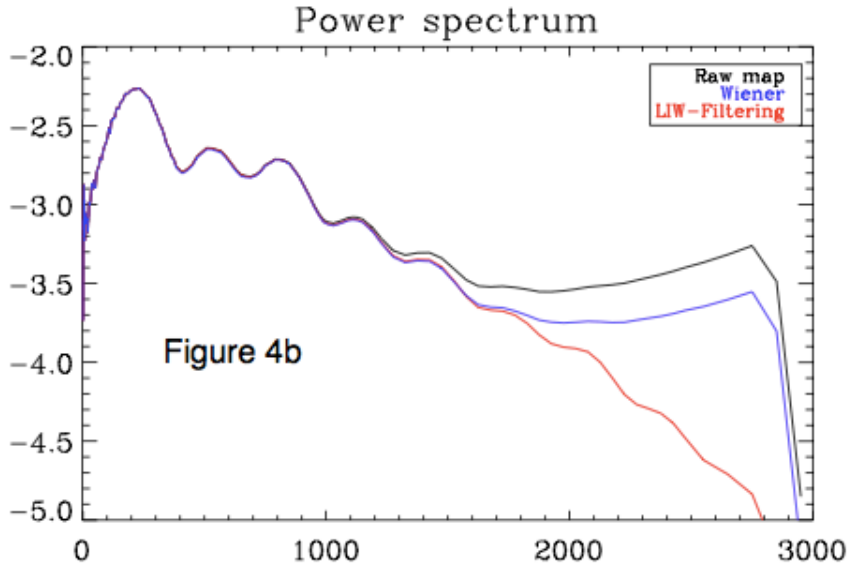




GMCA-LIW Filtering

CMB power spectrum of the CMB map computed by:
- GMCA CMB (black)
- GMCA + Wiener (blue)
- GMCA + LIW-filtering (in red).





Beyond Noise Reduction

Whatever the estimator used to estimate the CMB map, the CMB map will be contaminated by spurious foreground residuals

$$y = x + n + f$$

We use exactly the same wavelet-based Gaussian modeling.

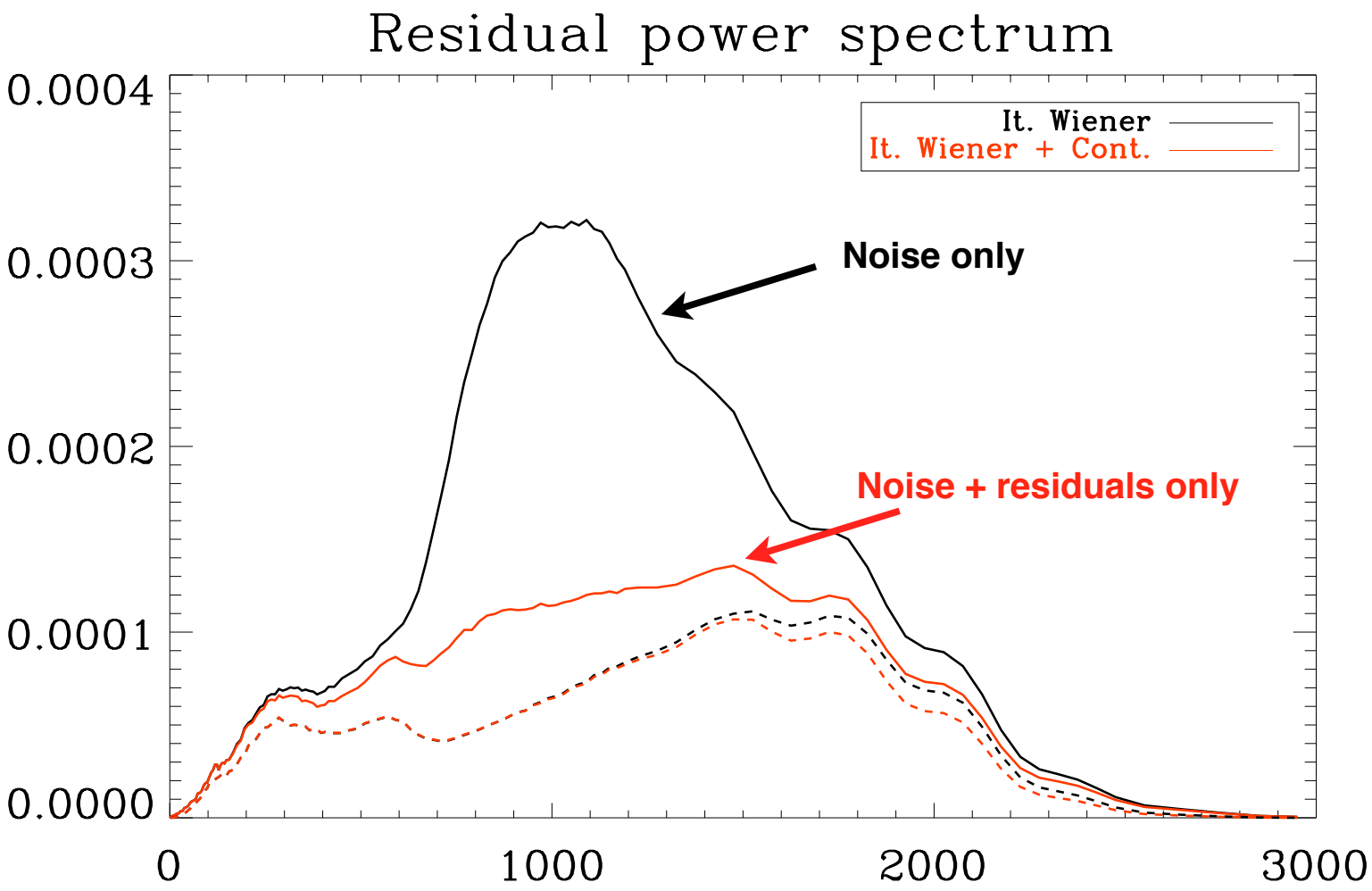
From the decorrelation of CMB, noise and foregrounds:

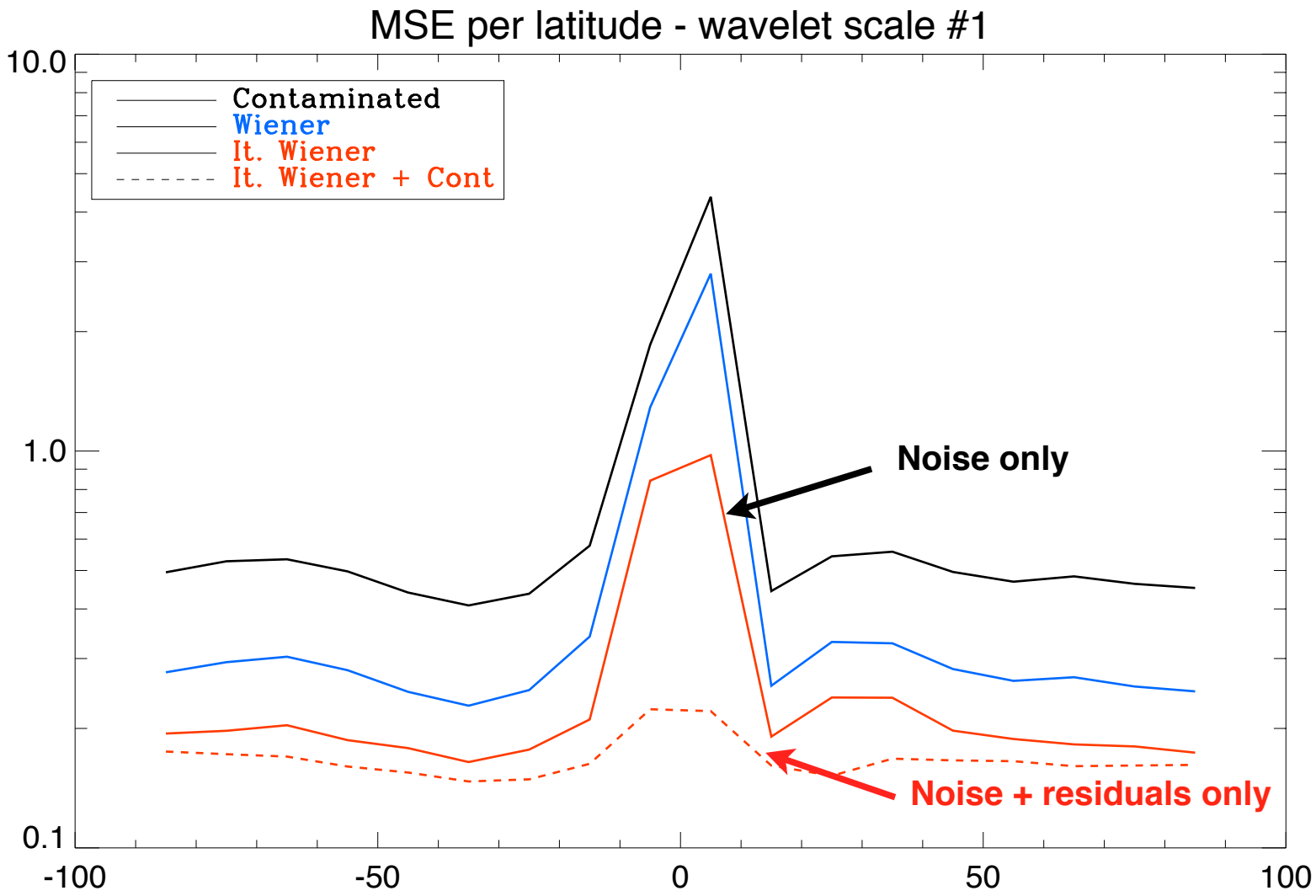
$$\hat{\sigma}_j^y[k]^2 = \sigma_j^2 + \hat{\sigma}_j^n[k]^2 + \hat{\sigma}_j^f[k]^2$$

CMB Noise Foregrounds

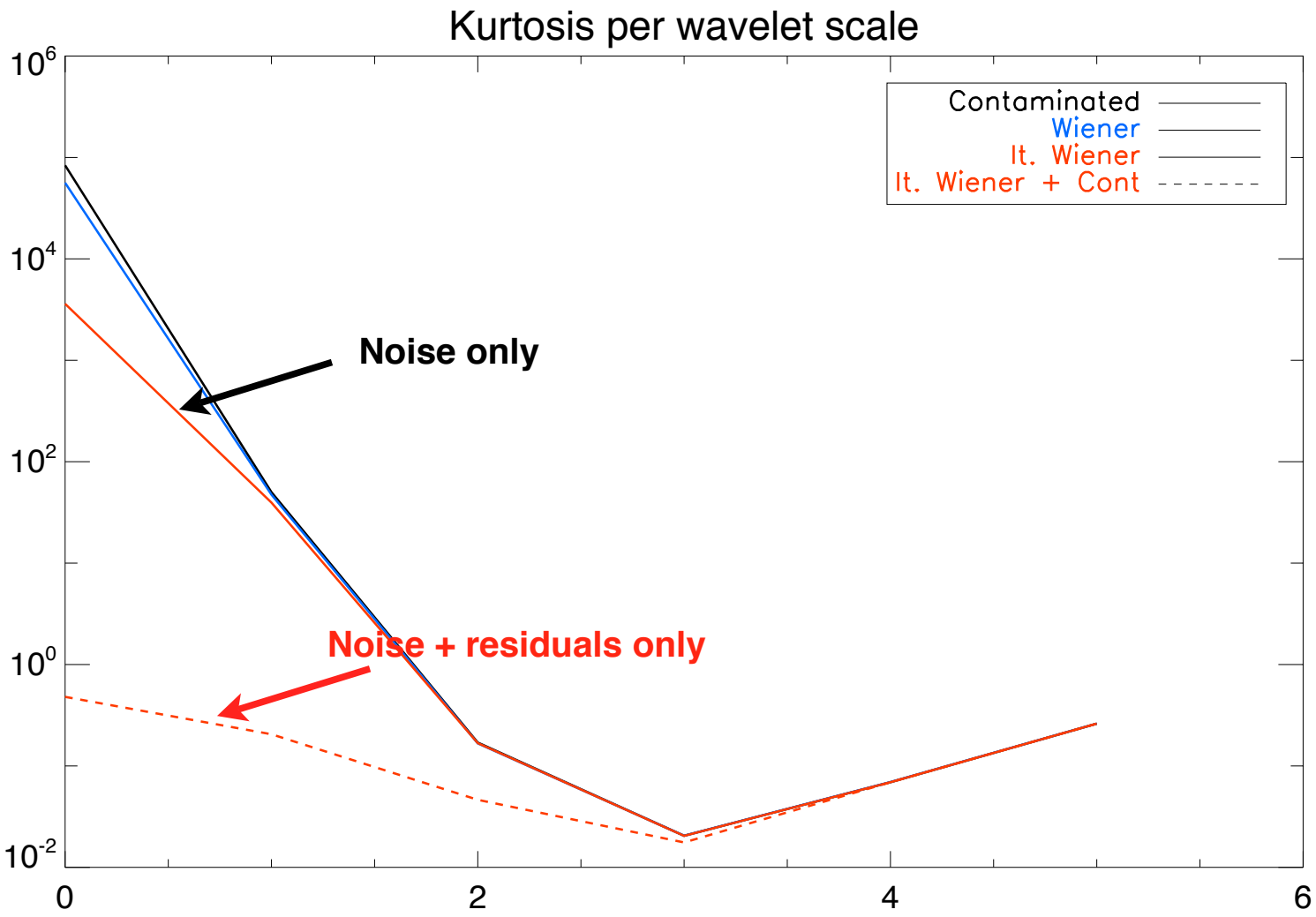
$$\hat{\sigma}_j^f[k]^2 = \left[\hat{\sigma}_j^y[k]^2 - \frac{1}{4\pi}\sum_\ell (2\ell+1)\psi_{j,\ell}^2 C_\ell - \hat{\sigma}_j^n[k]^2\right]_+$$

Some results

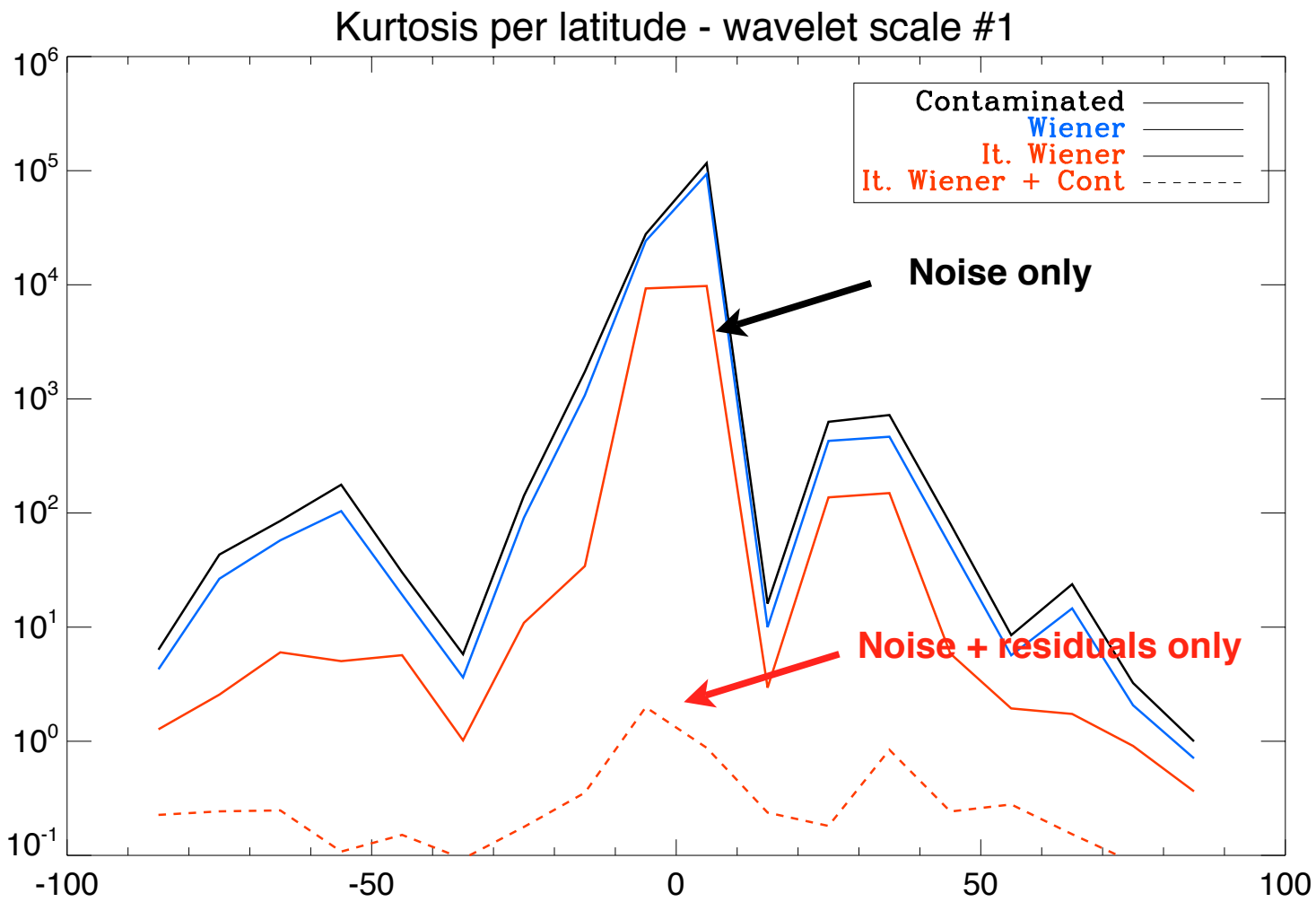



MSE per latitude - wavelet scale #1
10.0
1.0
0.1
-100
-50
0
50
100
Contaminated
Wiener
It. Wiener
It. Wiener + Cont
Noise only
Noise + residuals only

Some results



Some results



Sparsity and CMB

- Sparsity is very efficient for
  - Inverse problems (denoising, deconvolution, etc).
  - Inpainting
  - Component Separation.
  - Wiener Wiltering.
- Perspectives
  - Estimator Aggregation.
  - Dictionary Learning.