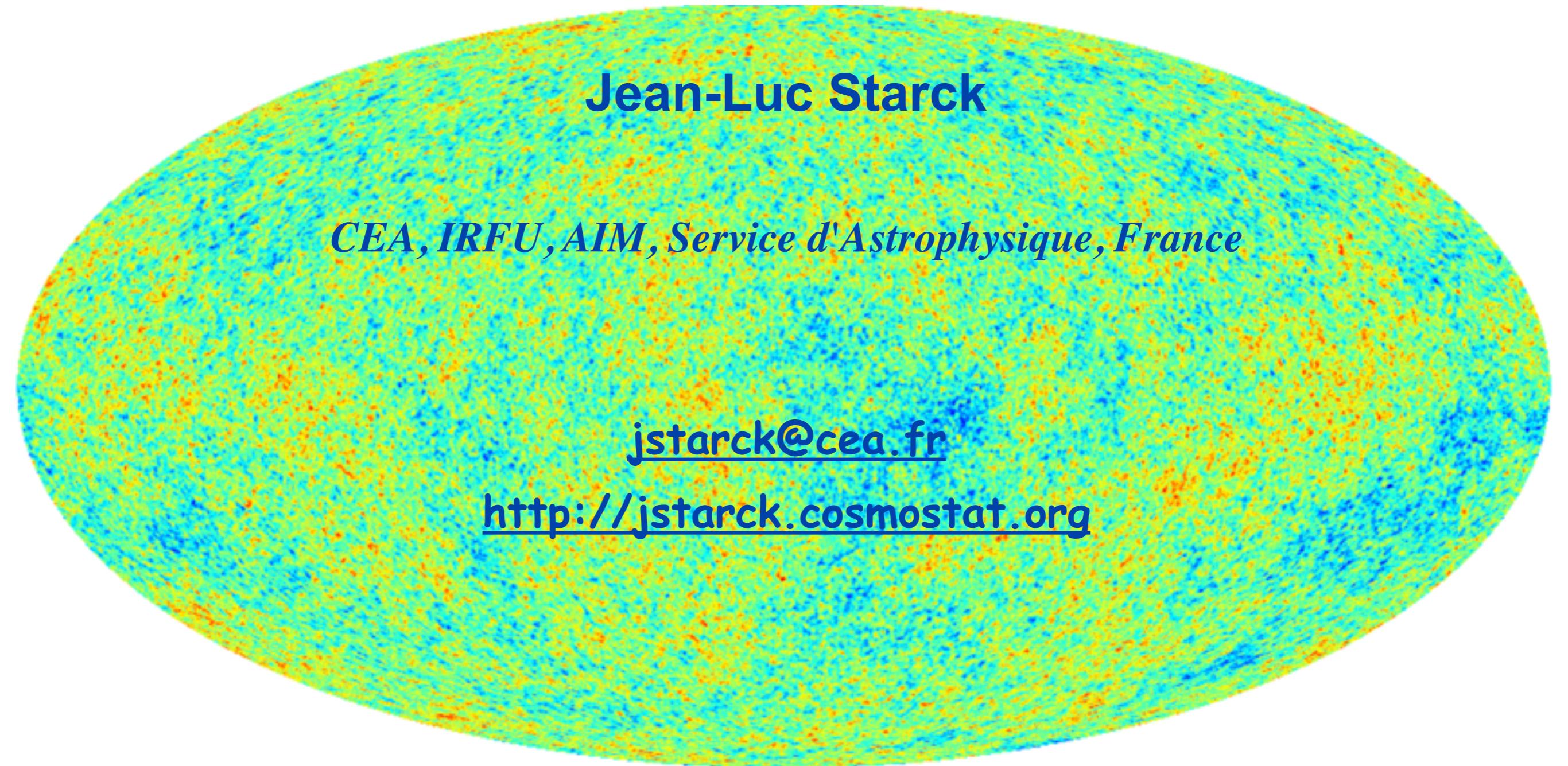


Sparsity & Inverse Problems in Astrophysics



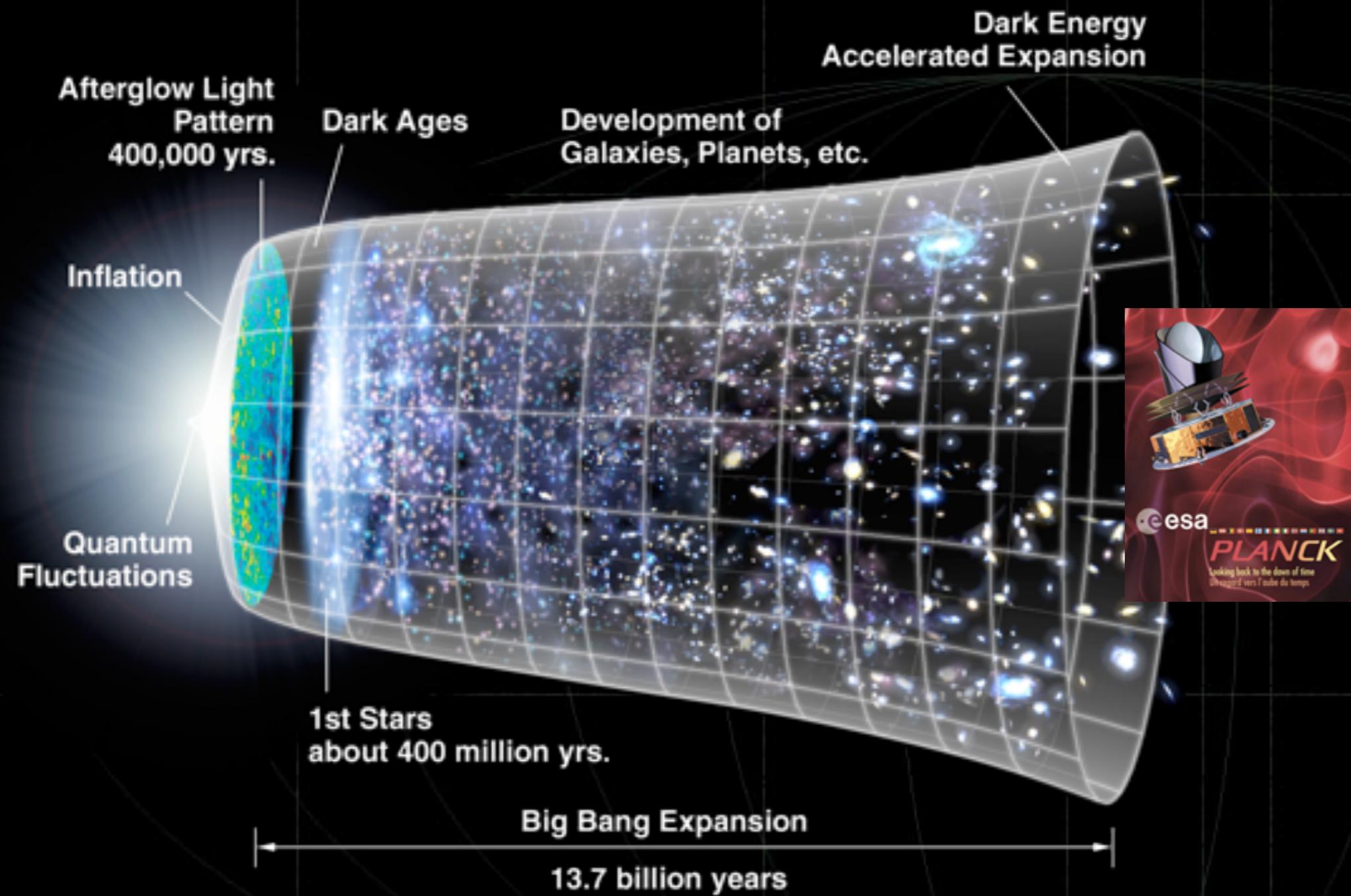
Jean-Luc Starck

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

jstarck@cea.fr

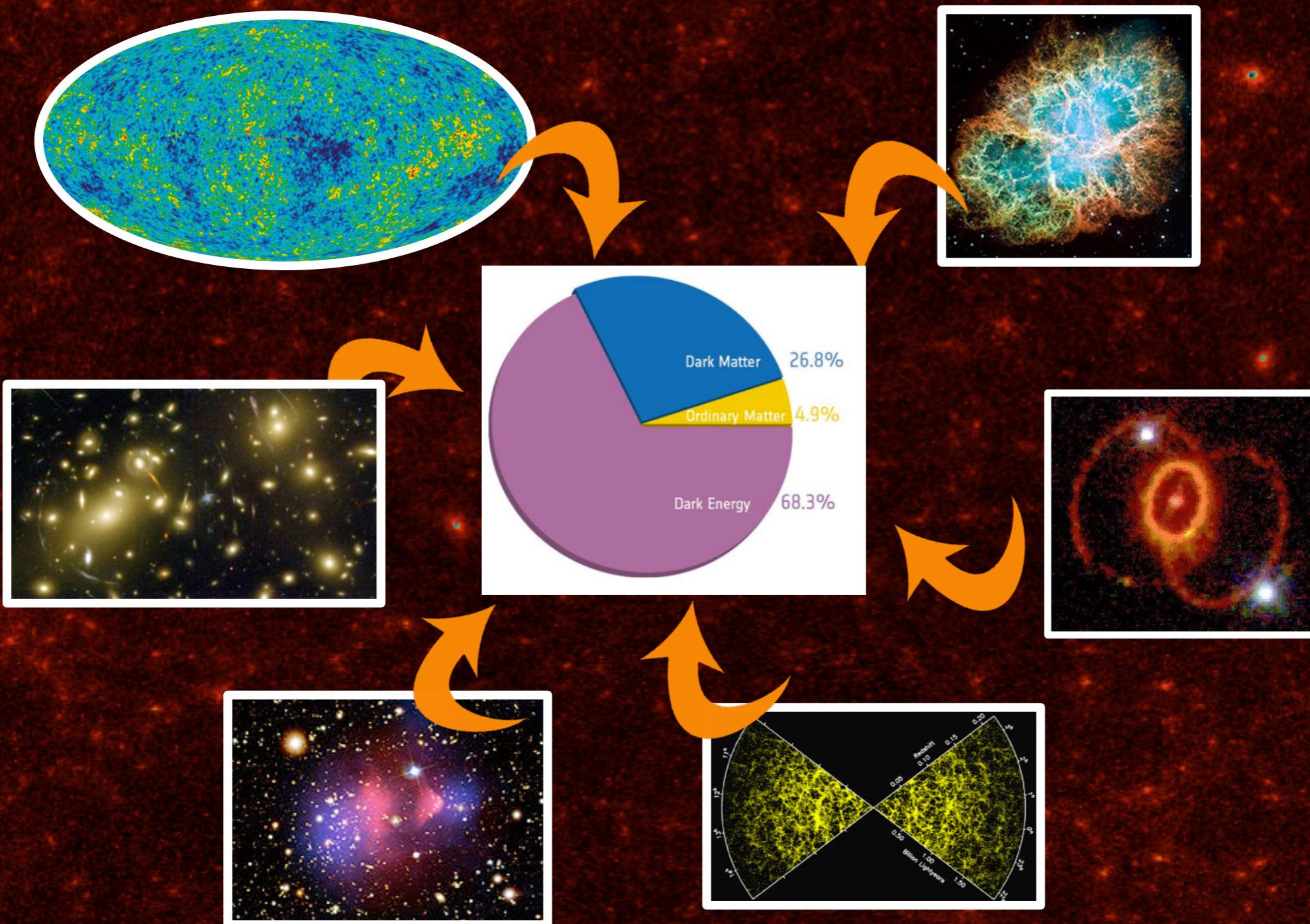
<http://jstarck.cosmostat.org>

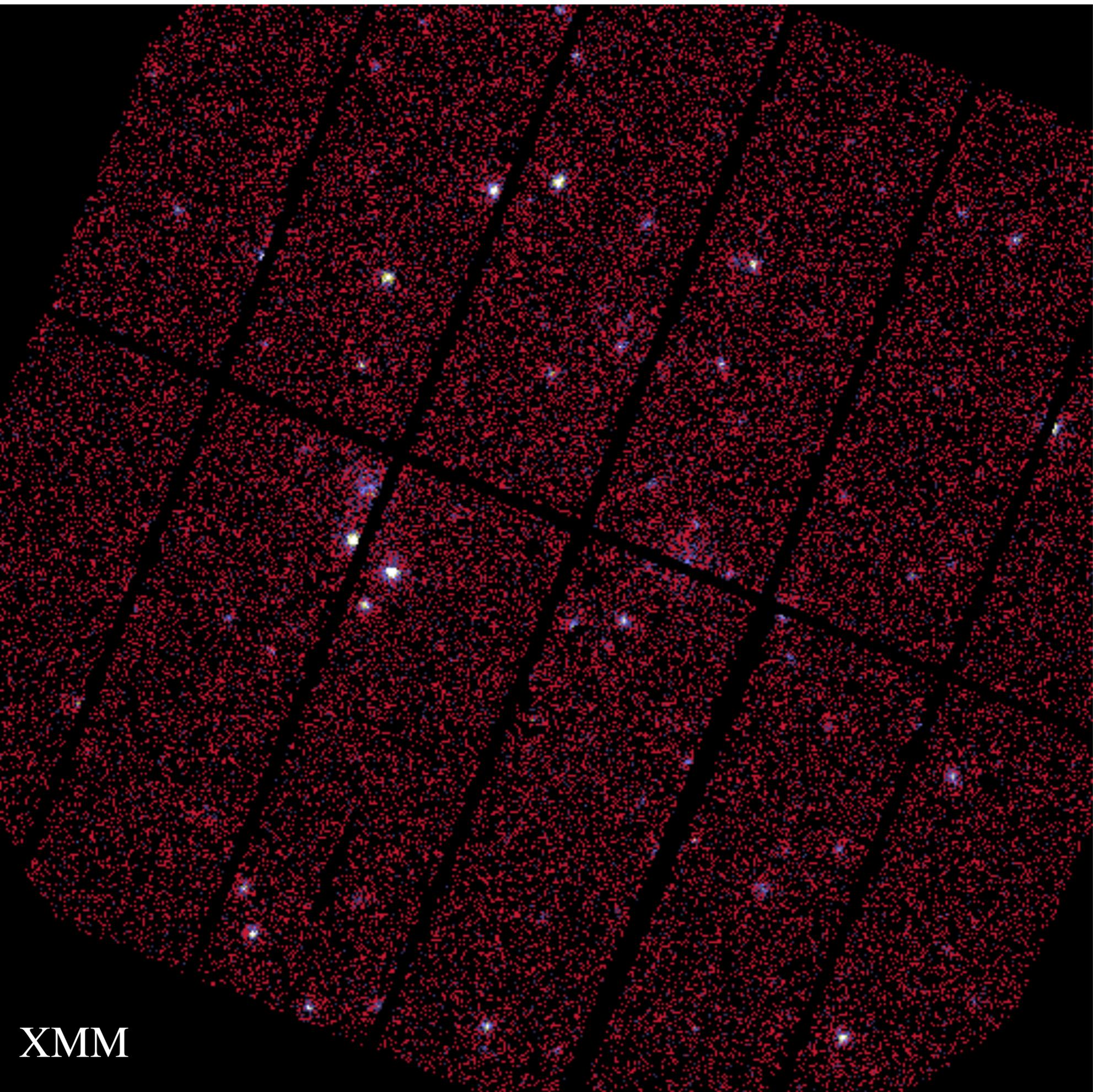
- Part I: Introduction to Inverse Problems in Astrophysics
- Part II: Sparsity & Dictionaries
- Part III: Sparse Regularization
- Part IV: Unmixing
- Part V: Sparsity for Planck and Euclid Space Missions



From observations to cosmological model

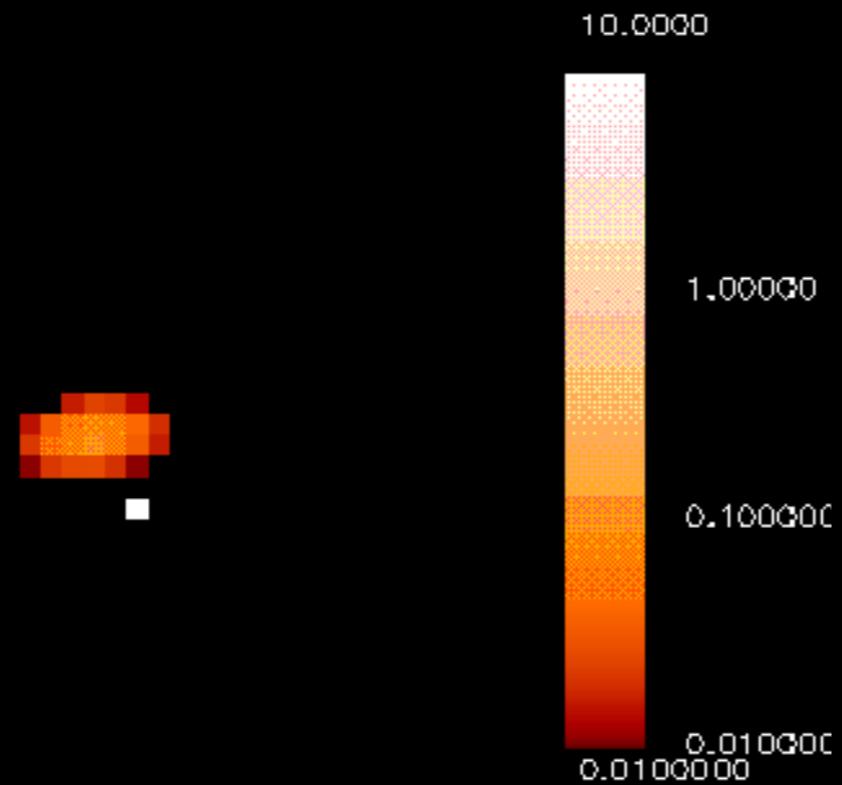
$$\mathcal{M}(\Omega_M, \Omega_\Lambda, \Omega_b, \sigma_8, \dots)$$



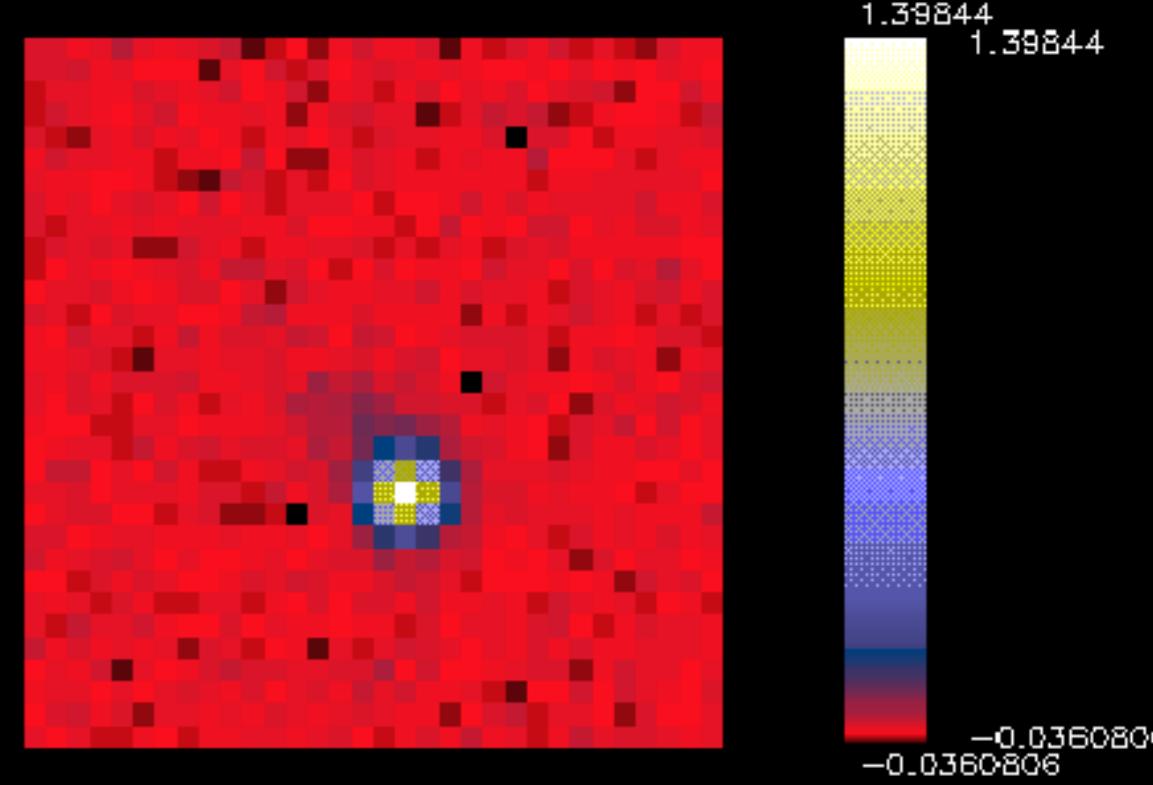


XMM

Simulation : faint galaxy nearby a bright star : origin

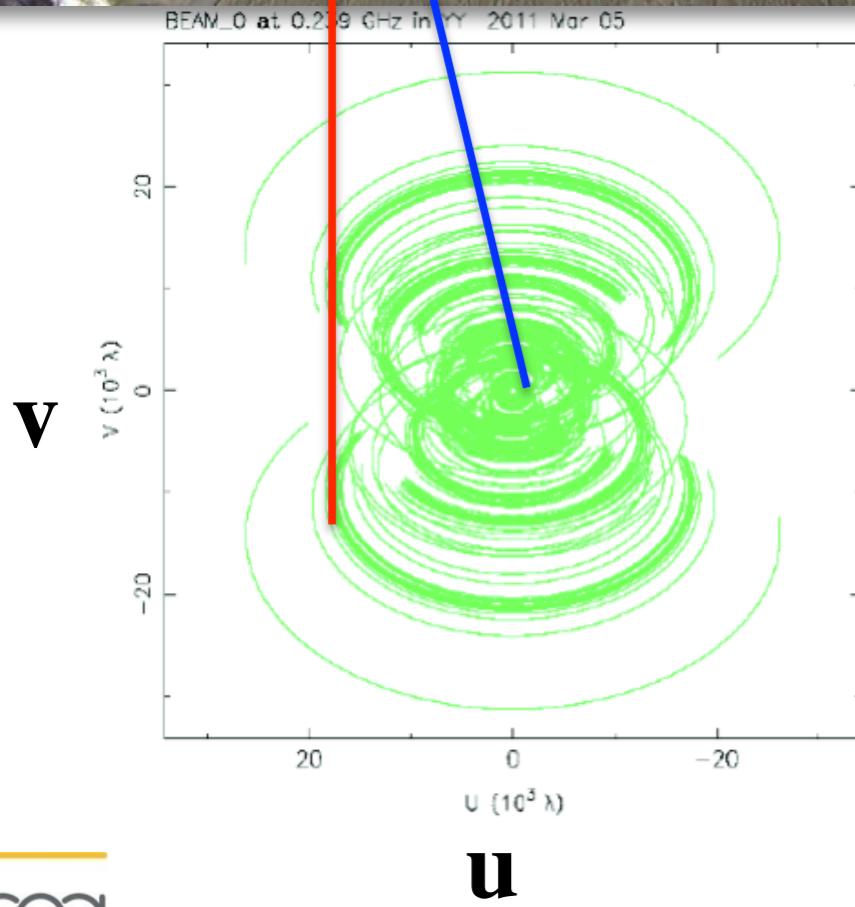
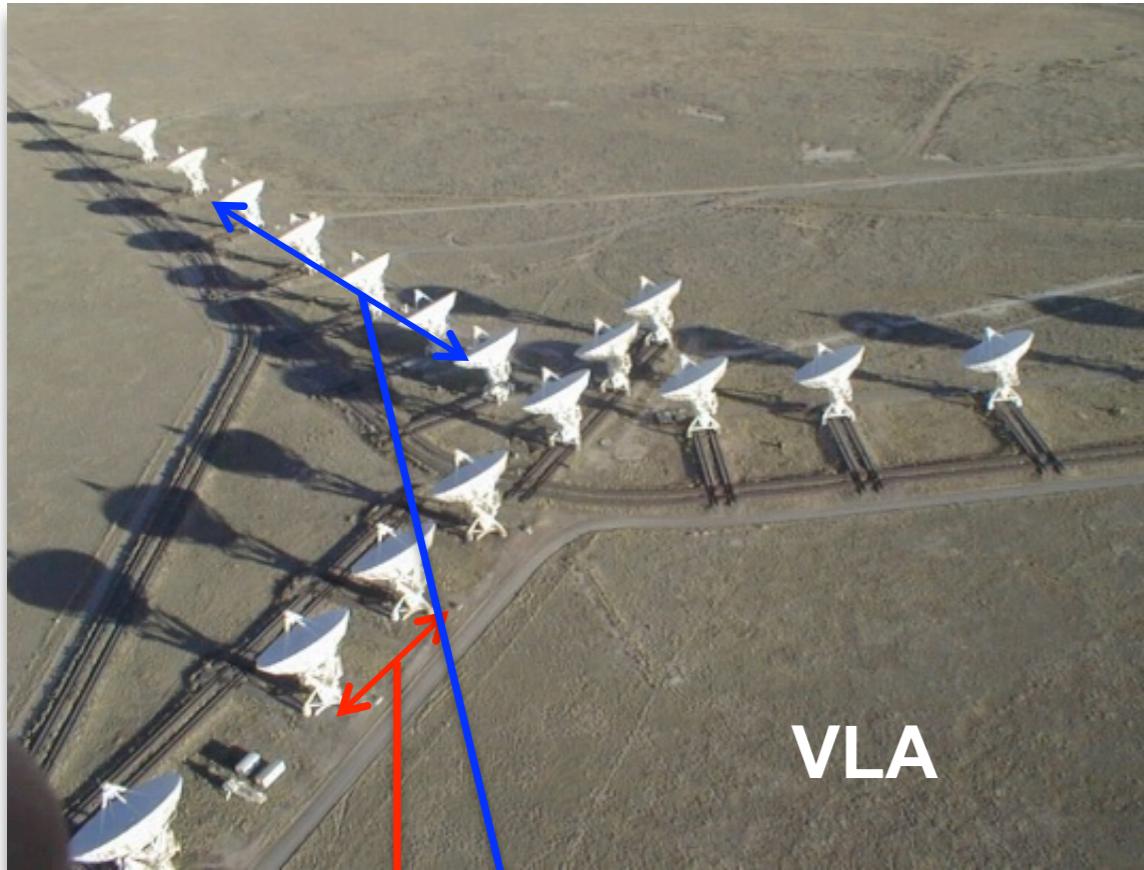


Simulation:weak galax. near a bright *, convolv. with IS



Max em 17 11

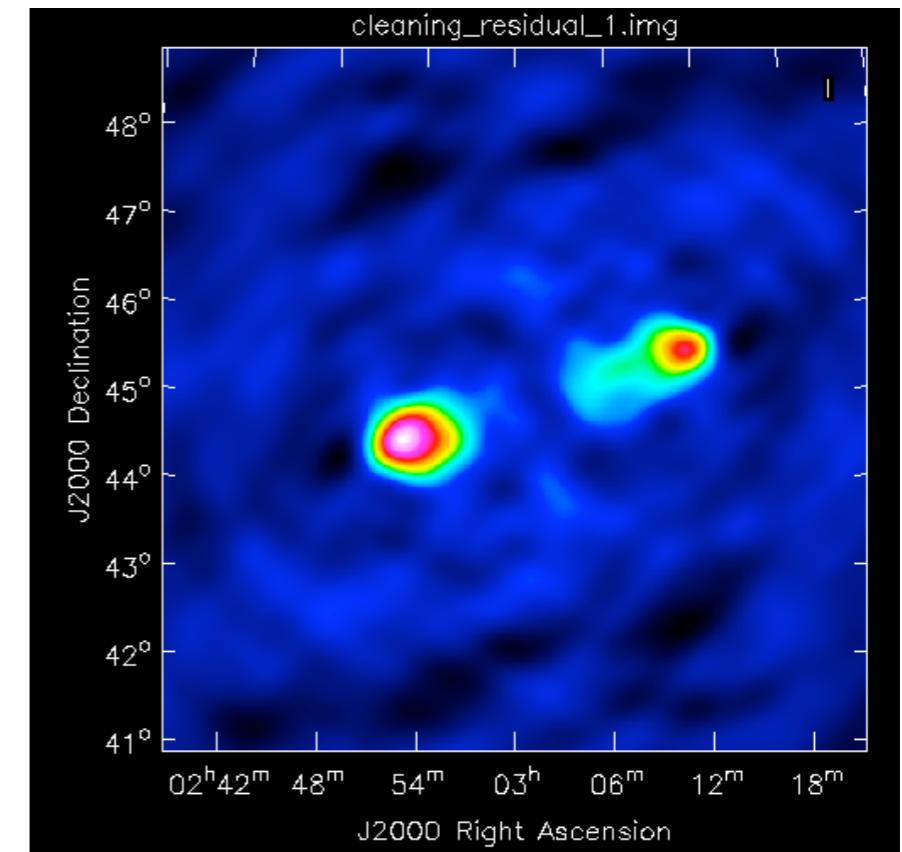
Imaging with interferometry



N antennas/telescopes

$\frac{N(N - 1)}{2}$ independent baselines

1 projected baseline
= 1 sample in the Fourier « u,v » plane

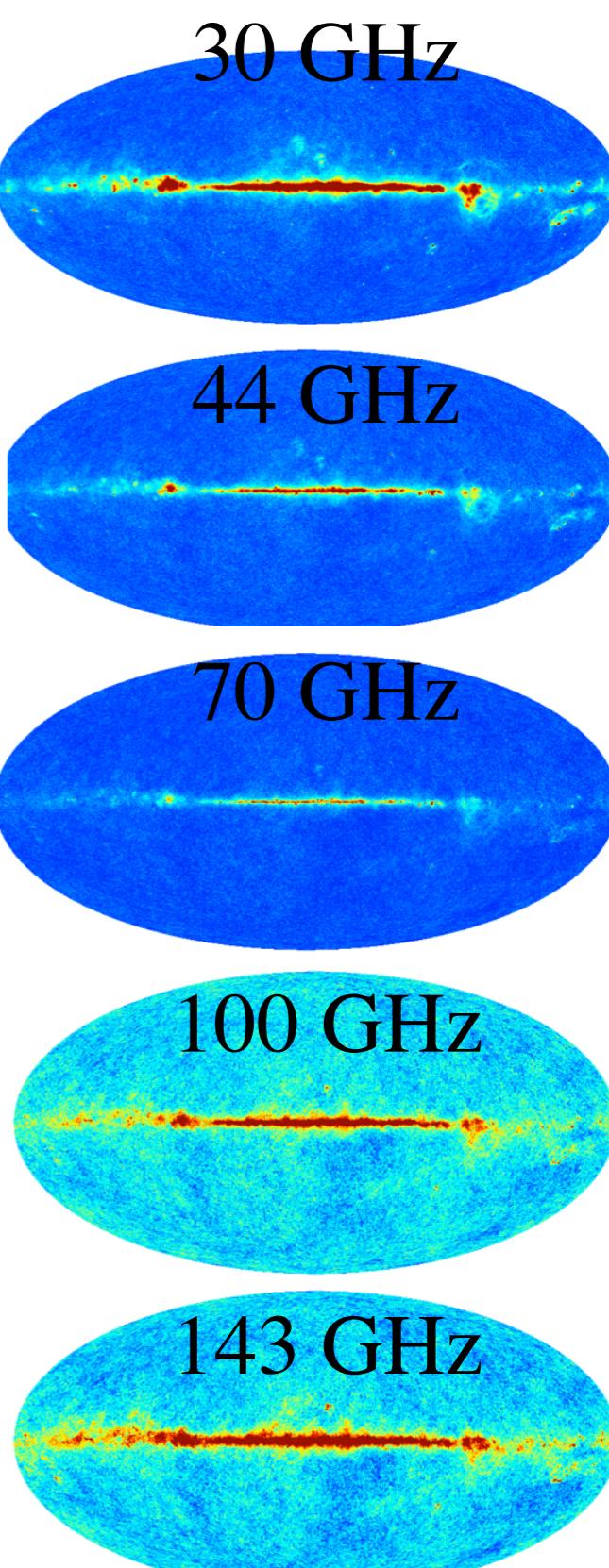
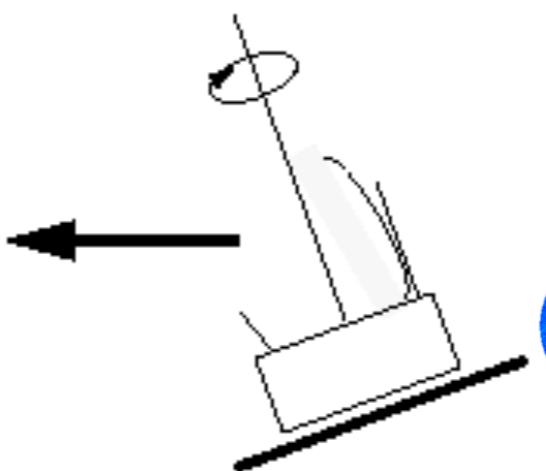
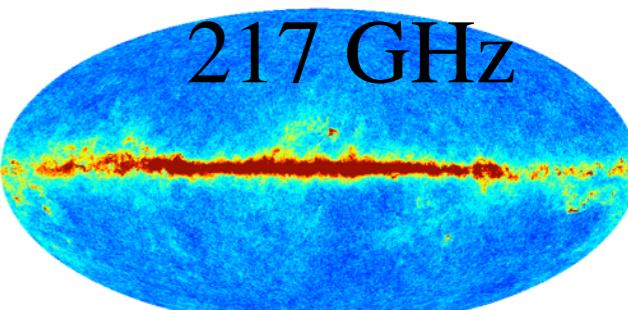
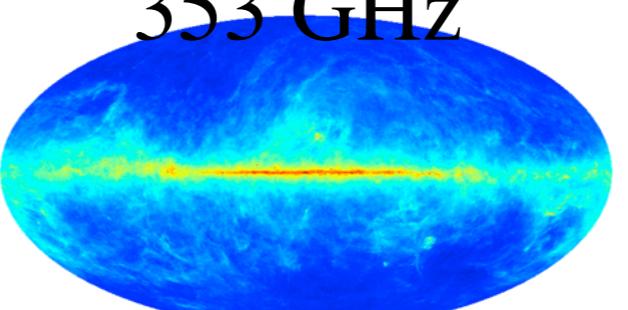
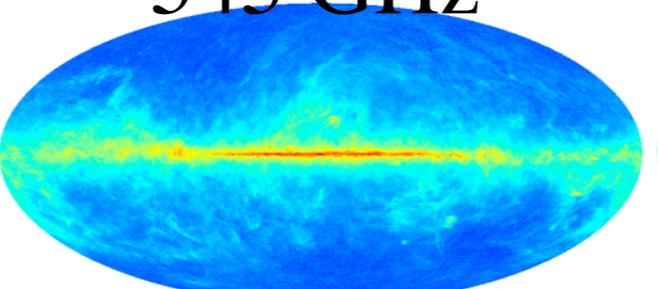
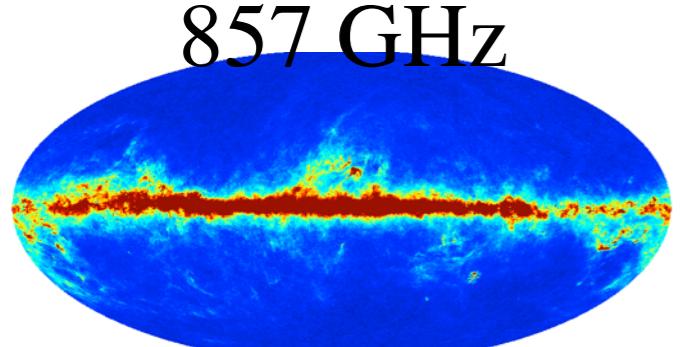
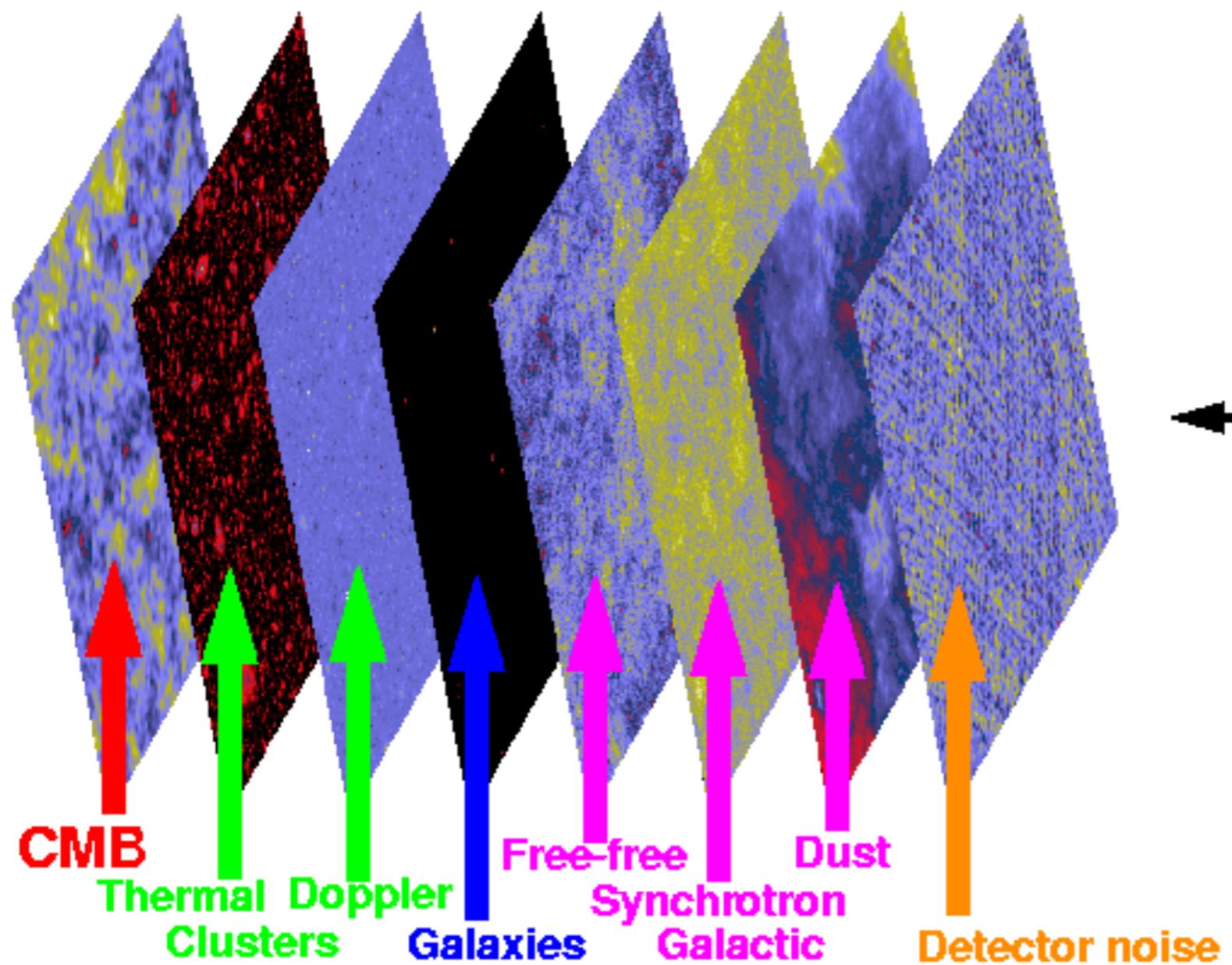


Gemini infrared data

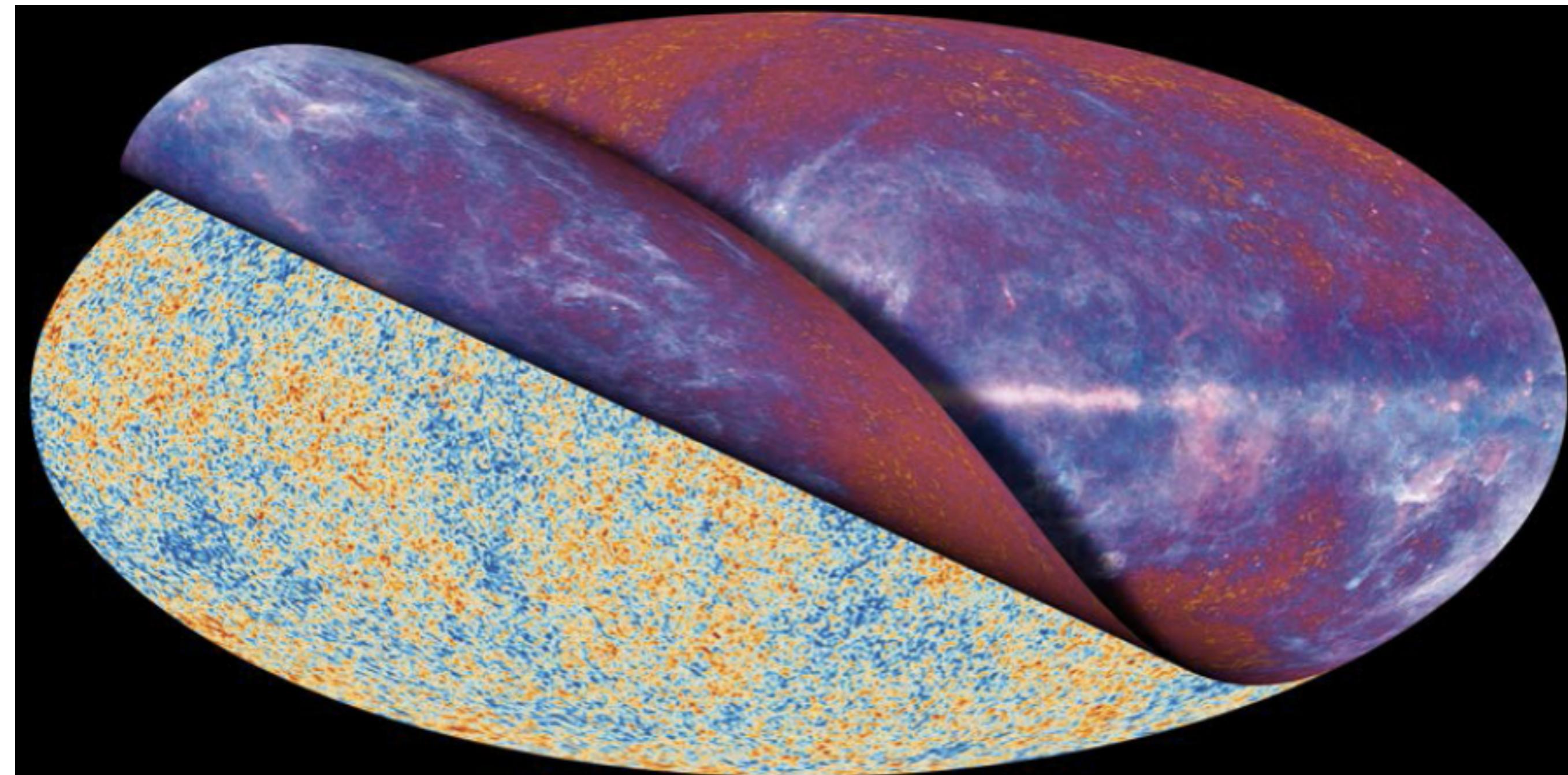
Galaxy SBS 0335-052



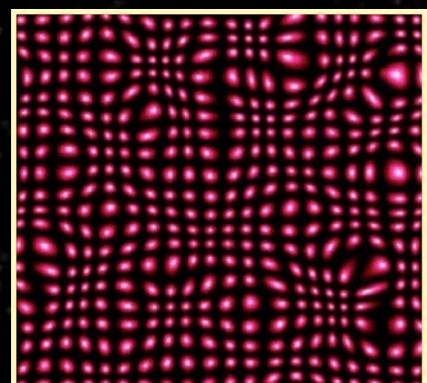
Planck Component Separation



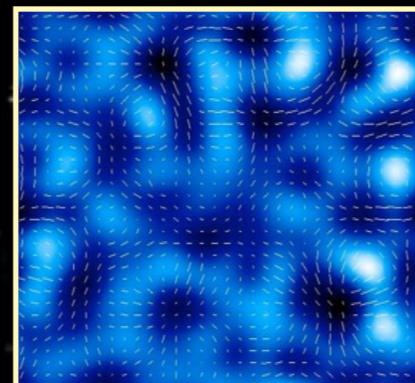
Foreground Removal



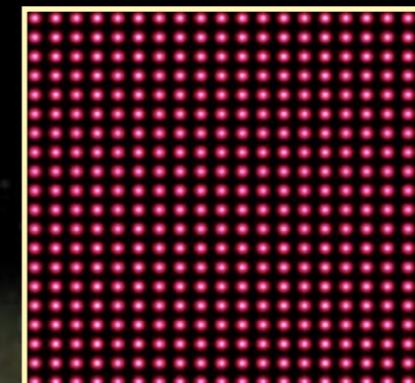
Weak Lensing



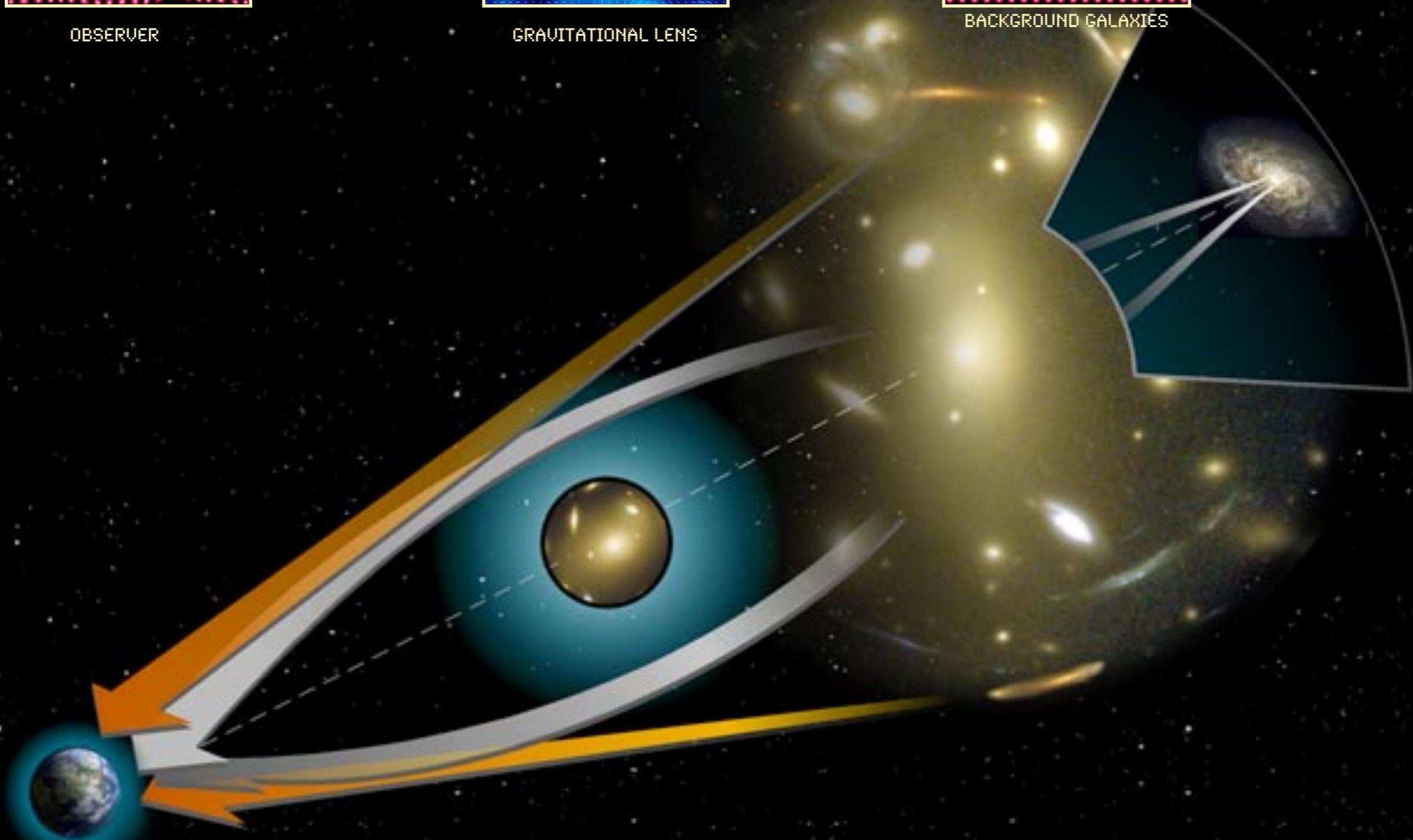
OBSERVER



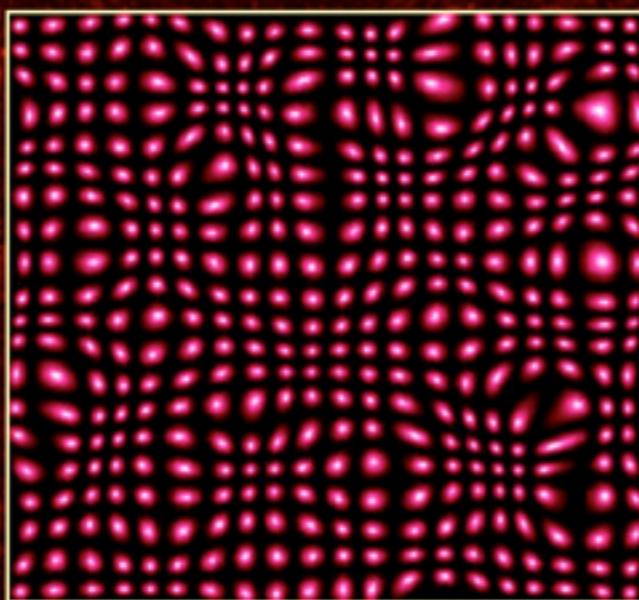
GRAVITATIONAL LENS



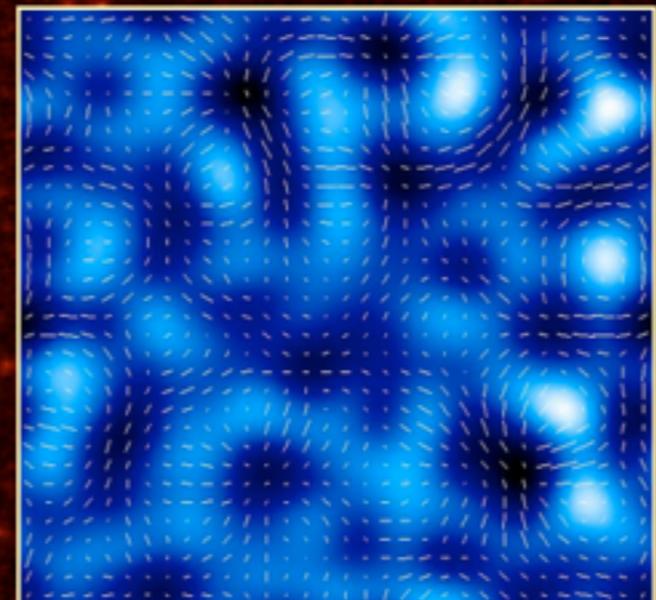
BACKGROUND GALAXIES



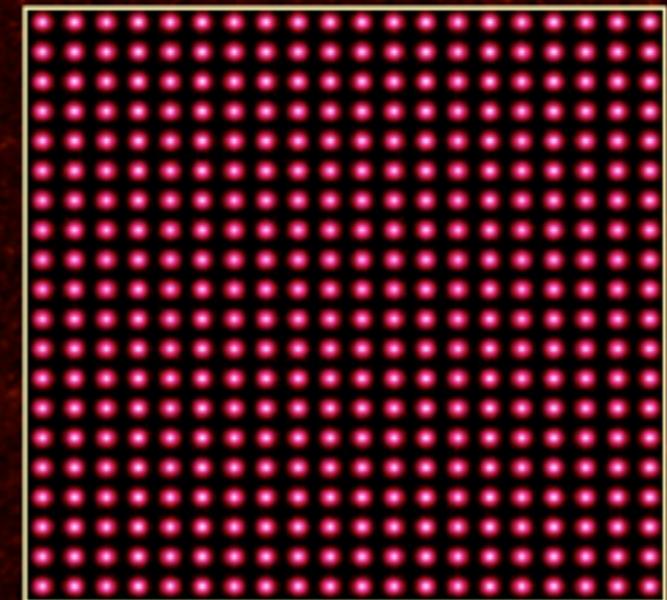
Weak Gravitational Lensing



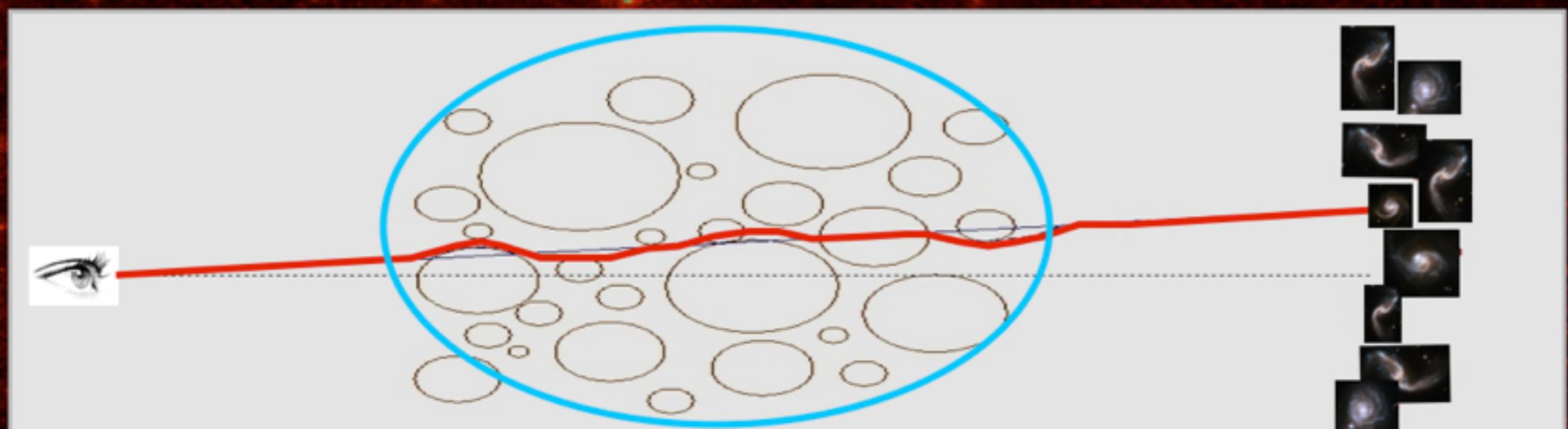
OBSERVER



GRAVITATIONAL LENS

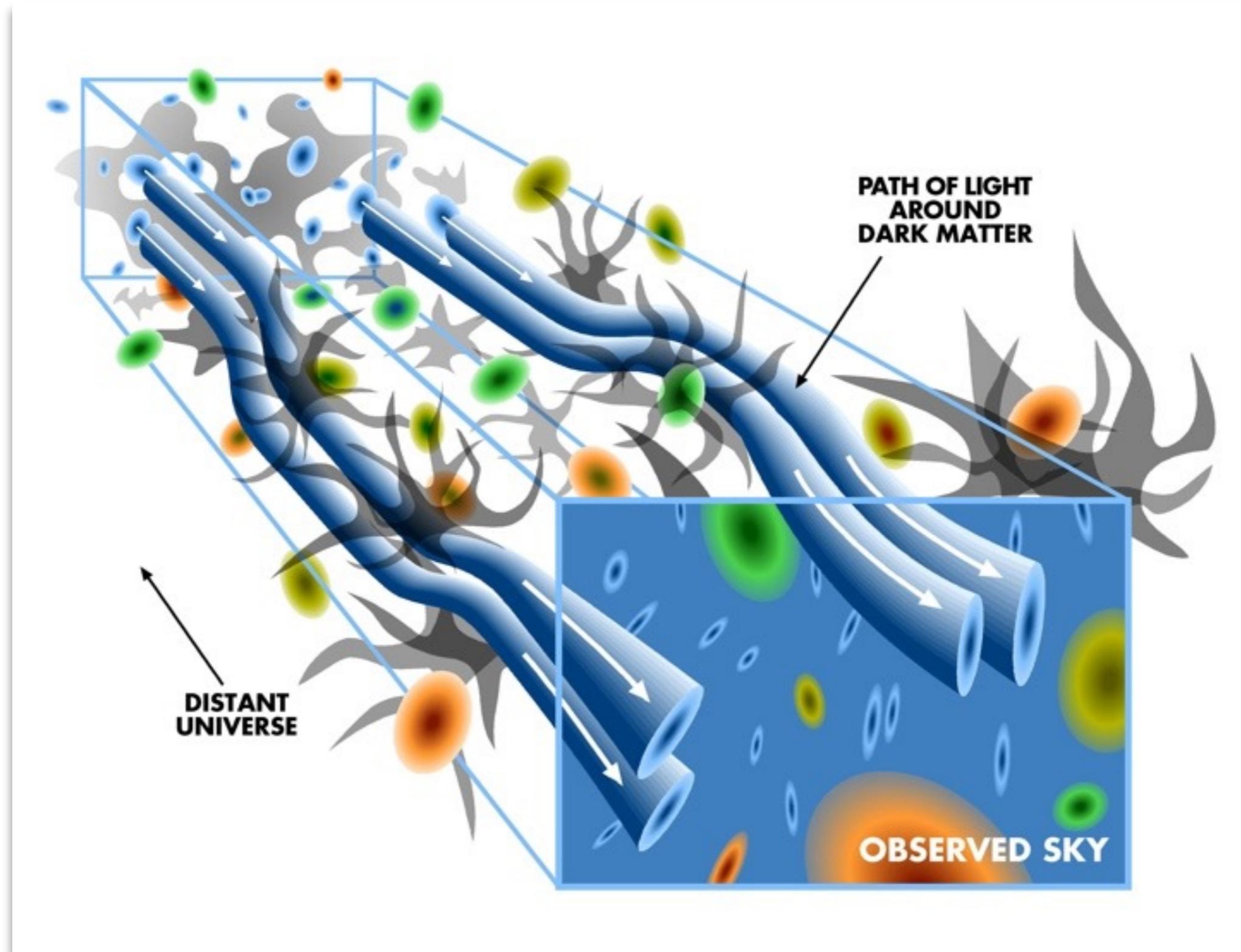


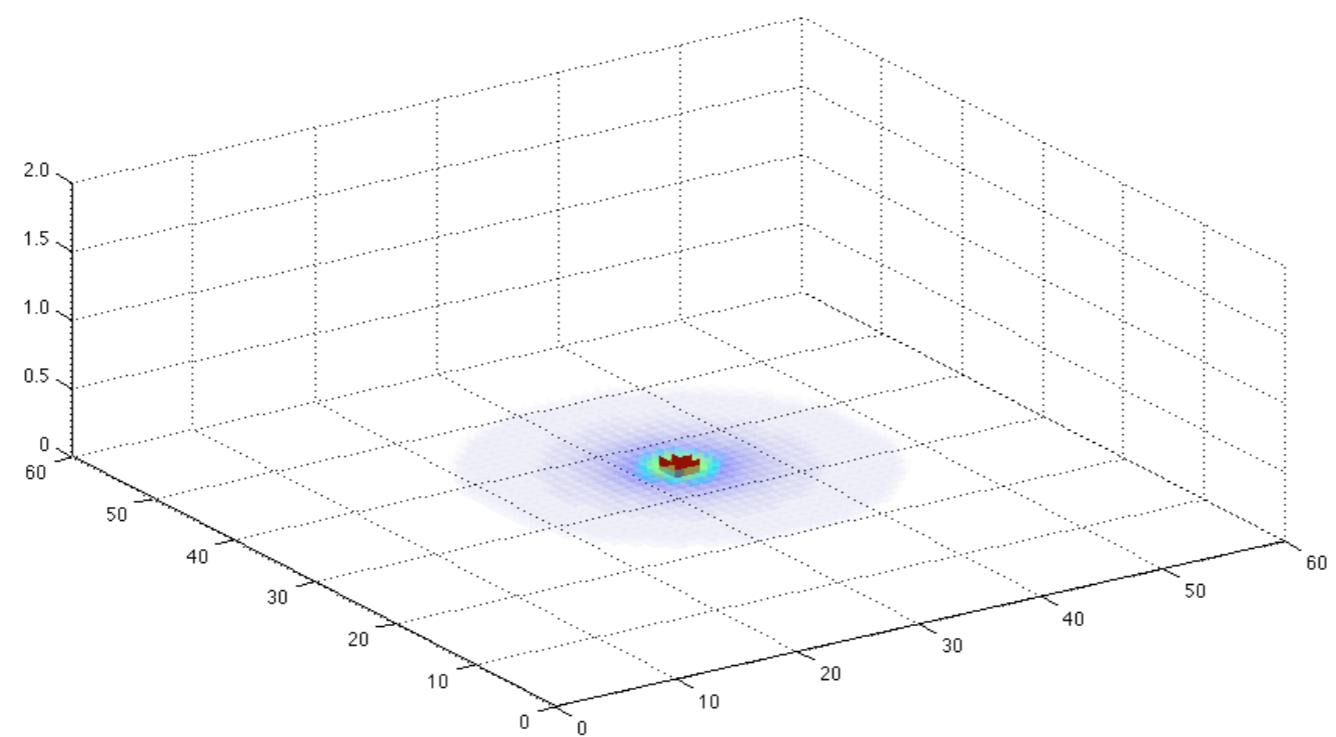
BACKGROUND GALAXIES

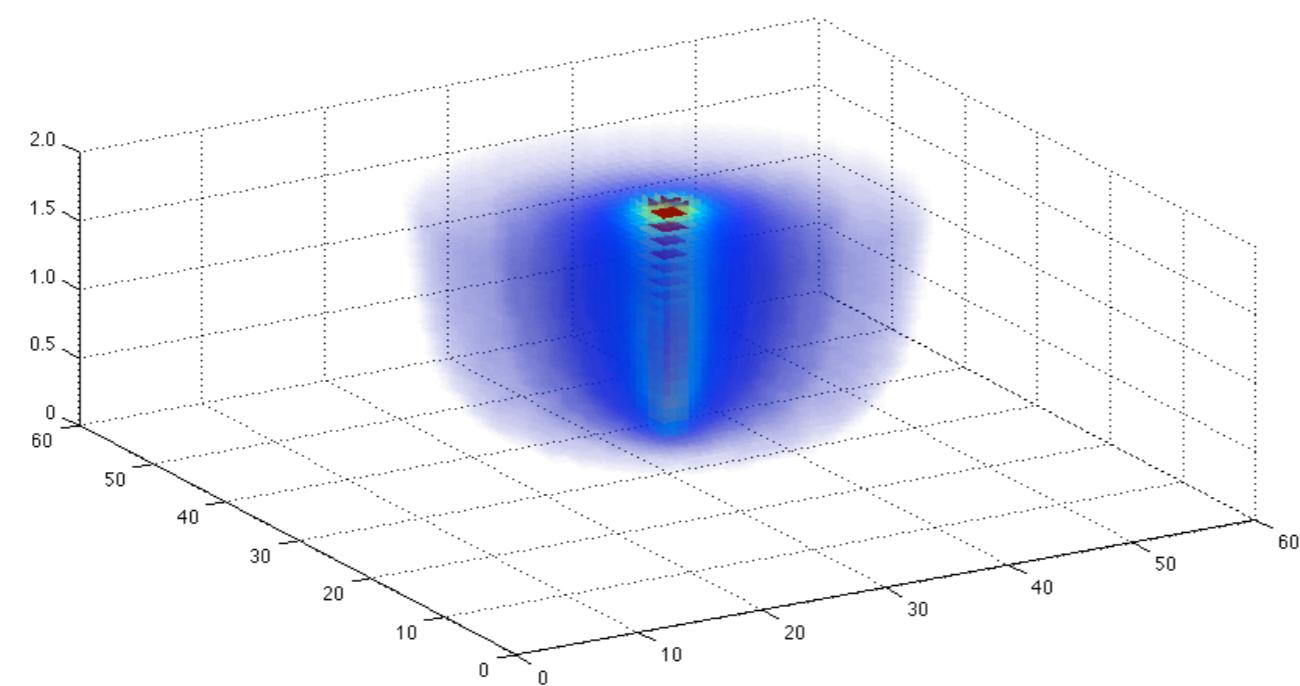


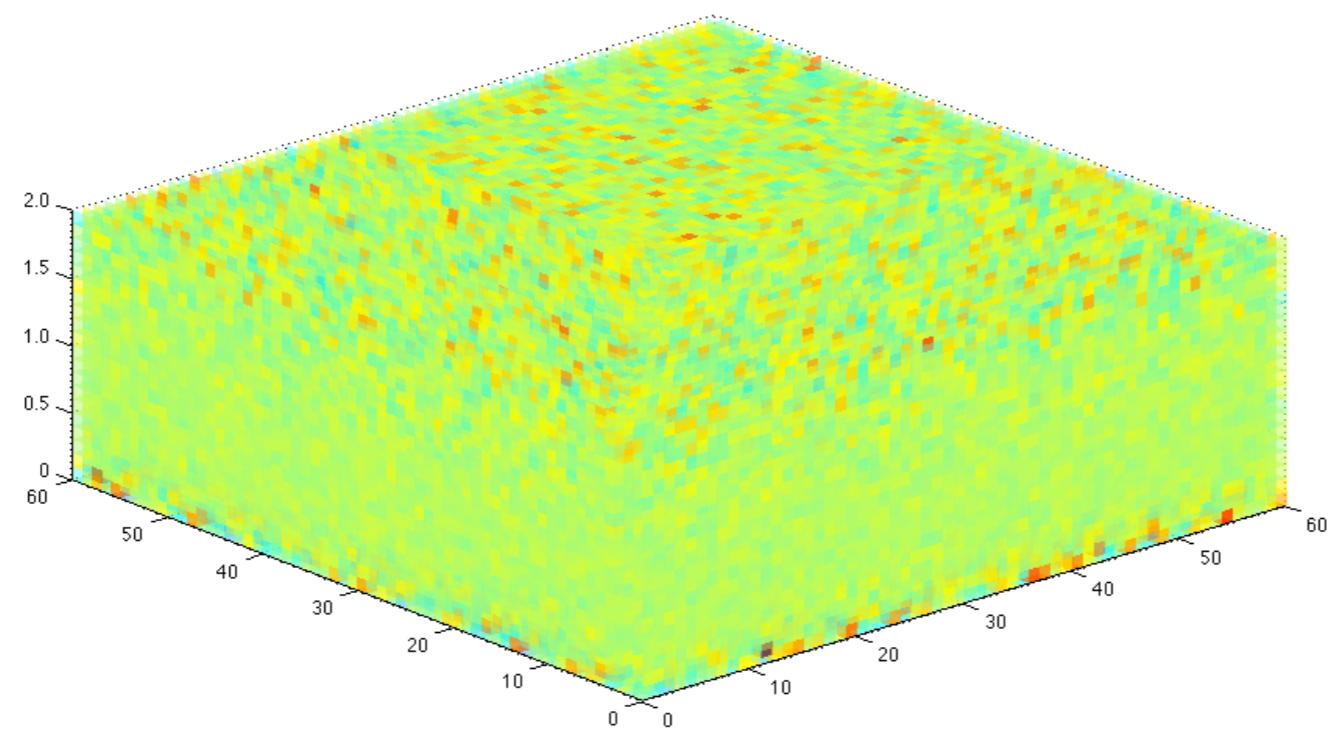
GRAVITATIONAL LENS

3D Weak Lensing









$$Y = HX + N$$

PB 1: find X knowing Y, H and the statistical properties of the noise N

Ex: Astronomical image deconvolution

Weak lensing

PB 2: find X and H knowing Y and the statistical properties of the noise N

Ex: Blind deconvolution

Ill posed problem, i.e. not an unique and stable solution ==> Regularization

$$\|Y - HX\|^2 \quad \text{with some constraints on } X$$

- Part I: Introduction to Inverse Problems in Astrophysics
- **Part II: Sparsity & Dictionaries**
- Part III: Sparse Regularization
- Part IV: Unmixing
- Part V: Sparsity for Planck and Euclid Space Missions

What is Sparsity?

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

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

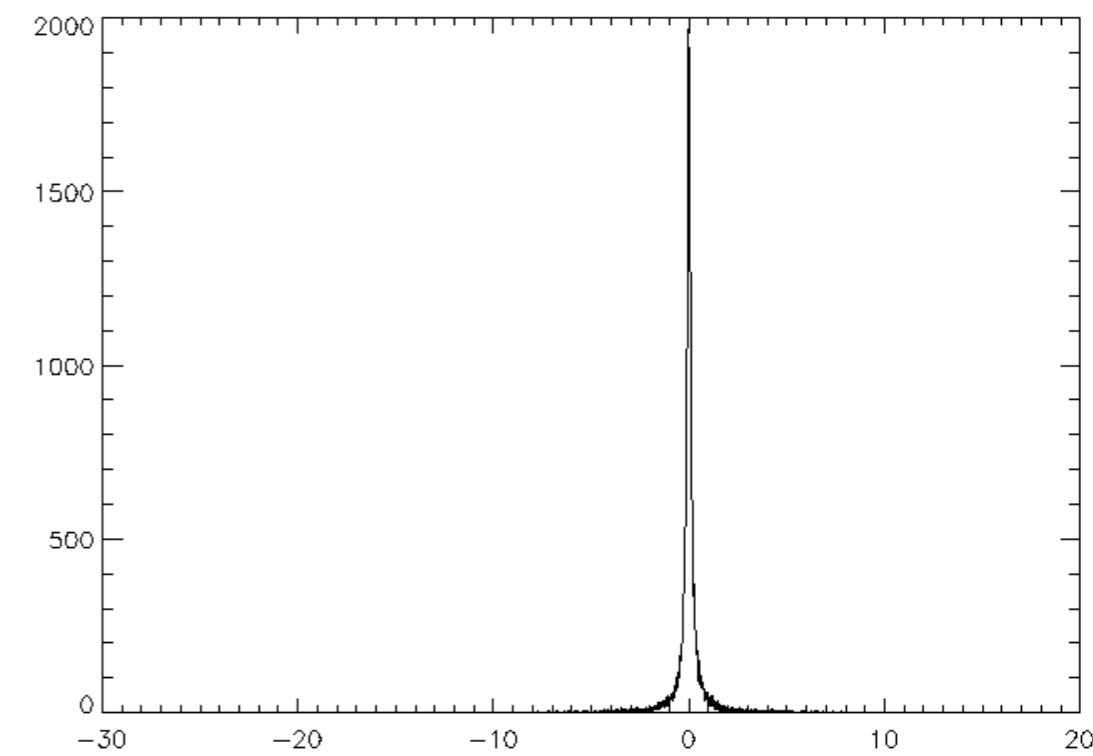
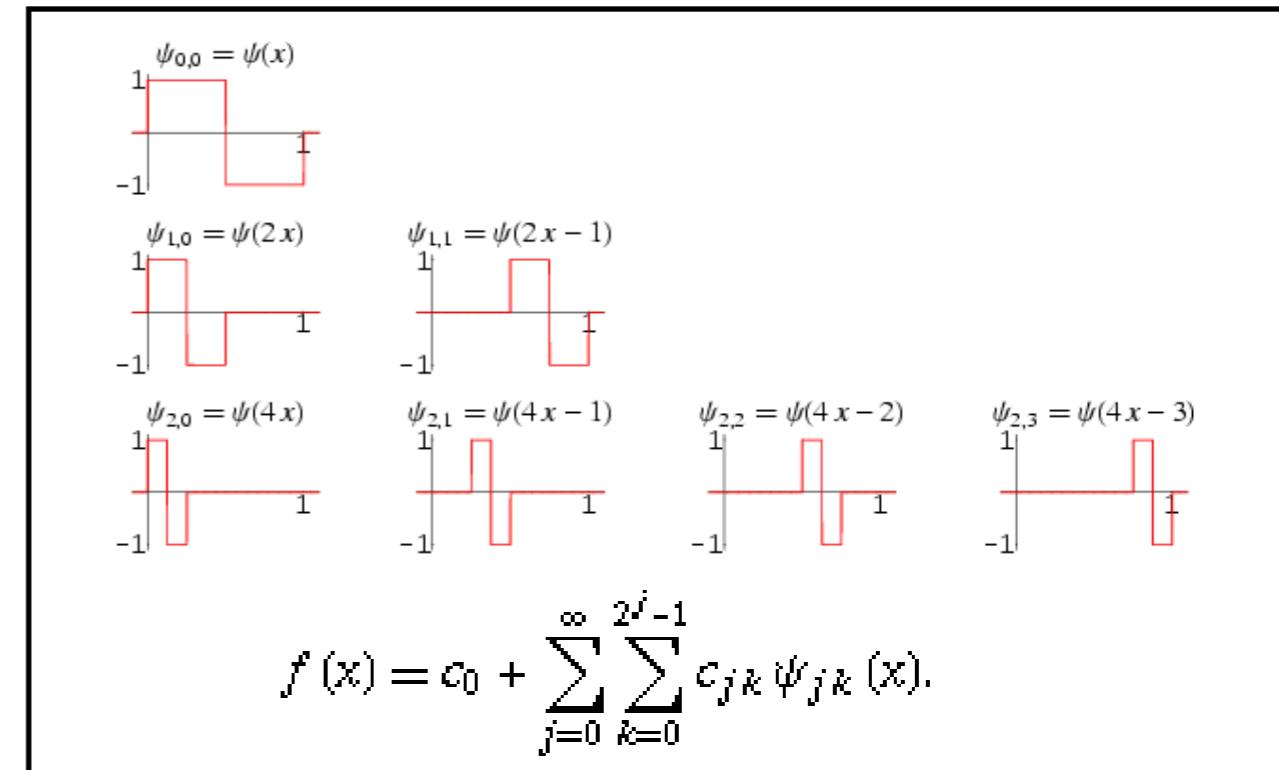
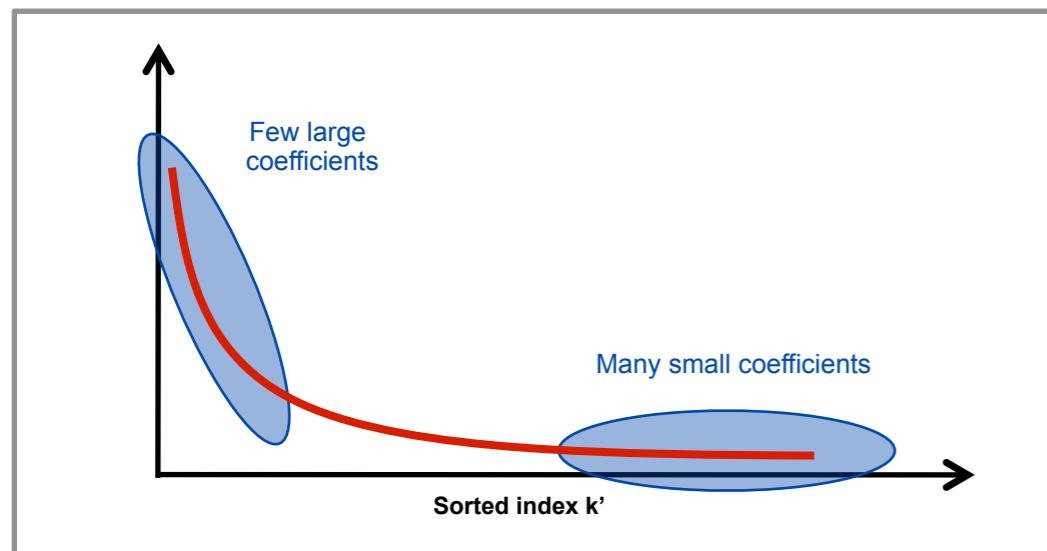
Dictionary
(basis, frame)

K

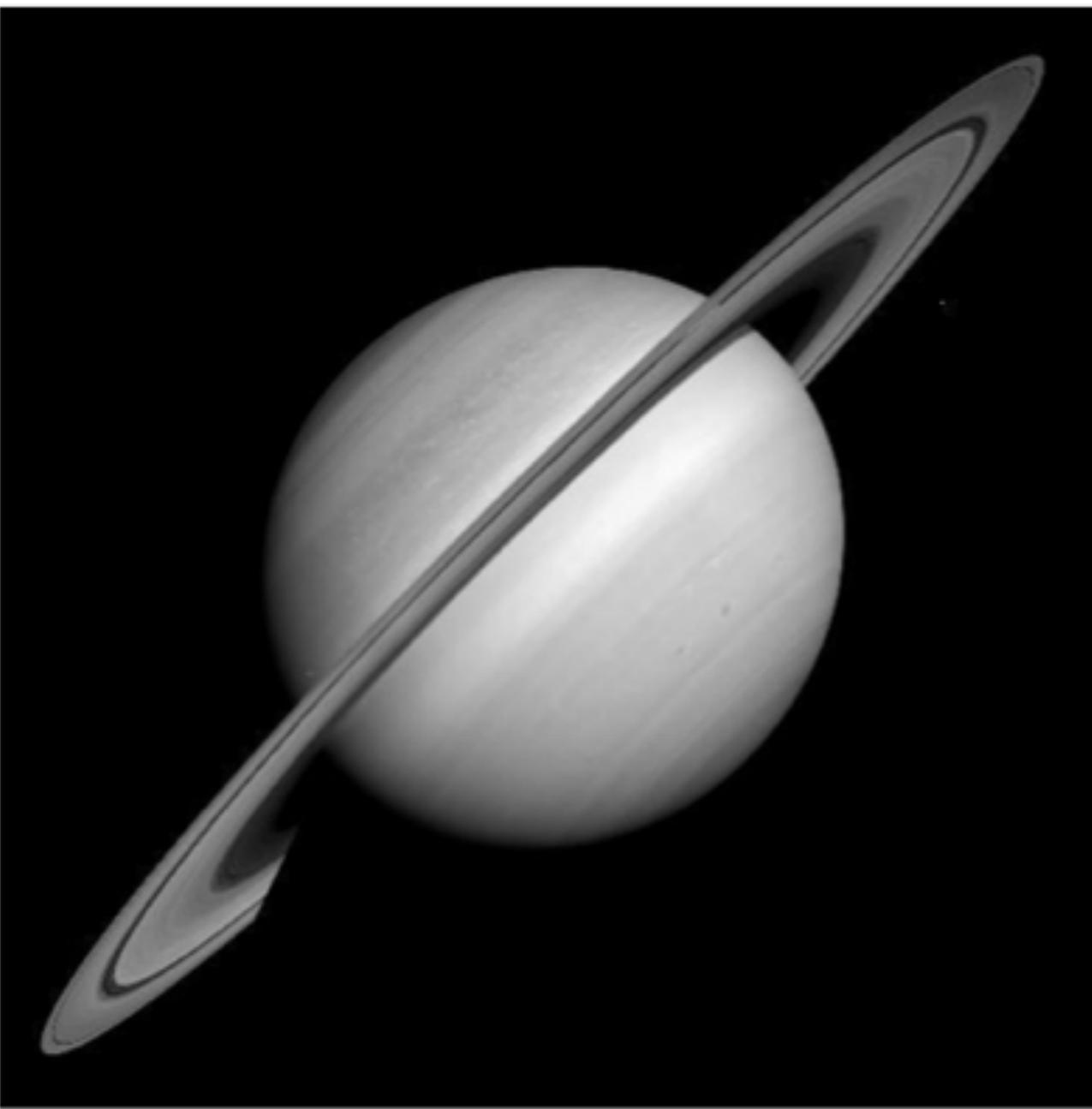
Atoms

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

coefficients



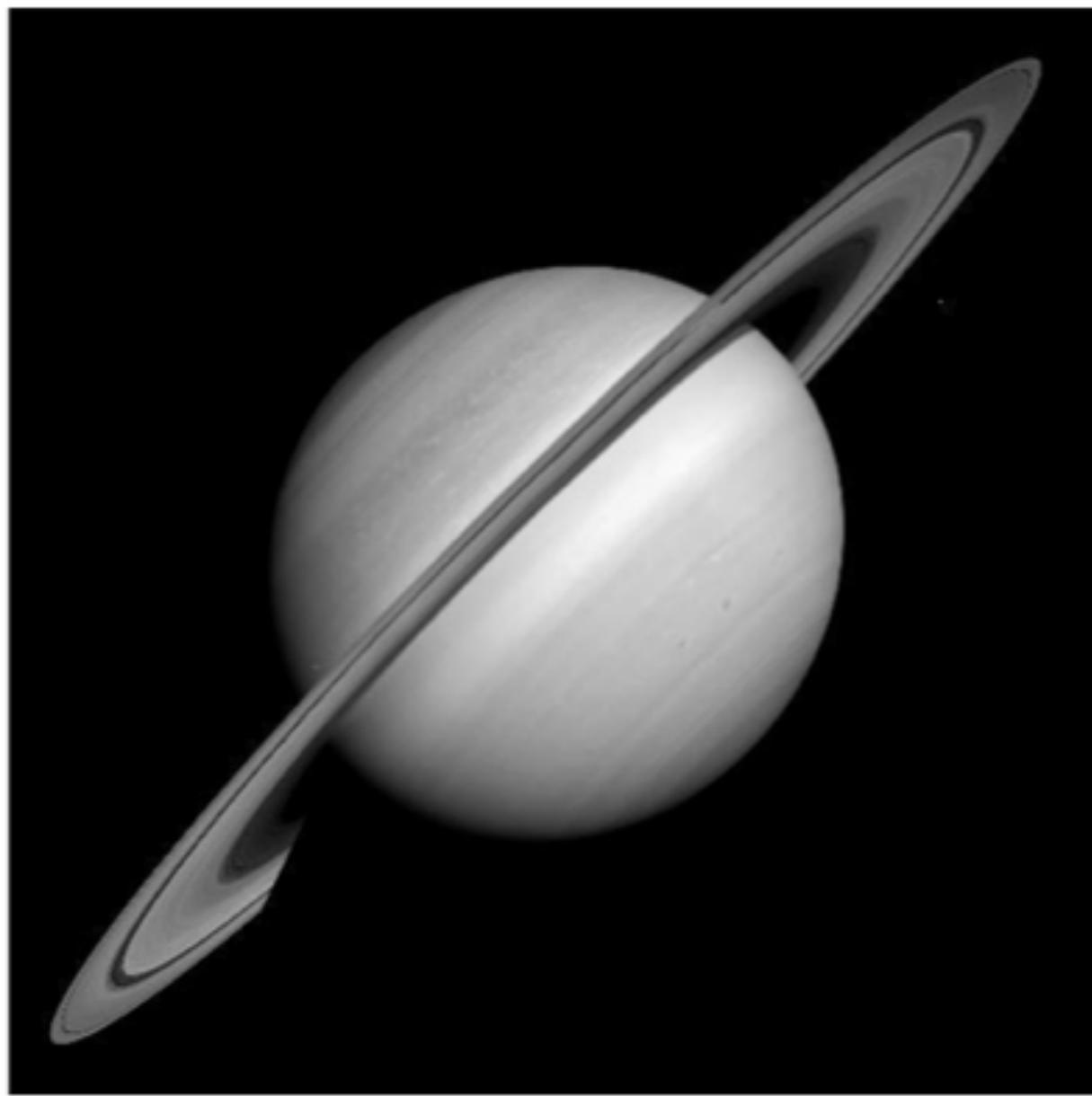
- Fast calculation of the coefficients
- Analyze the signal through the statistical properties of the coefficients
- Approximation theory uses the sparsity of the coefficients



The top 1% of the coefficients concentrate only 8.66% of the energy.
Not sparse...



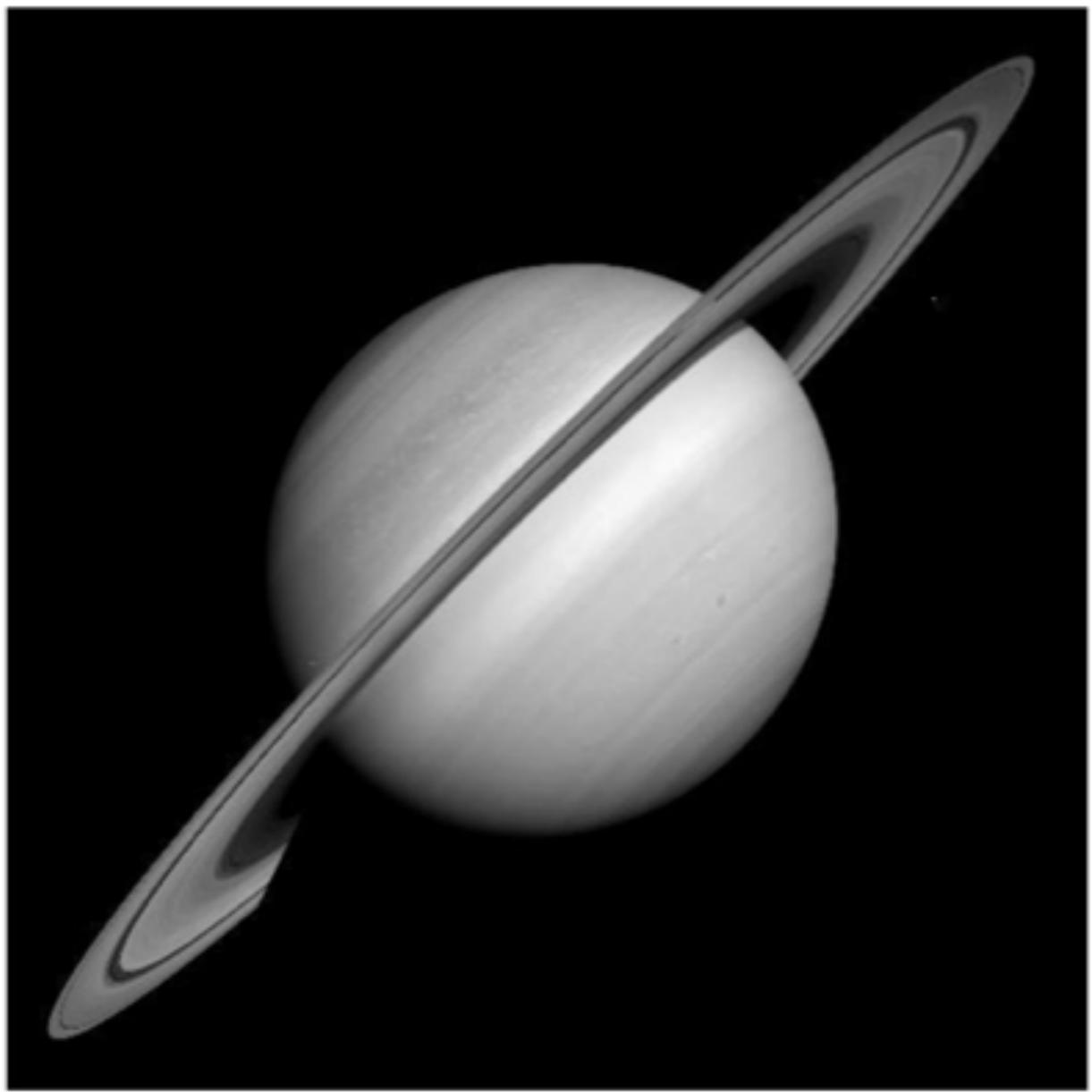
1% largest coefficients in real space
(the others are set to 0)



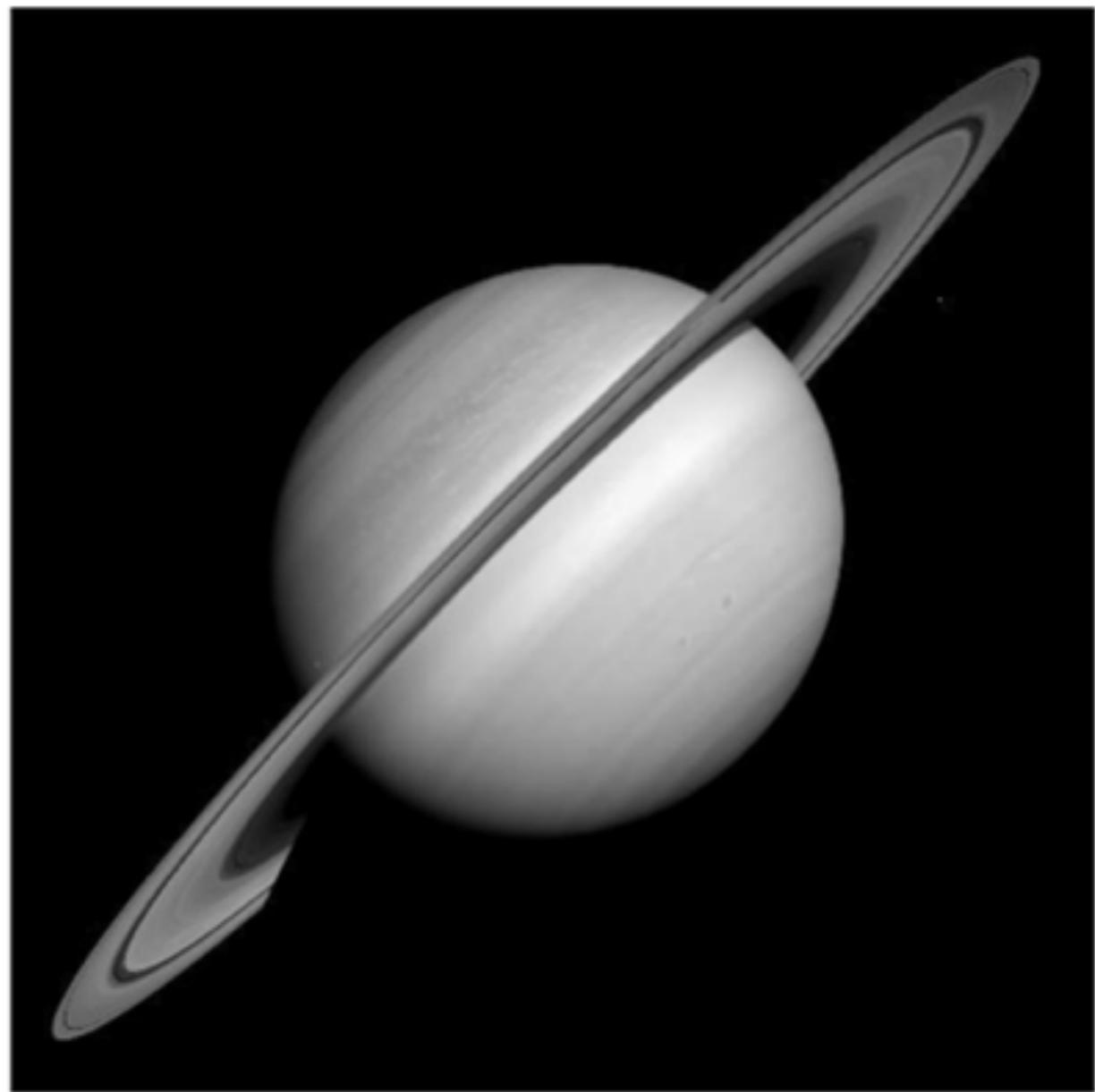
The wavelet
coefficients encode
edges and large scale
information.

1% largest coefficients in wavelet space
(the others are set to 0)
Wavelet transform



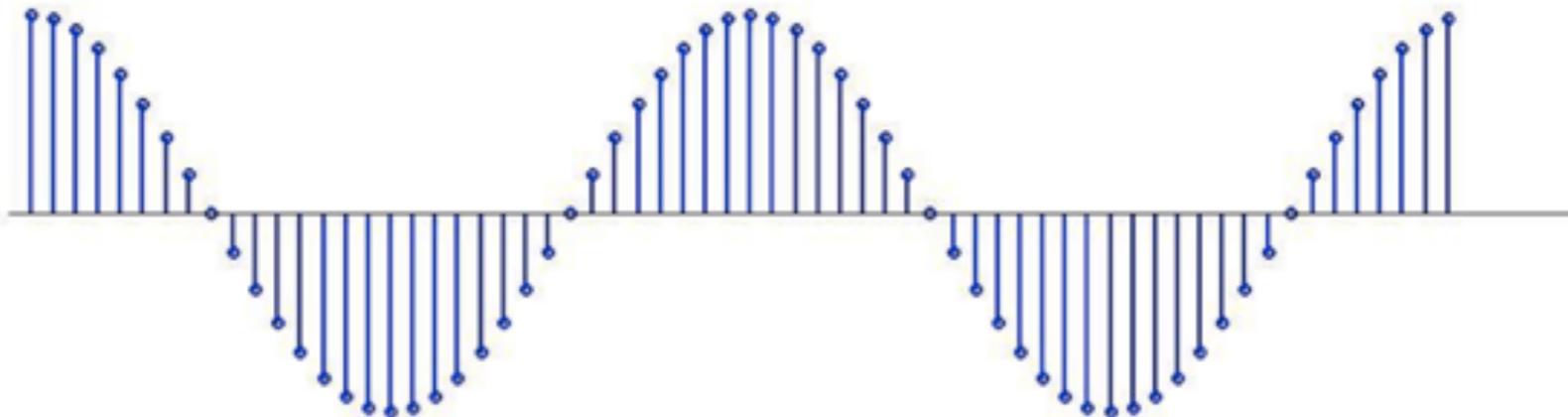


**1% of the wavelet coefficients
concentrate 99.96% of the energy:
This can be used as a *prior*.**



Reconstruction, after throwing away
99% of the wavelet coefficients

Strict Sparsity: k-sparse signals



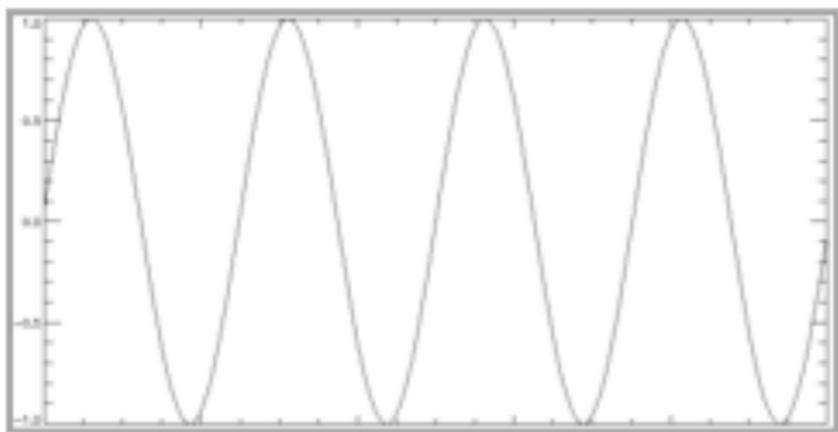
A sine wave in
real space...

...can be a Dirac
in Fourier space.

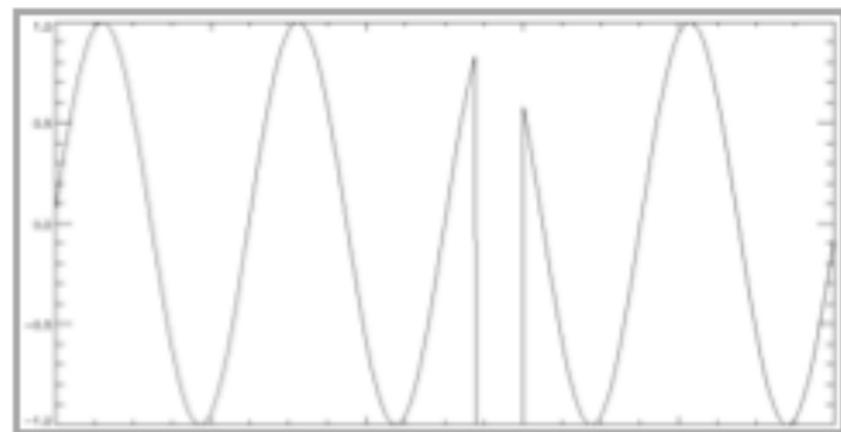


Sinusoids are
sparse in the
Fourier domain.

Minimizing the ℓ_0 norm



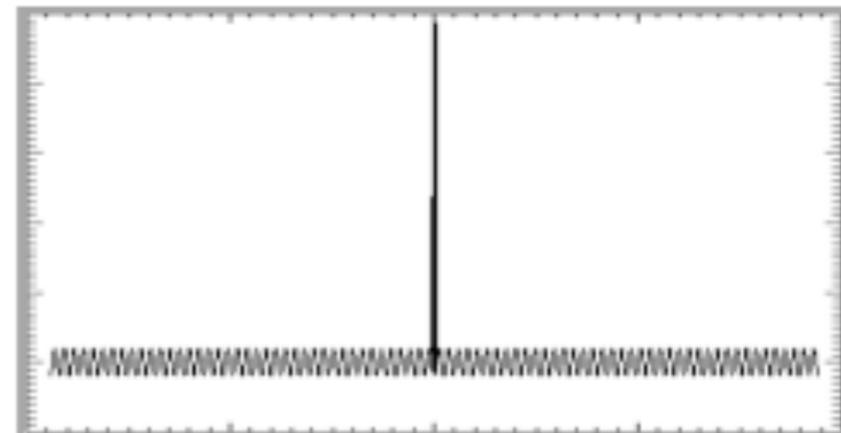
Sine curve



Truncated sine curve



TF of a sine curve



□ TF of a truncated sine curve □

with $0^0 = 0$, $\| \alpha \|_0 = \sum_k \alpha_k^0 = \# \{ \alpha_k \neq 0 \}$

How to measure sparsity ?

with $0^0 = 0$, $\|\alpha\|_0 = \sum_k \alpha_k^0 = \#\{\alpha_k \neq 0\}$

Formally, the sparsest coefficients are obtained by solving the optimization problem:

$$(P0) \text{ Minimize } \|\alpha\|_0 \text{ subject to } S = \phi\alpha$$

It has been proposed (*to relax and*) to replace the l_0 norm by the l_1 norm (Chen, 1995):

$$(P1) \text{ Minimize } \|\alpha\|_1 \text{ subject to } S = \phi\alpha$$

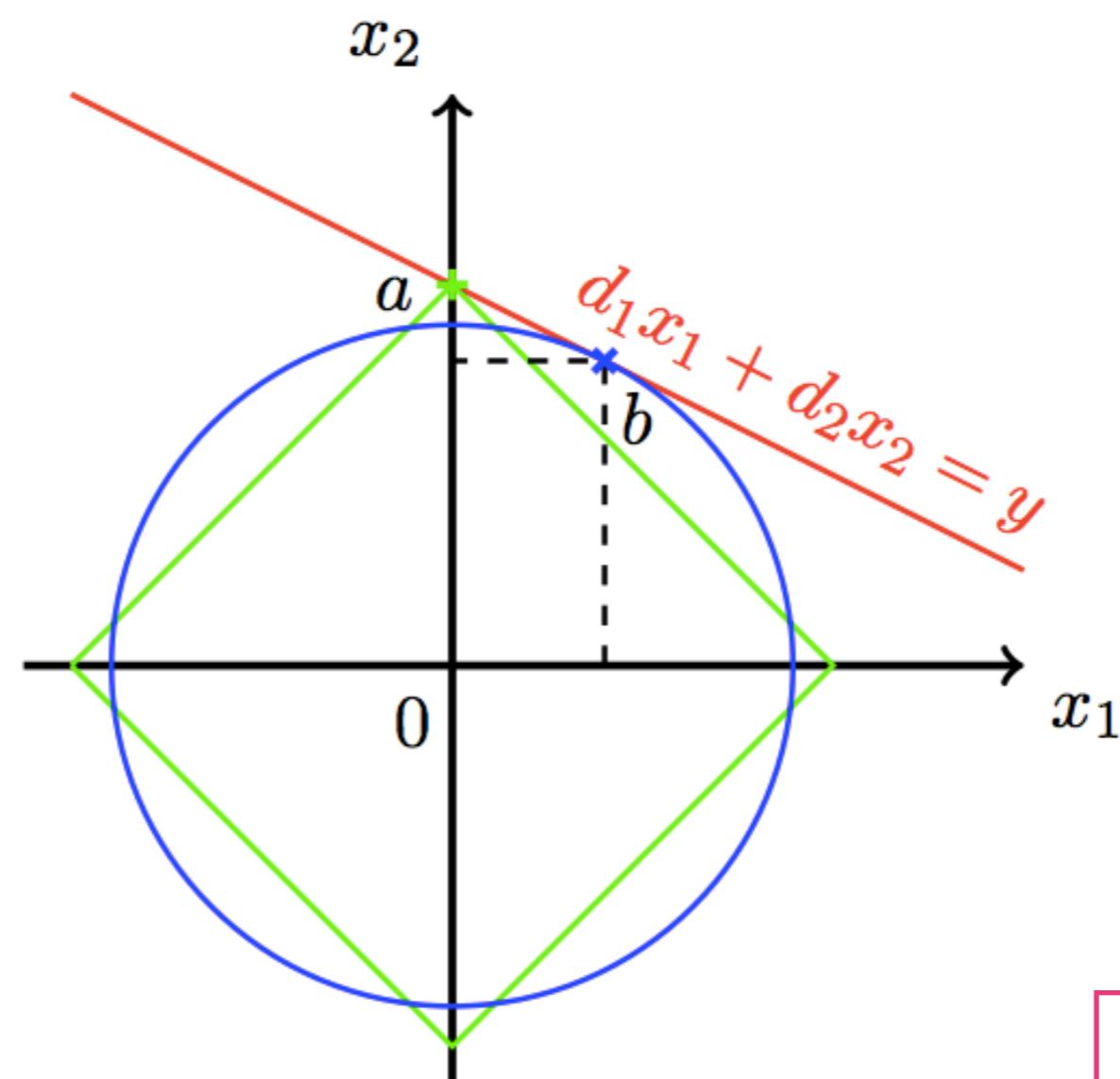
It can be seen as a kind of convexification of (P0).

It has been shown (Donoho and Huo, 1999) that for certain dictionary, if there exists a highly sparse solution to (P0), then it is identical to the solution of (P1).

==> Link the sparsity and the sampling through the Compressed Sensing.

L1 Norm & Sparsity

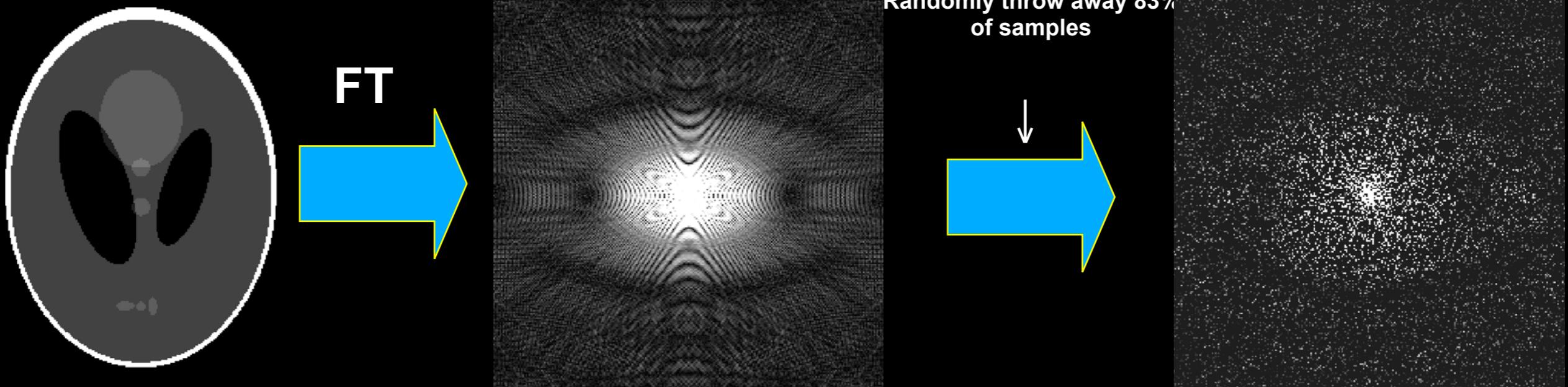
$$\|X\|_p = \left(\sum_i |X_i|^p \right)^{\frac{1}{p}}$$



$$p < 2$$

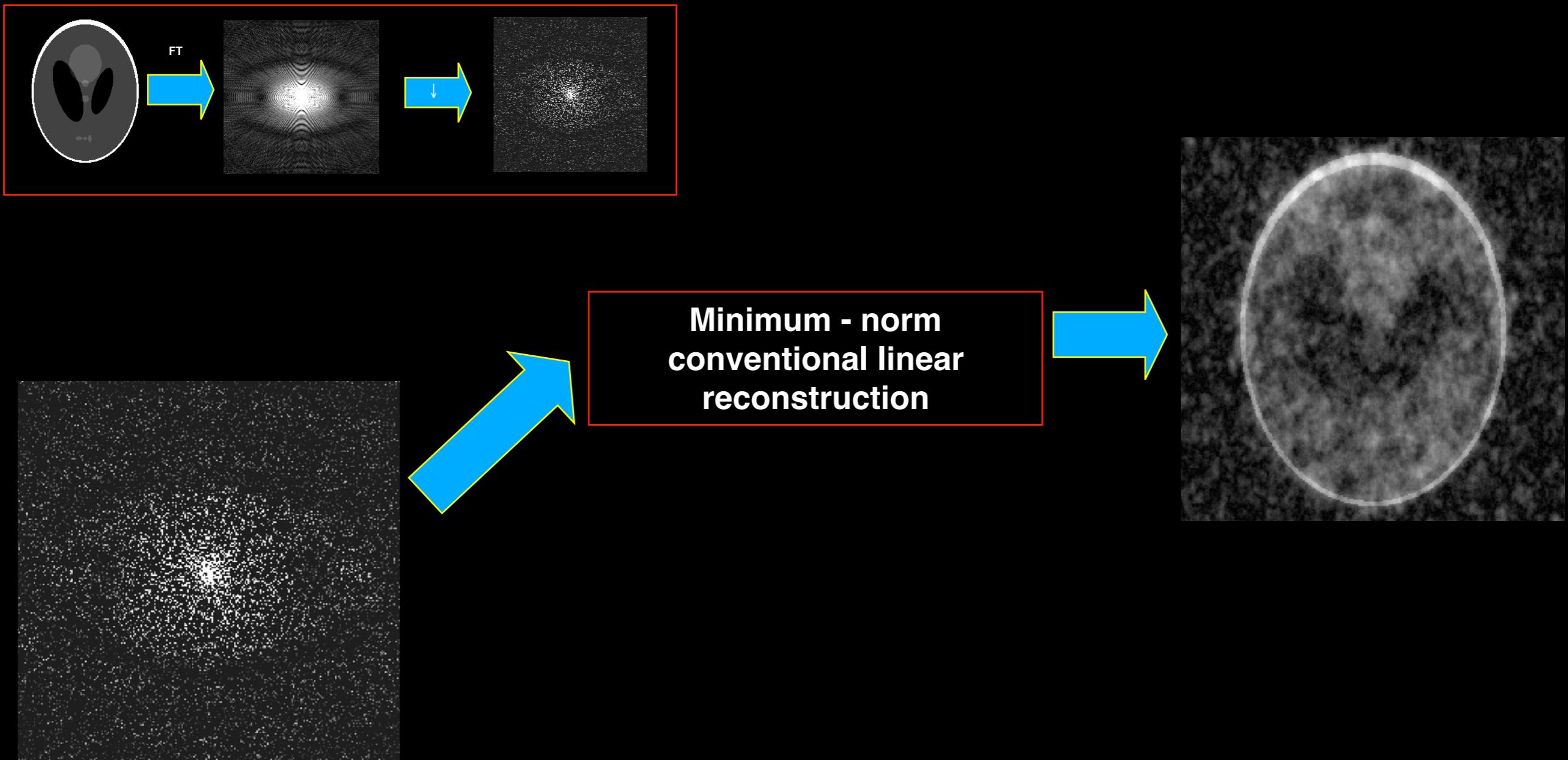
Sparsity & Sampling

A Surprising Experiment*



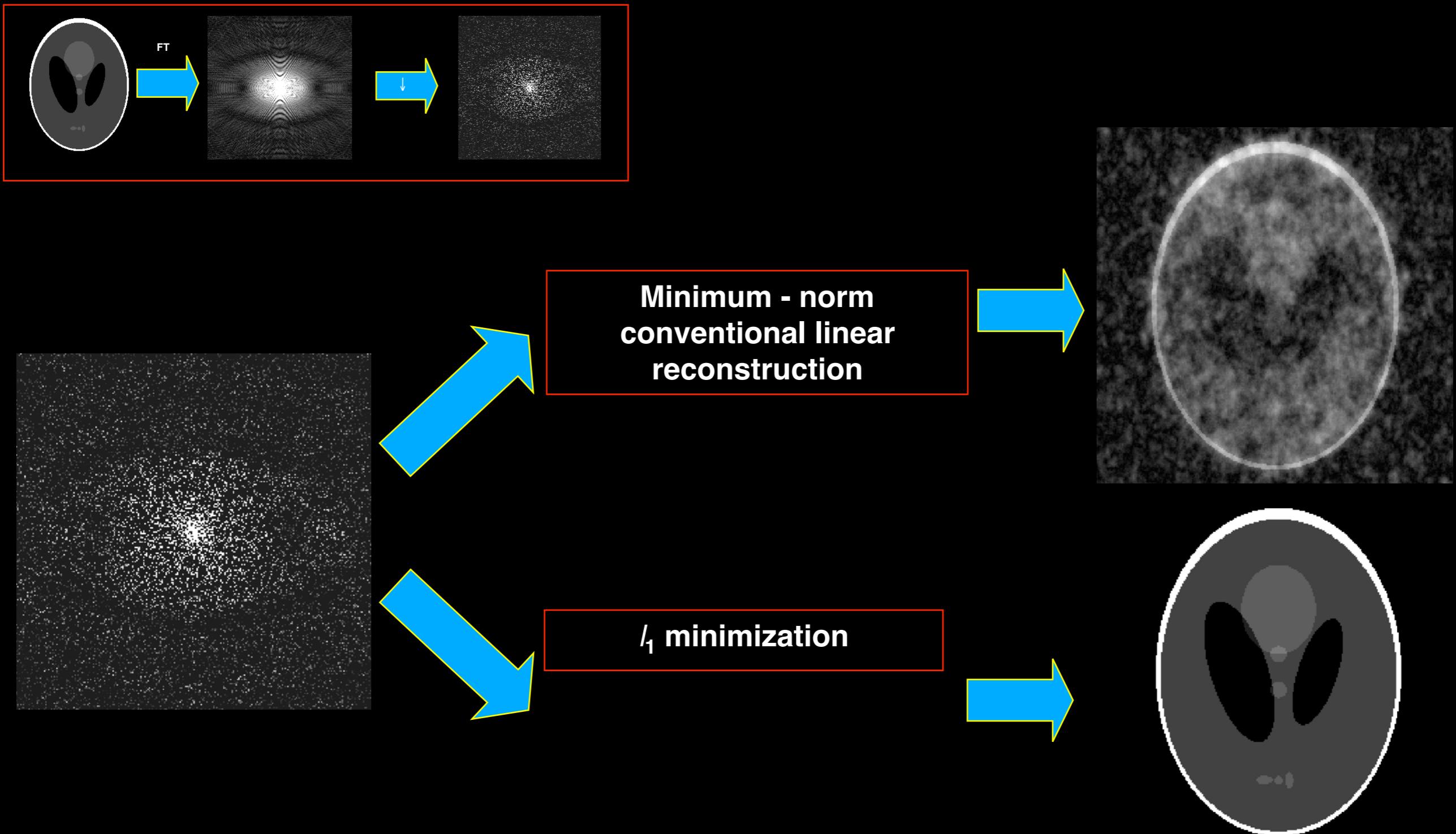
* E.J. Candes, J. Romberg and T. Tao.

A Surprising Result*



* E.J. Candes, J. Romberg and T. Tao.

A Surprising Result*



E.J. Candes



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 $\sim K \log N$ incoherent measurements”

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

$$y = \Theta x$$

$M \times 1$ measurements $M \times N$ $N \times 1$ sparse signal

$K < M \ll N$

K nonzero entries

y

Θ

x

$M \times 1$
measurements

$M \times N$

$N \times 1$
sparse
signal

K
nonzero
entries

$y = \Theta x$

Reconstruction via non linear processing:

$$\min_x \|x\|_1 \text{ s.t. } y = \Theta x$$

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

Compressed Sensing Reconstruction

Measurements: $y_k = \langle x, \theta_k \rangle$

Reconstruction via non linear processing: $\min_x \|x\|_1 \text{ s.t. } y = \Theta_\Lambda x$

In practice, x is sparse in a given **dictionary**:

$$x = \Phi \alpha$$

and we need to solve:

$$\min_{\alpha} \|\alpha\|_1 \text{ s.t. } y = \Theta_\Lambda \Phi \alpha$$

The mutual incoherence is defined as

$$\mu_{\Theta, \Phi} = \sqrt{N} \max_{i,k} |\langle \phi_i, \theta_k \rangle|$$

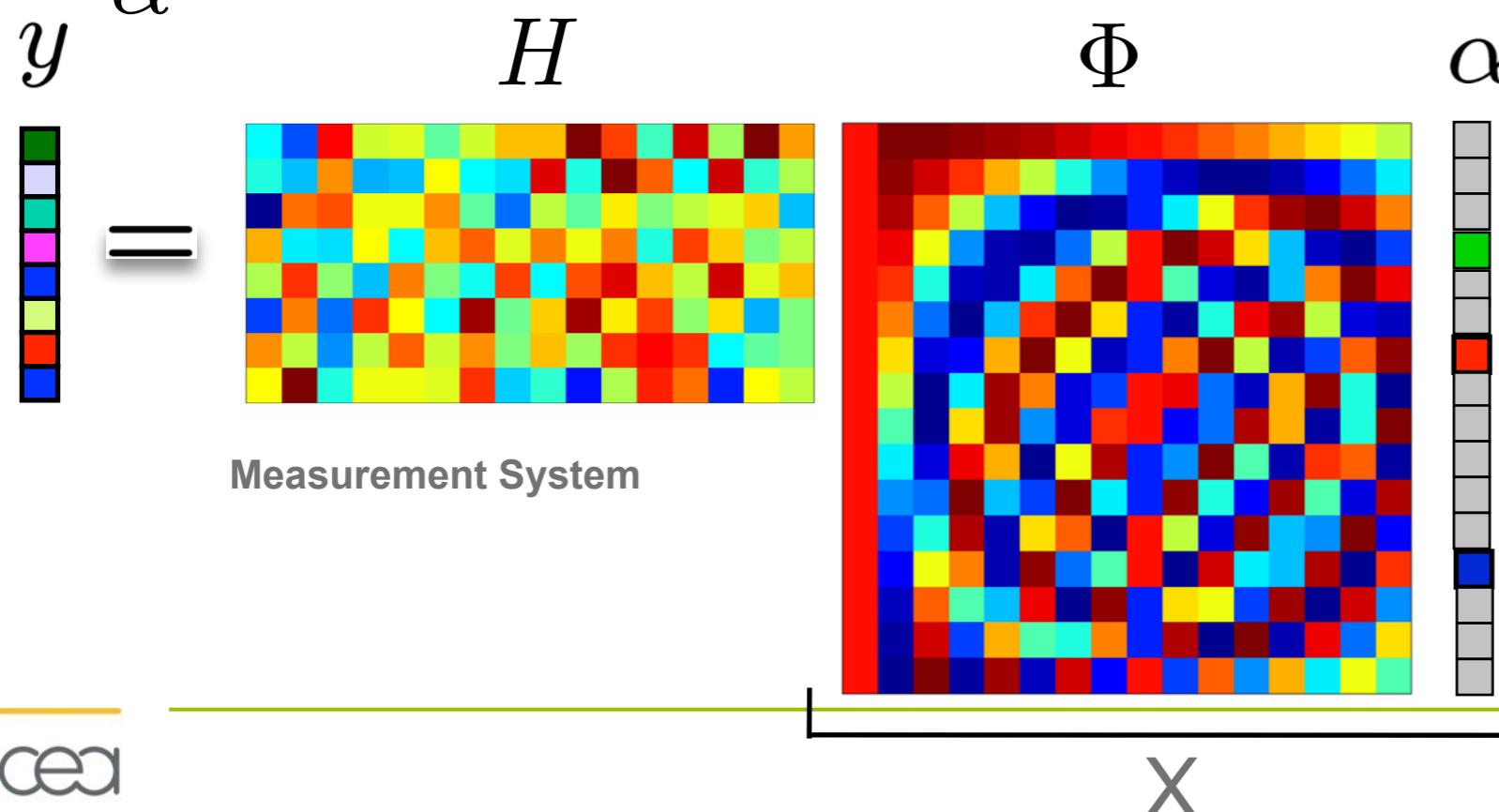
the number of required measurements is :

$$m \geq C \mu_{\Theta, \Phi}^2 K \log n$$

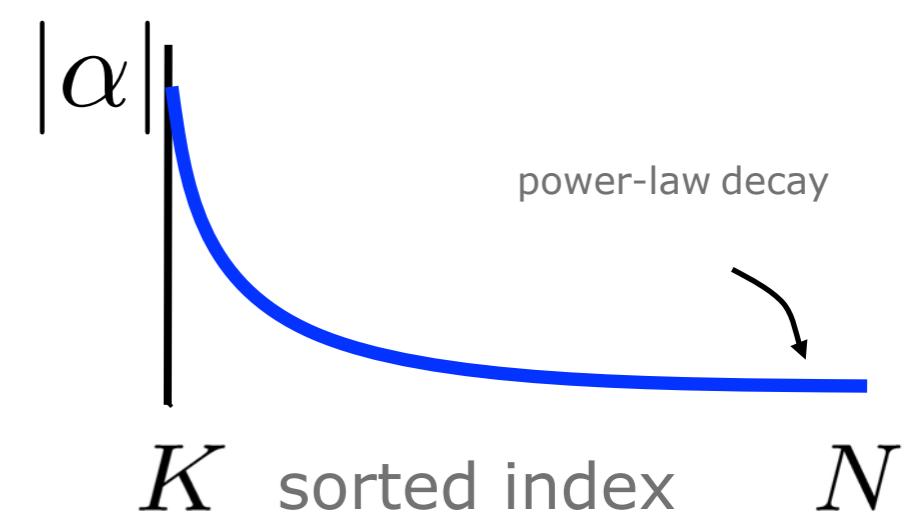
$$Y = HX + N$$

$X = \Phi\alpha$, and α is sparse

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



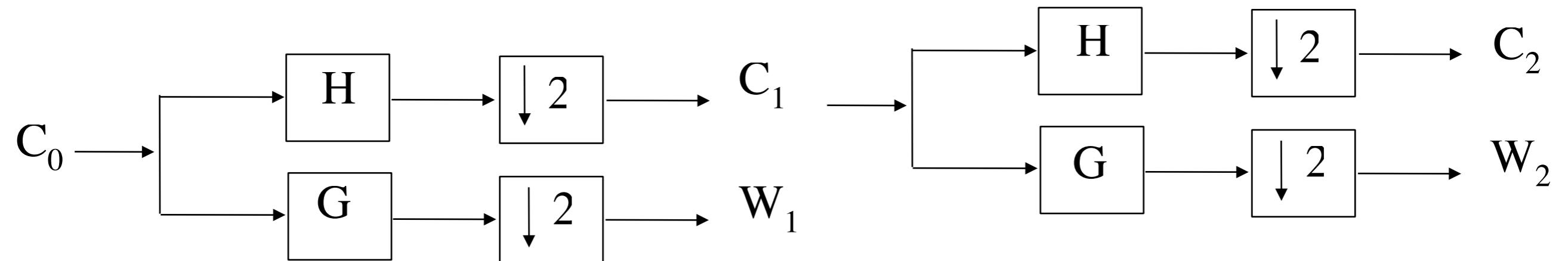
- Denoising
- Deconvolution
- Component Separation
- Inpainting
- Blind Source Separation
- Minimization algorithms
- Compressed Sensing



The Orthogonal Wavelet Transform (OWT)

$$s_l = \sum_k c_{J,k} \phi_{J,l}(k) + \sum_k \sum_{j=1}^J \psi_{j,l}(k) w_{j,k}$$

Transformation



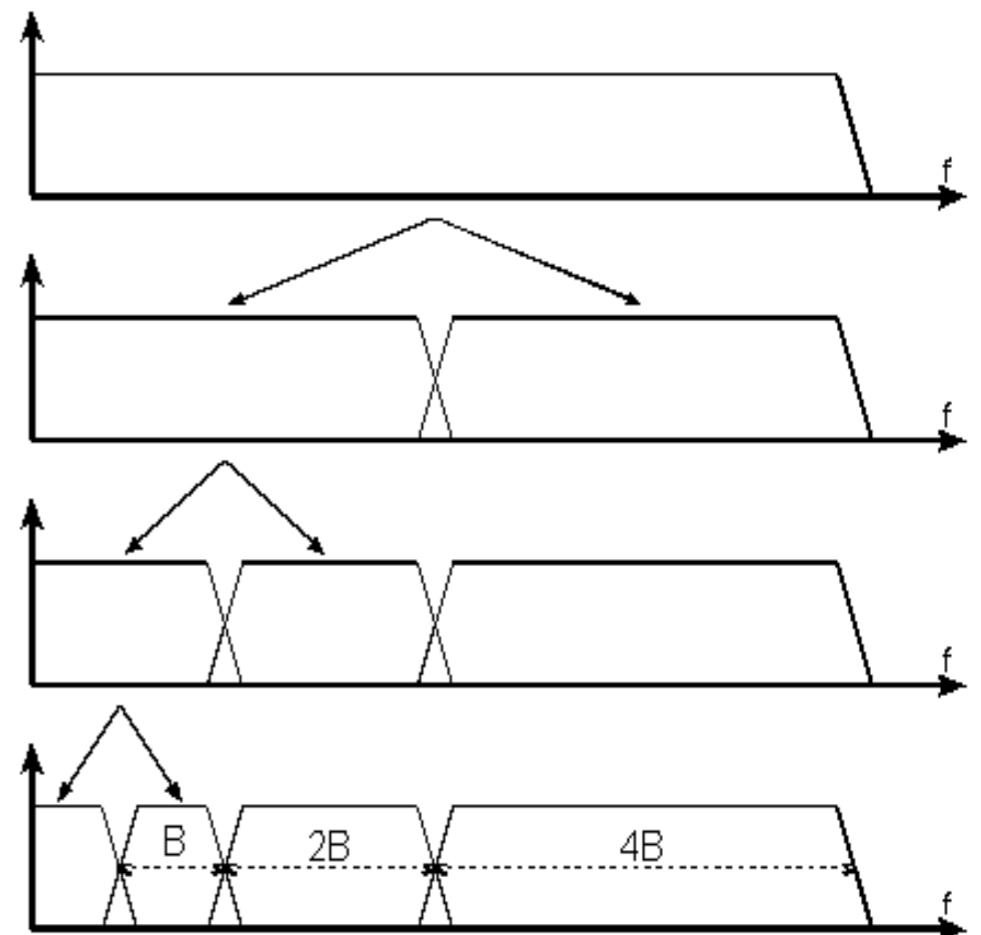
$$c_{j+1,l} = \sum_h h_{k-2l} c_{j,k} = (\bar{h} * c_j)_{2l}$$

$$w_{j+1,l} = \sum_h g_{k-2l} c_{j,k} = (\bar{g} * c_j)_{2l}$$

Reconstruction:

$$c_{j,l} = \sum_k \tilde{h}_{k+2l} c_{j+1,k} + \tilde{g}_{k+2l} w_{j+1,k} = \tilde{h} * c_{j+1} + \tilde{g} * w_{j+1}$$

$$x = (x_1, 0, x_2, 0, x_3, \dots, 0, x_j, 0, \dots, x_{n-1}, 0, x_n)$$



At two dimensions, we separate the variables x,y:

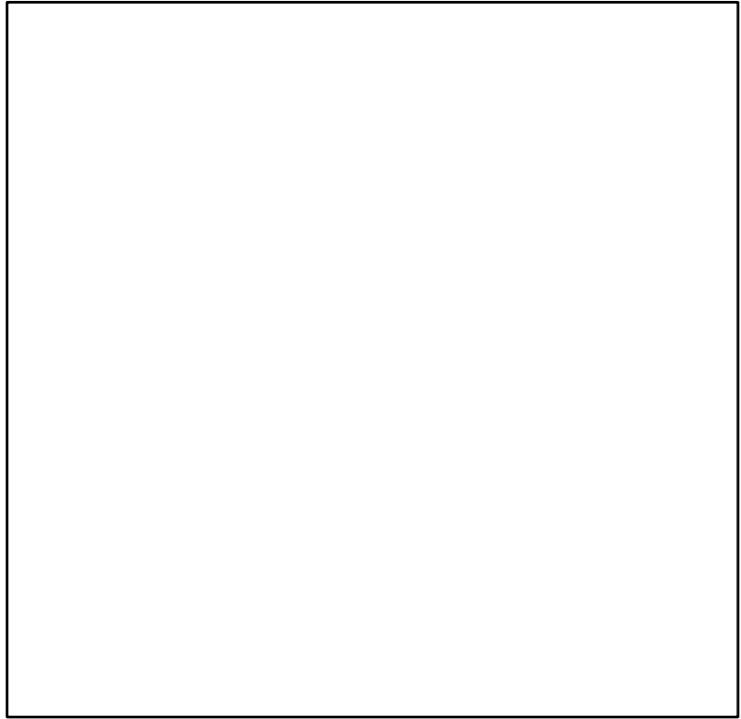
- vertical wavelet: $\psi^1(x, y) = \phi(x)\psi(y)$
- horizontal wavelet: $\psi^2(x, y) = \psi(x)\phi(y)$
- diagonal wavelet: $\psi^3(x, y) = \psi(x)\psi(y)$

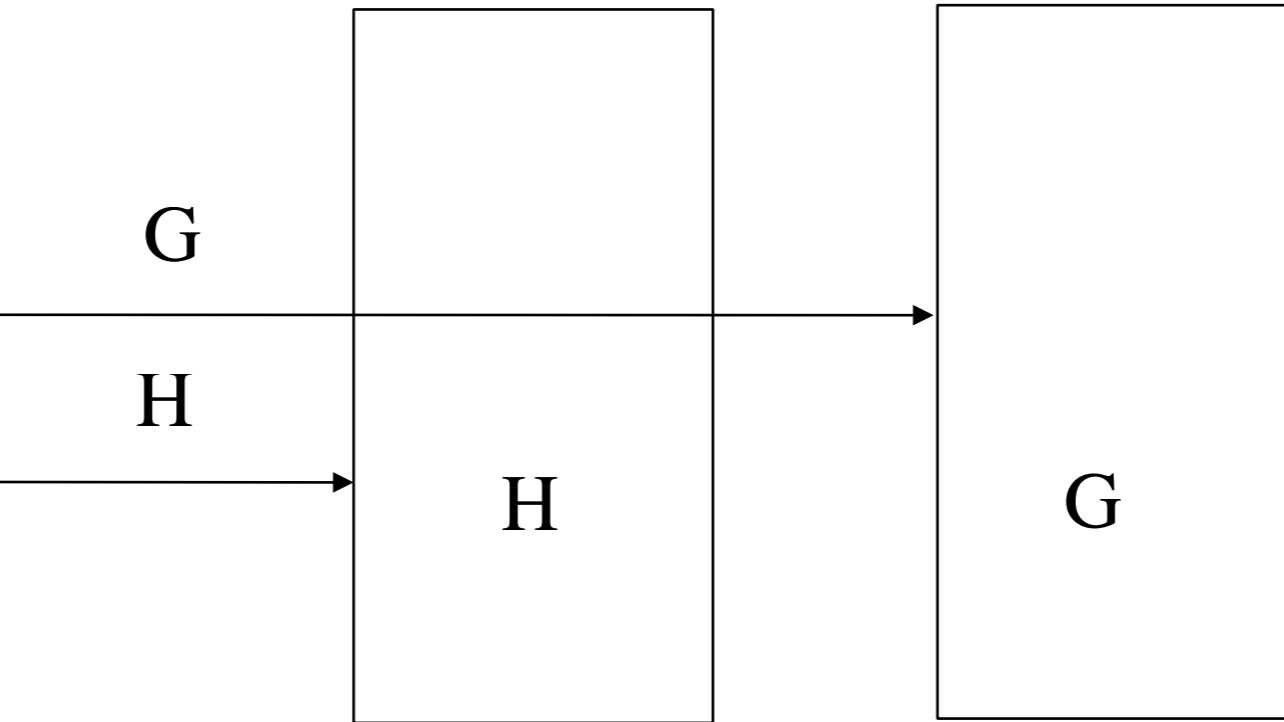
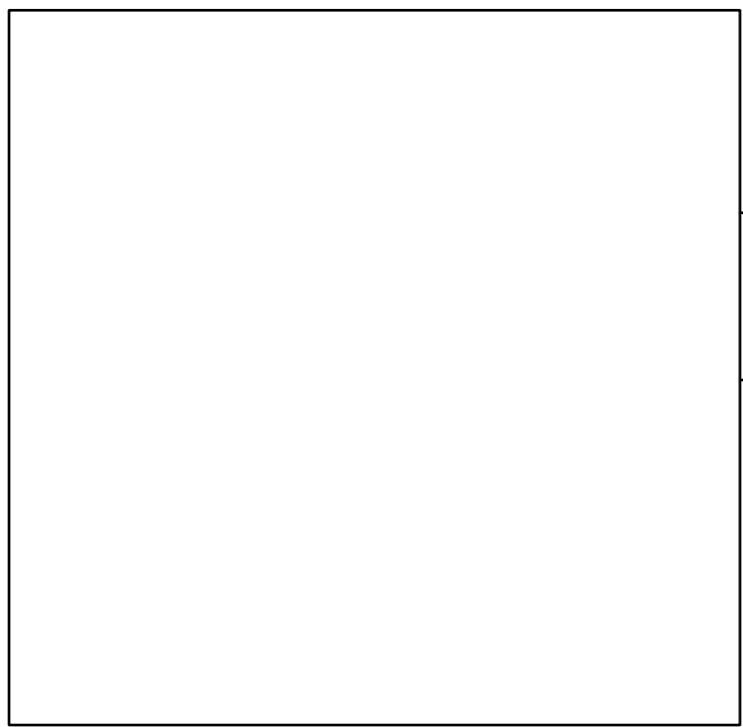
The detail signal is contained in three sub-images

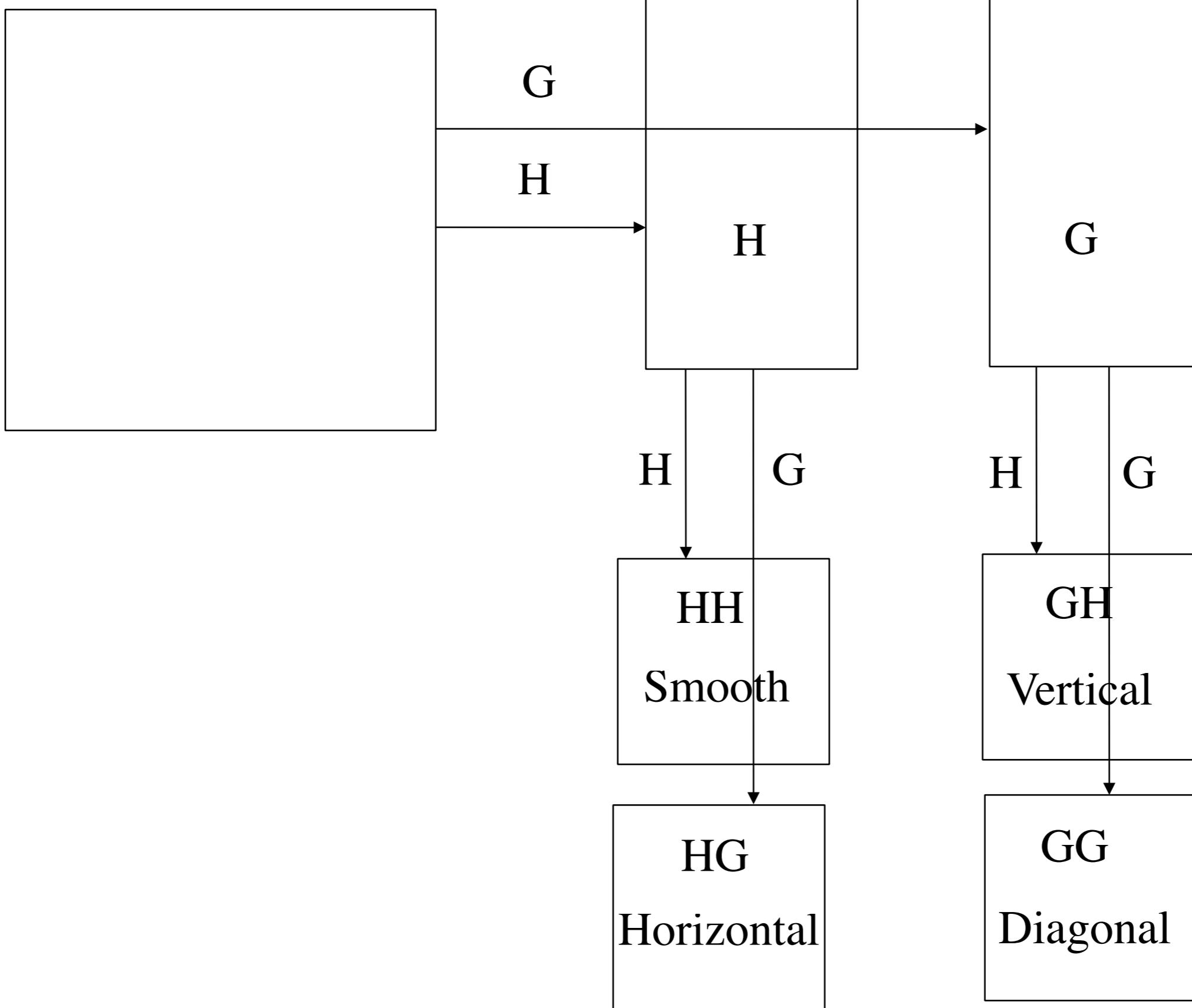
$$w_j^1(k_x, k_y) = \sum_{l_x=-\infty}^{+\infty} \sum_{l_y=-\infty}^{+\infty} g(l_x - 2k_x)h(l_y - 2k_y)c_{j+1}(l_x, l_y)$$

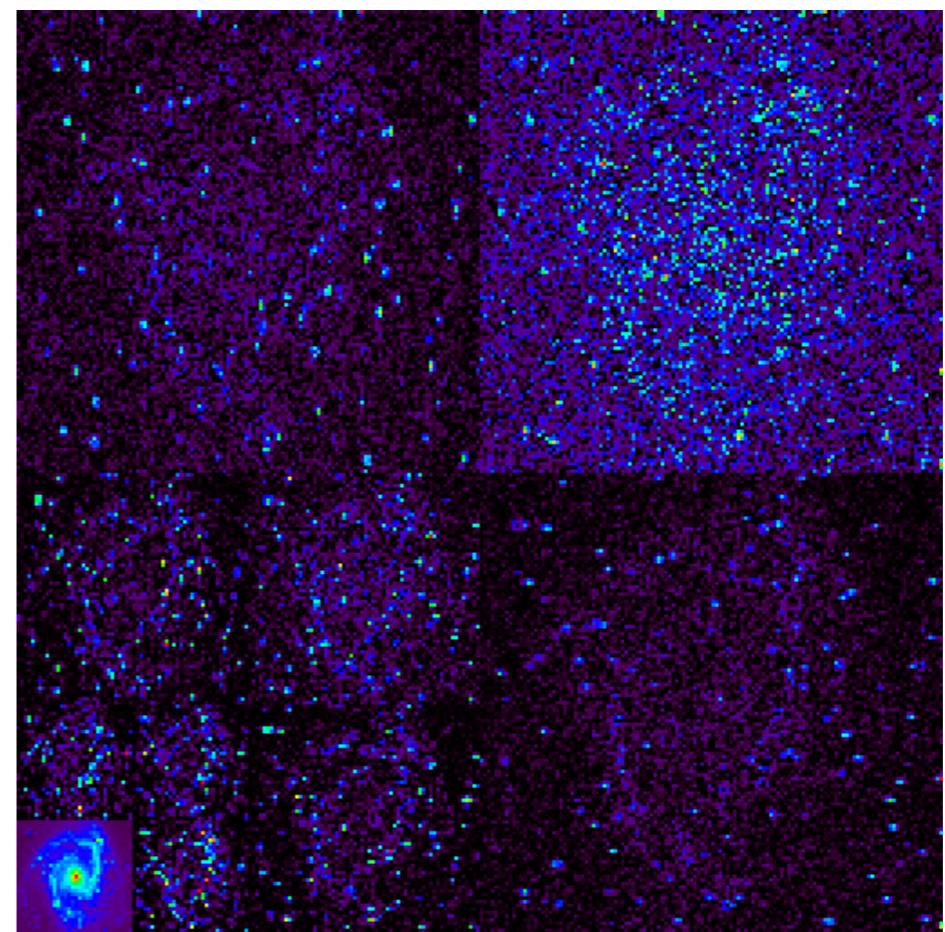
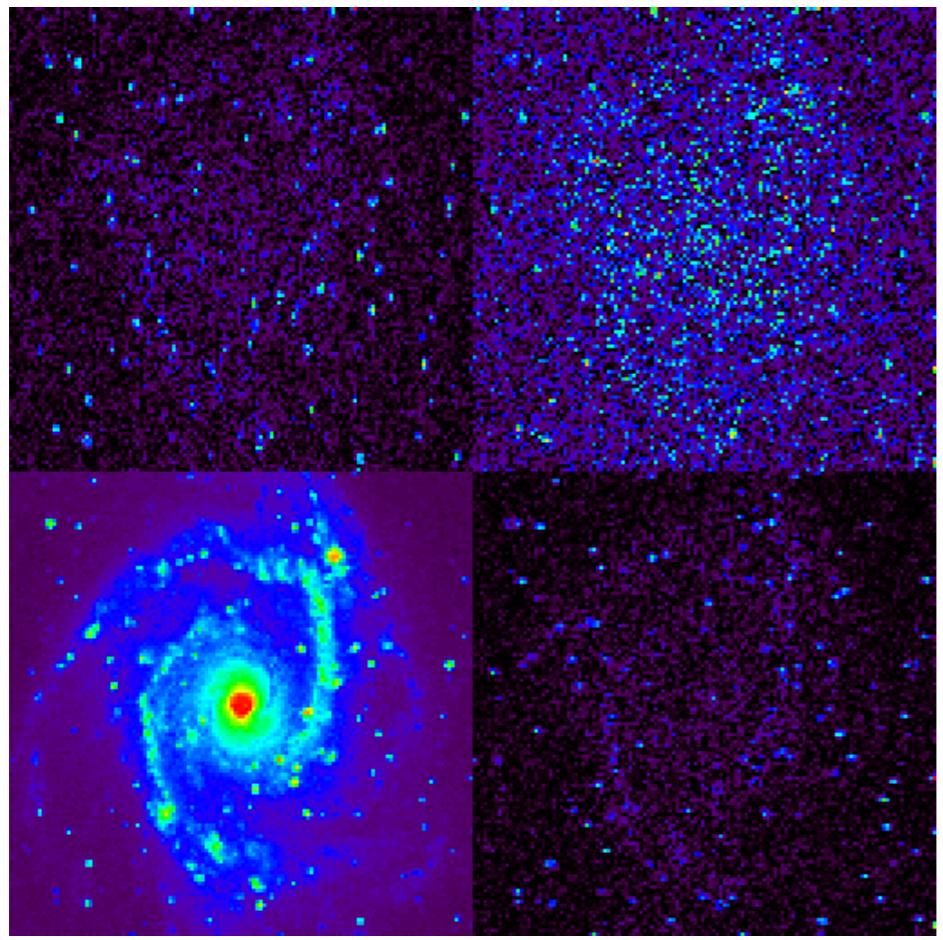
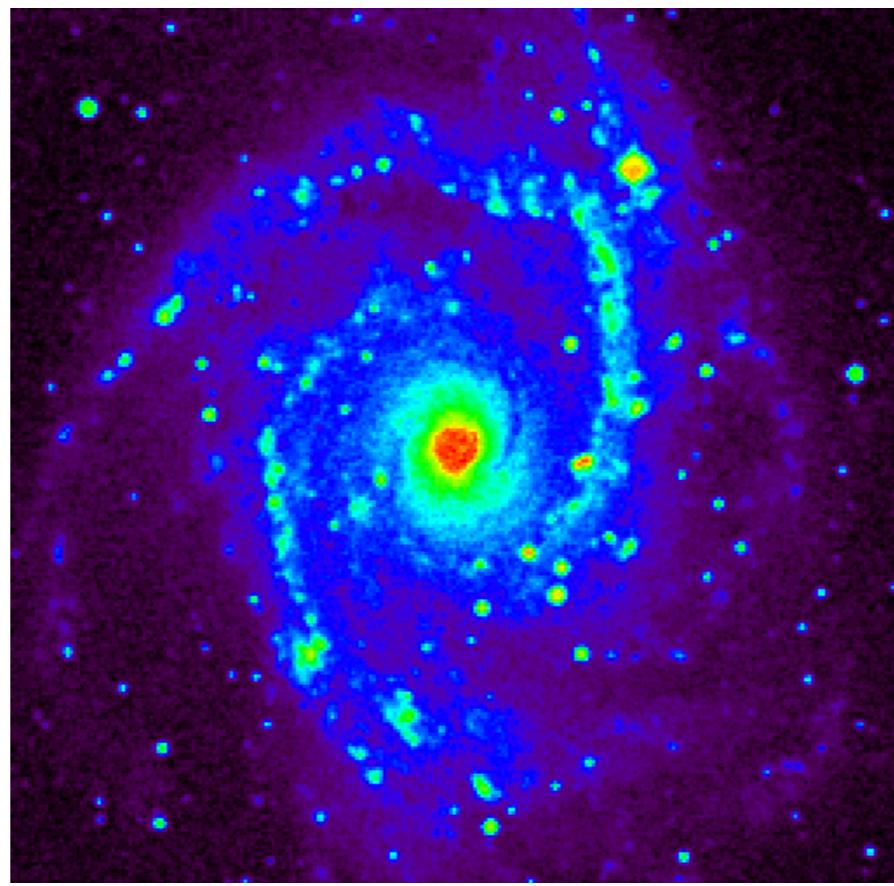
$$w_j^2(k_x, k_y) = \sum_{l_x=-\infty}^{+\infty} \sum_{l_y=-\infty}^{+\infty} h(l_x - 2k_x)g(l_y - 2k_y)c_{j+1}(l_x, l_y)$$

$$w_j^3(k_x, k_y) = \sum_{l_x=-\infty}^{+\infty} \sum_{l_y=-\infty}^{+\infty} g(l_x - 2k_x)g(l_y - 2k_y)c_{j+1}(l_x, l_y)$$



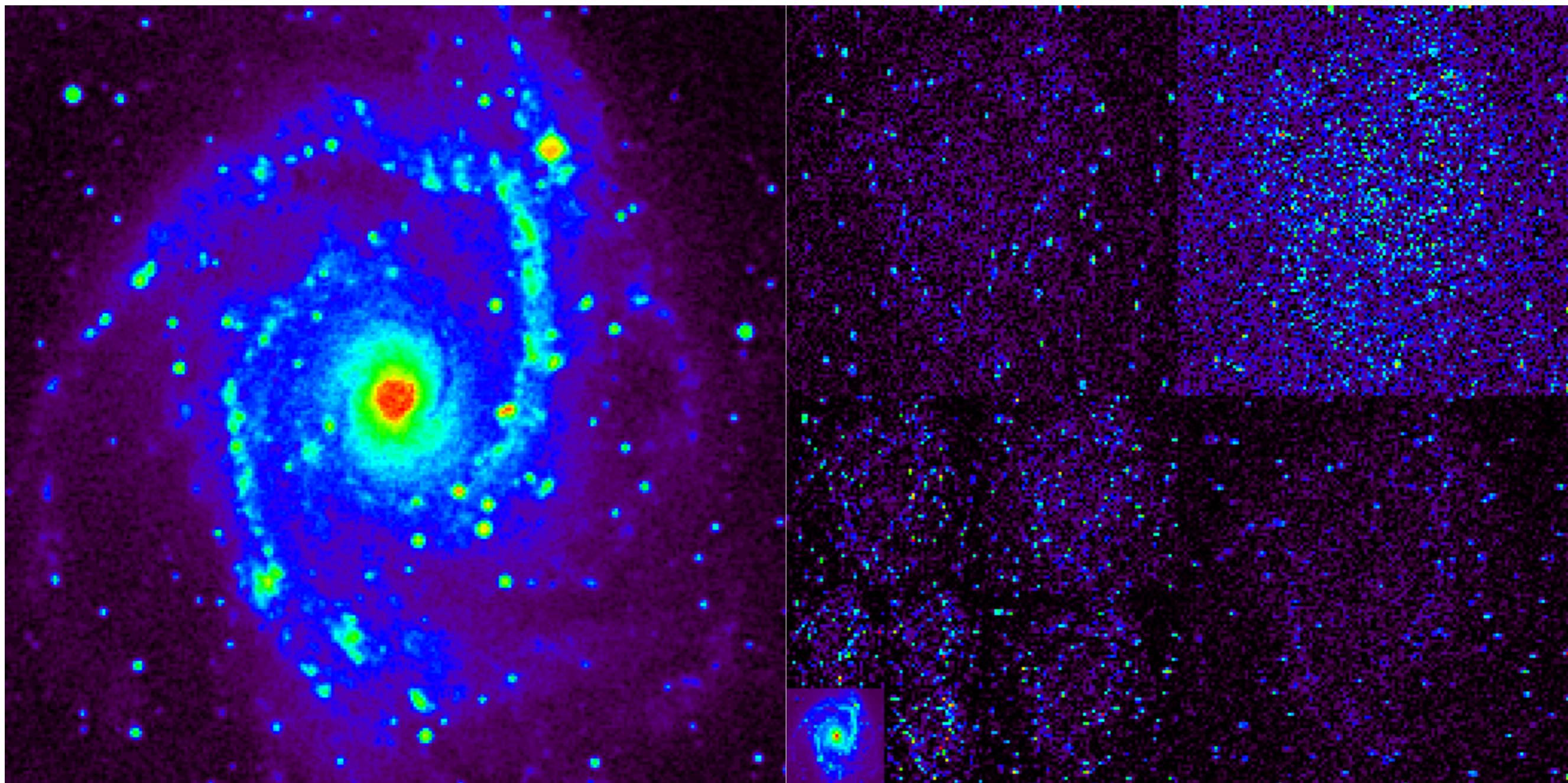


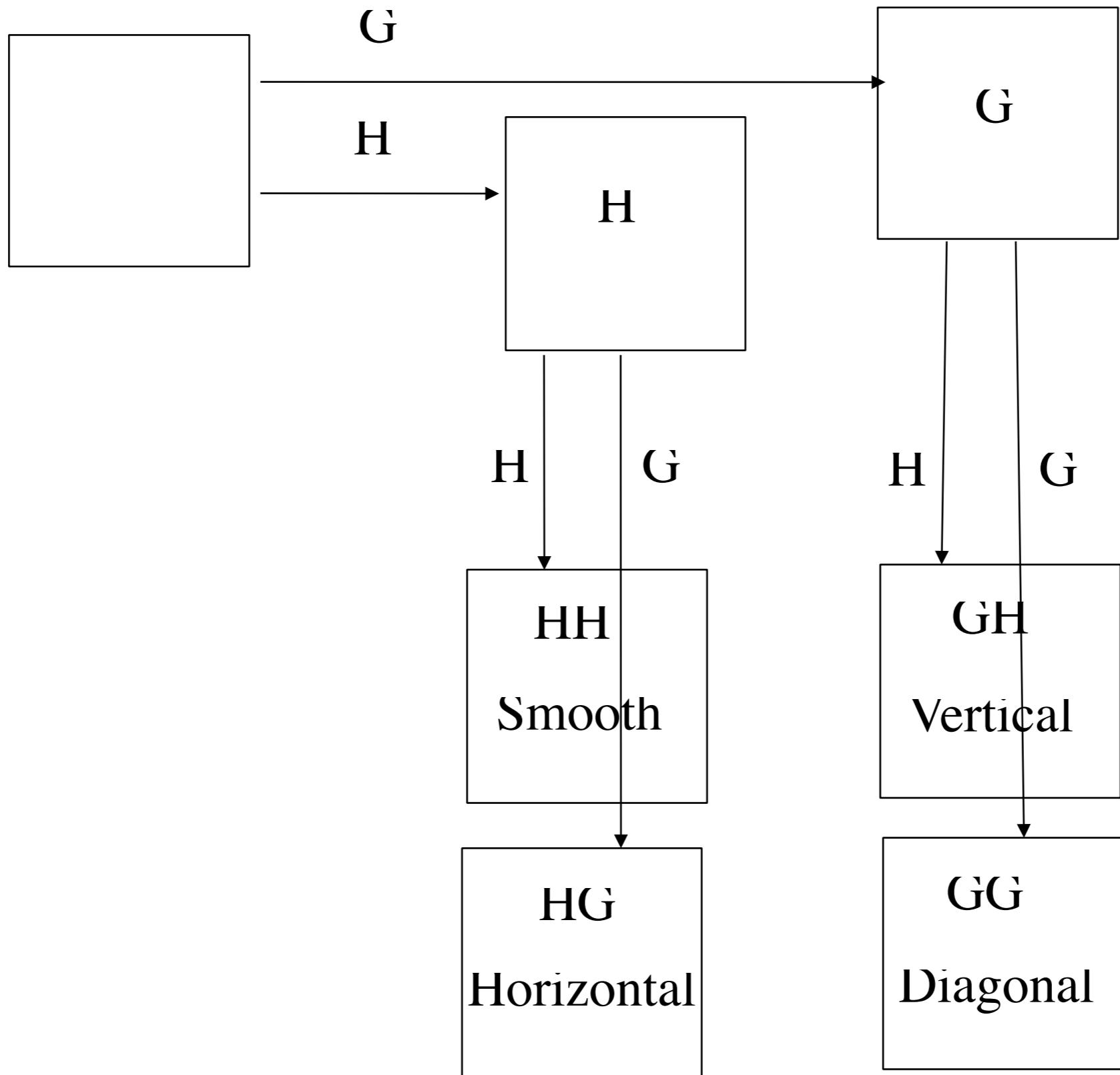




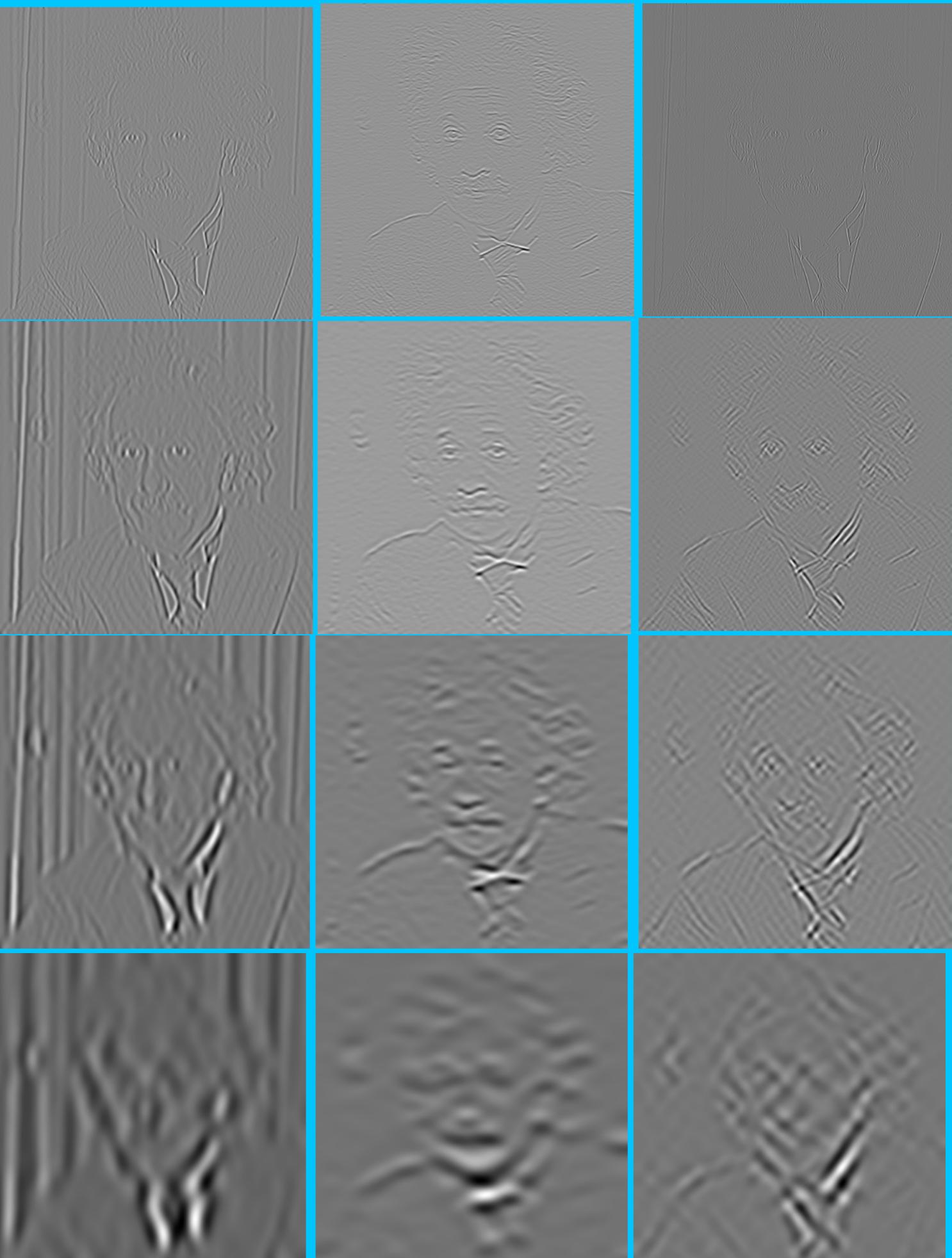
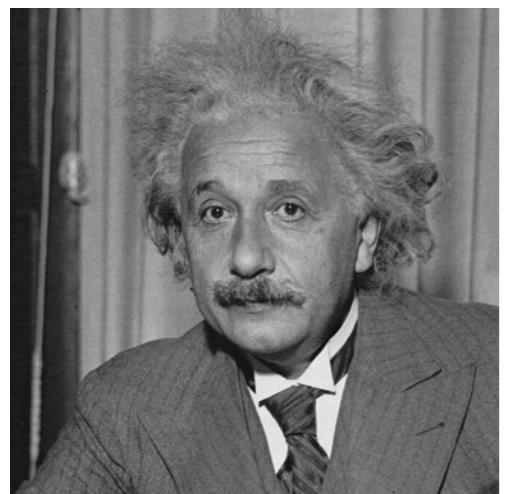
NGC2997

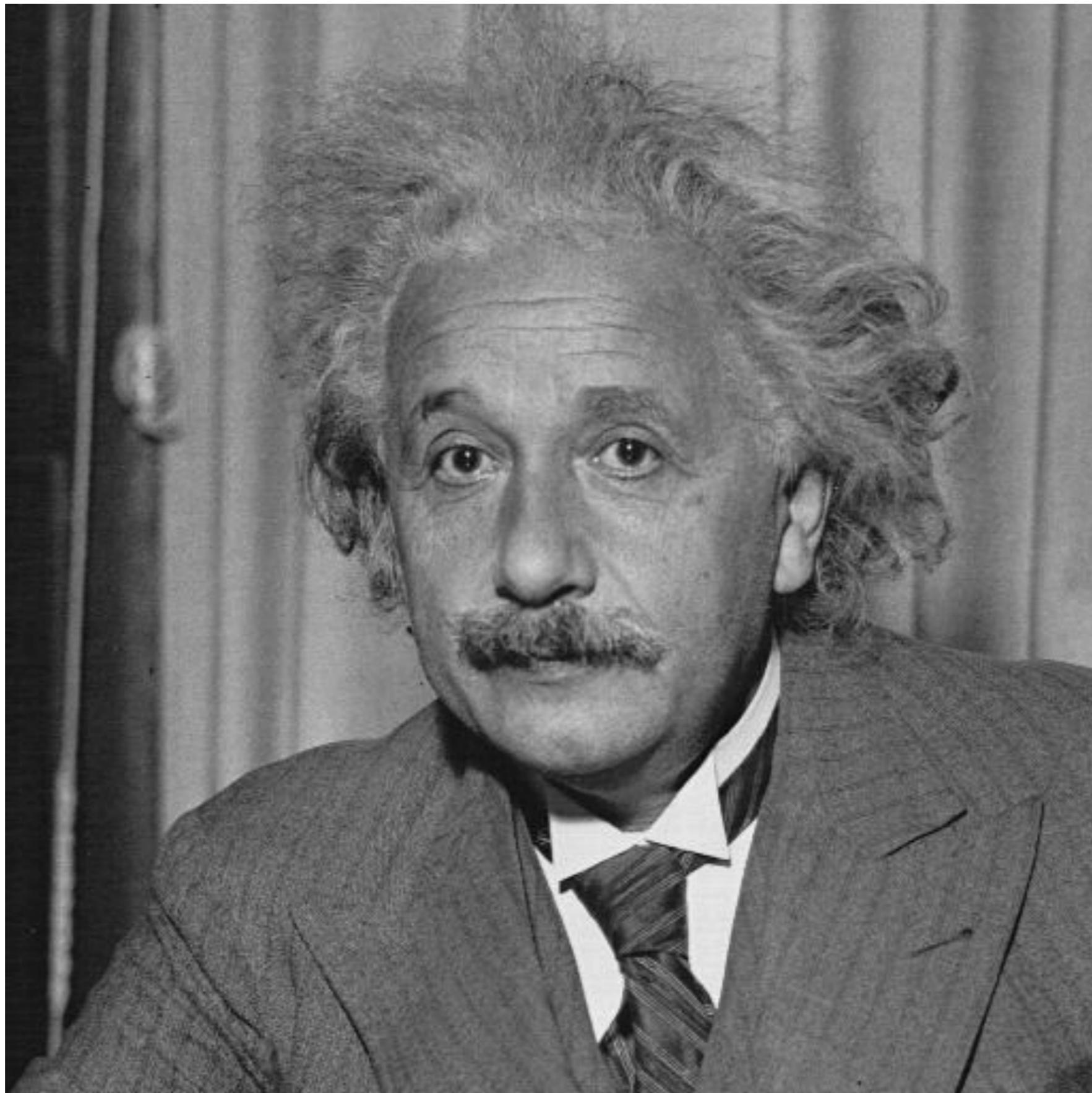
NGC2997 WT

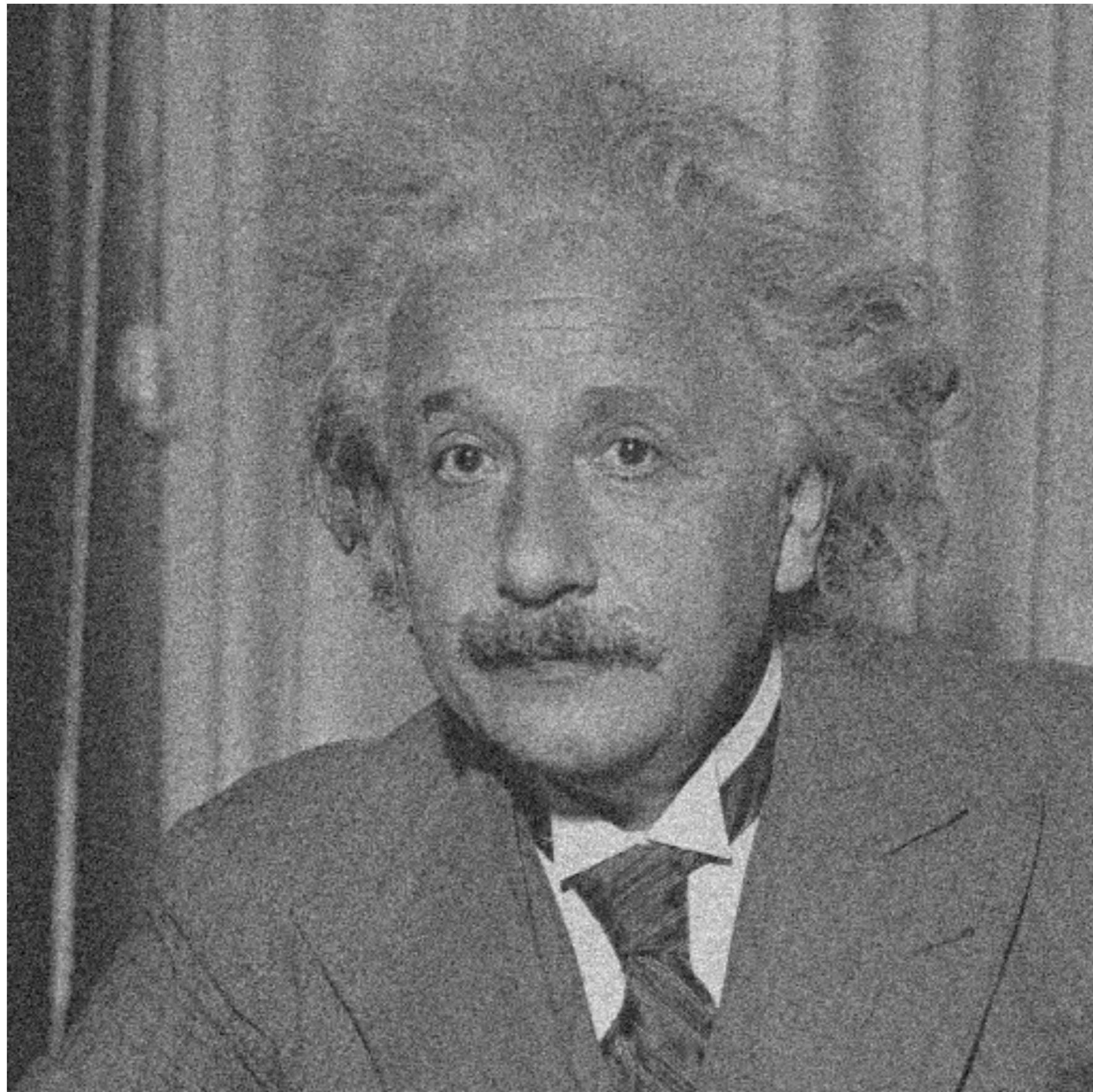


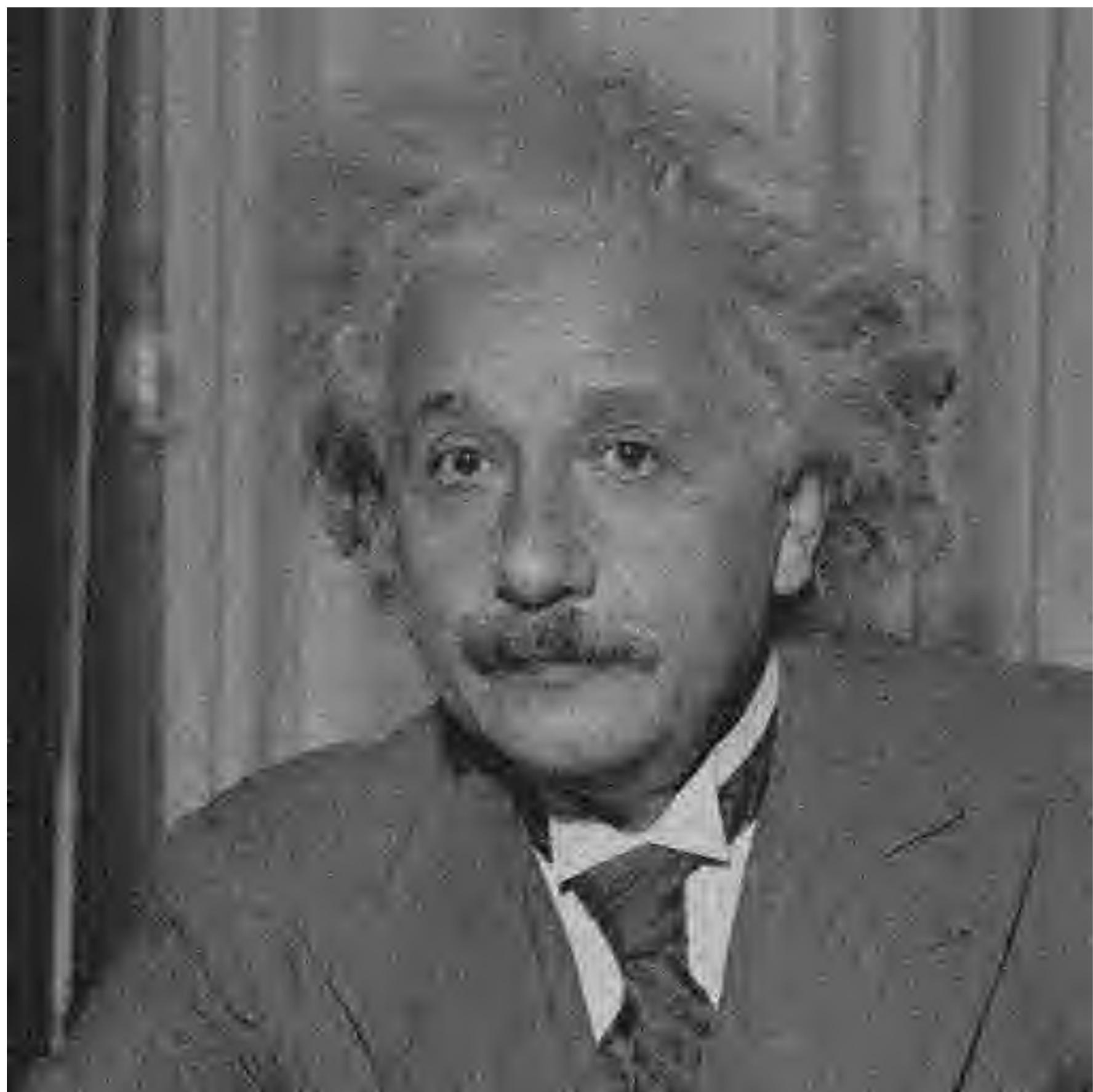


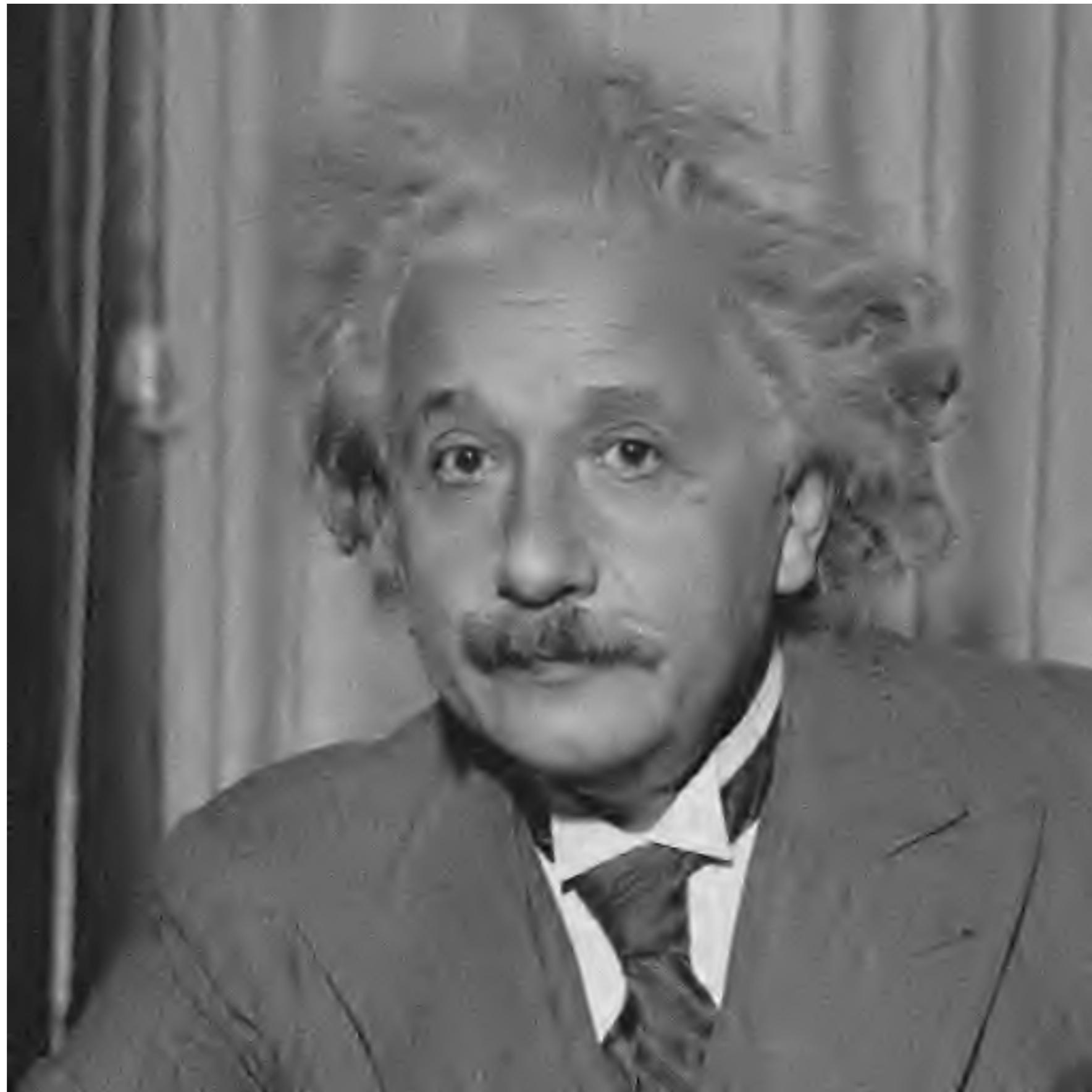
Undecimated Wavelet Transform











ISOTROPIC UNDECIMATED WAVELET TRANSFORM

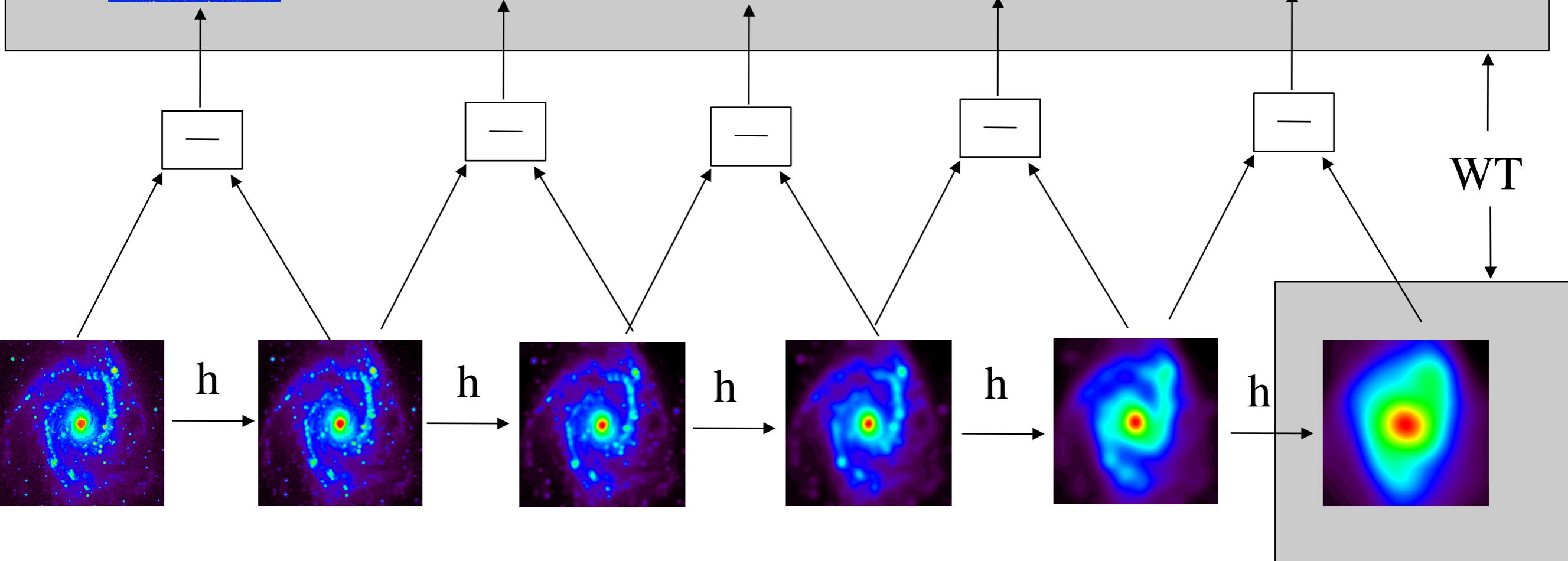
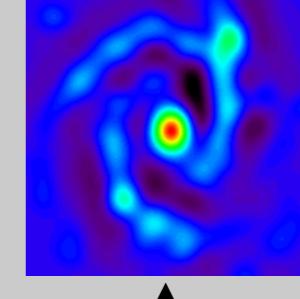
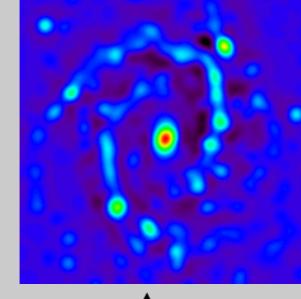
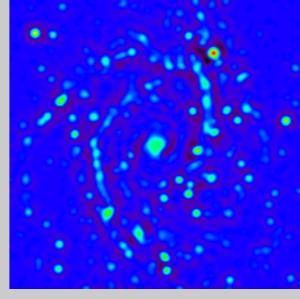
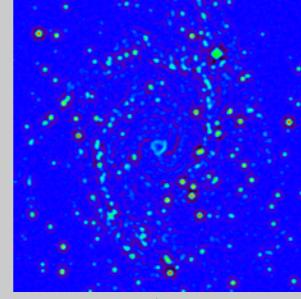
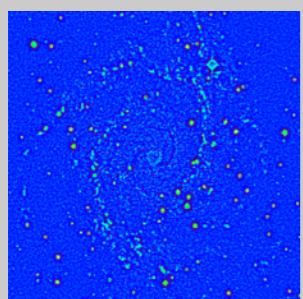
Scale 1

Scale 2

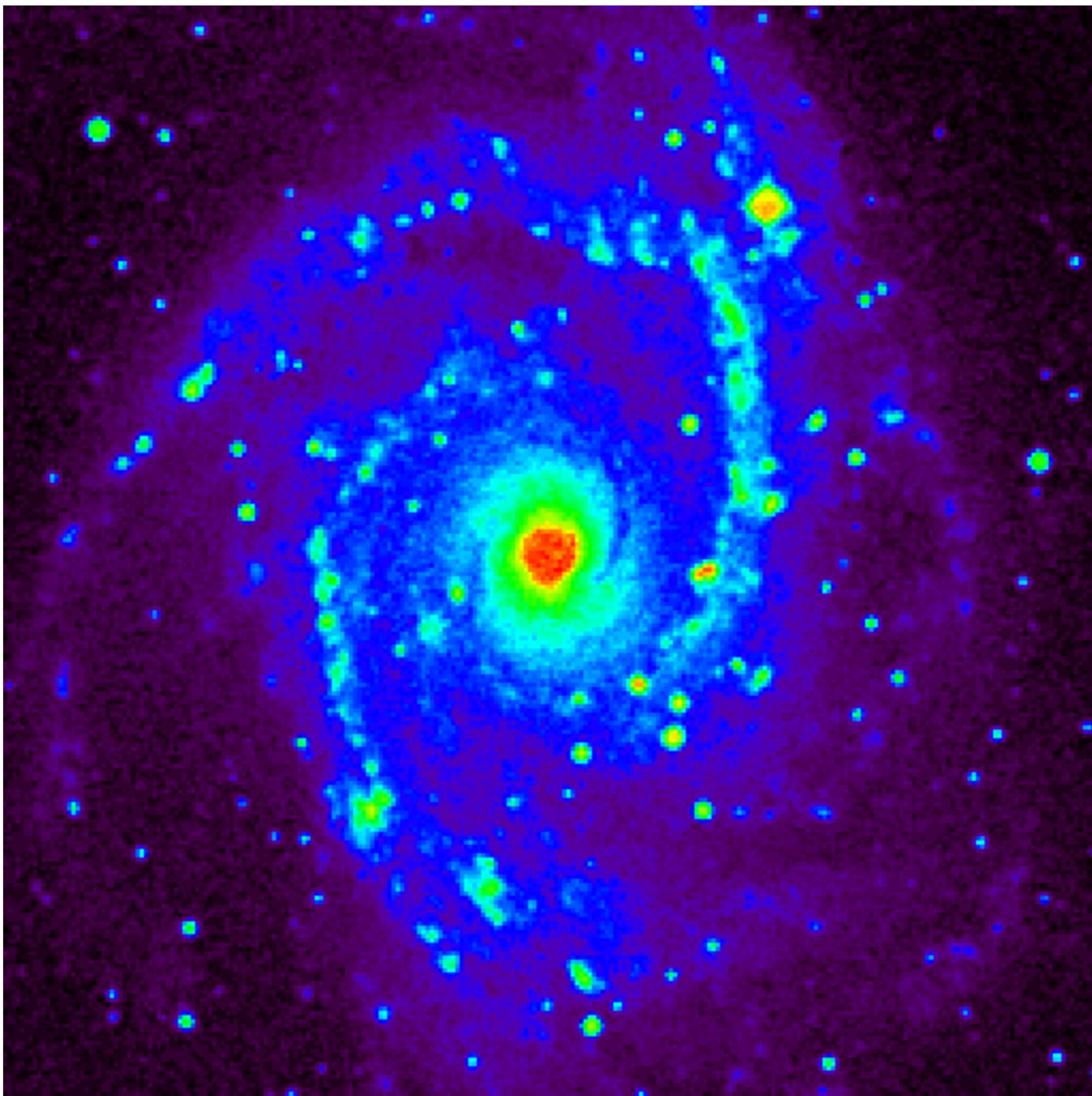
Scale 3

Scale 4

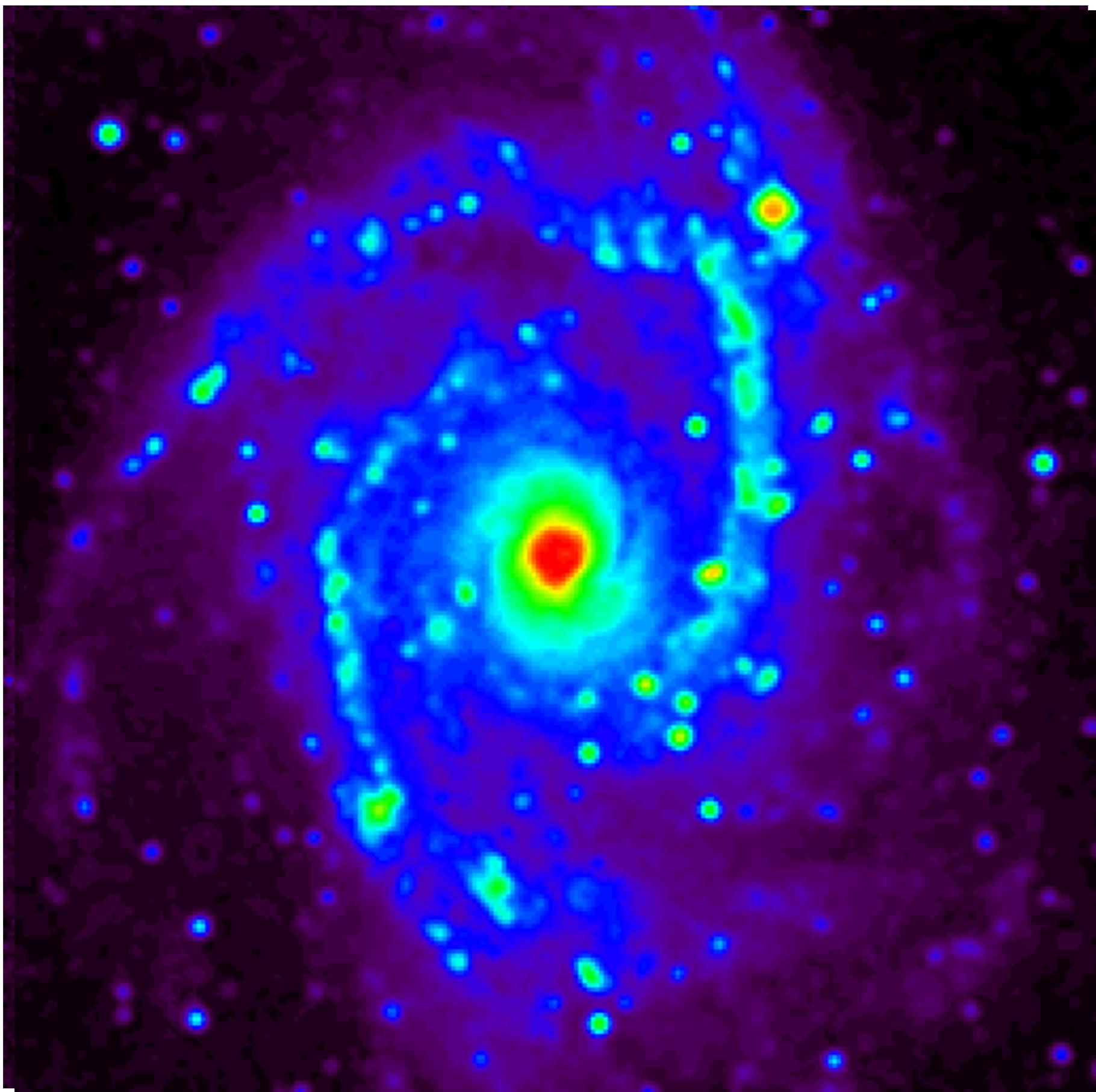
Scale 5



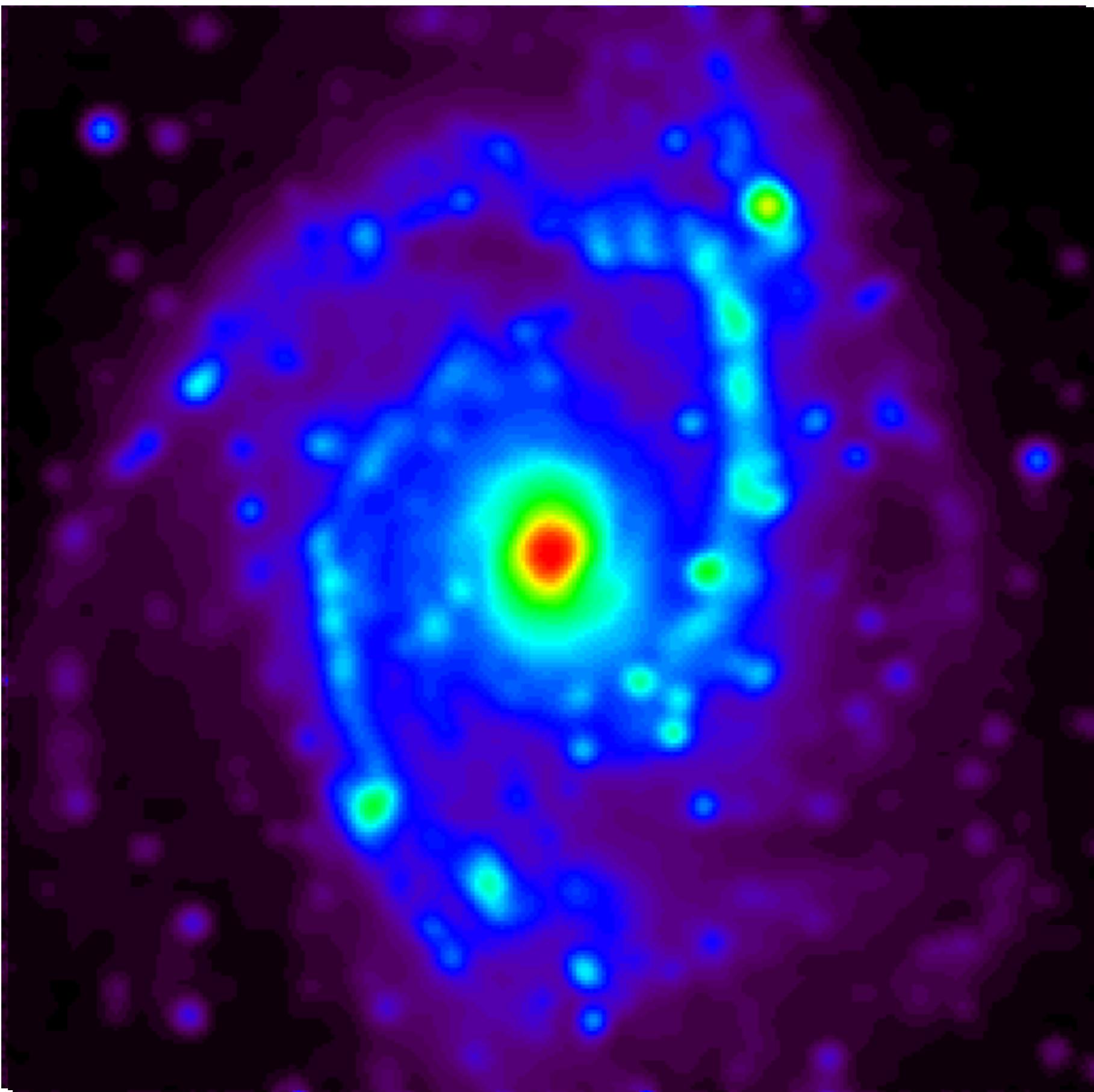
NGC2997



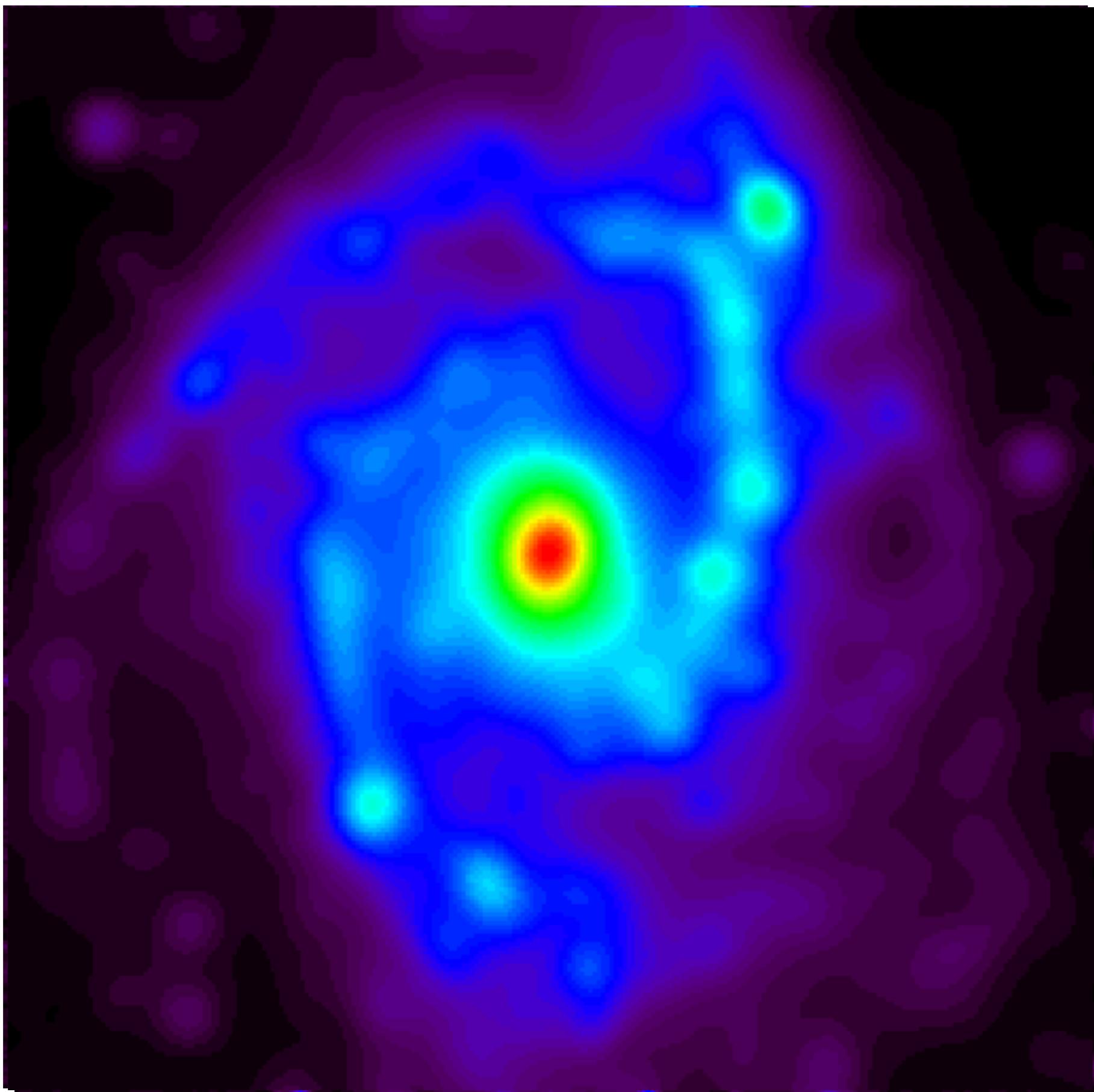
NGC2997



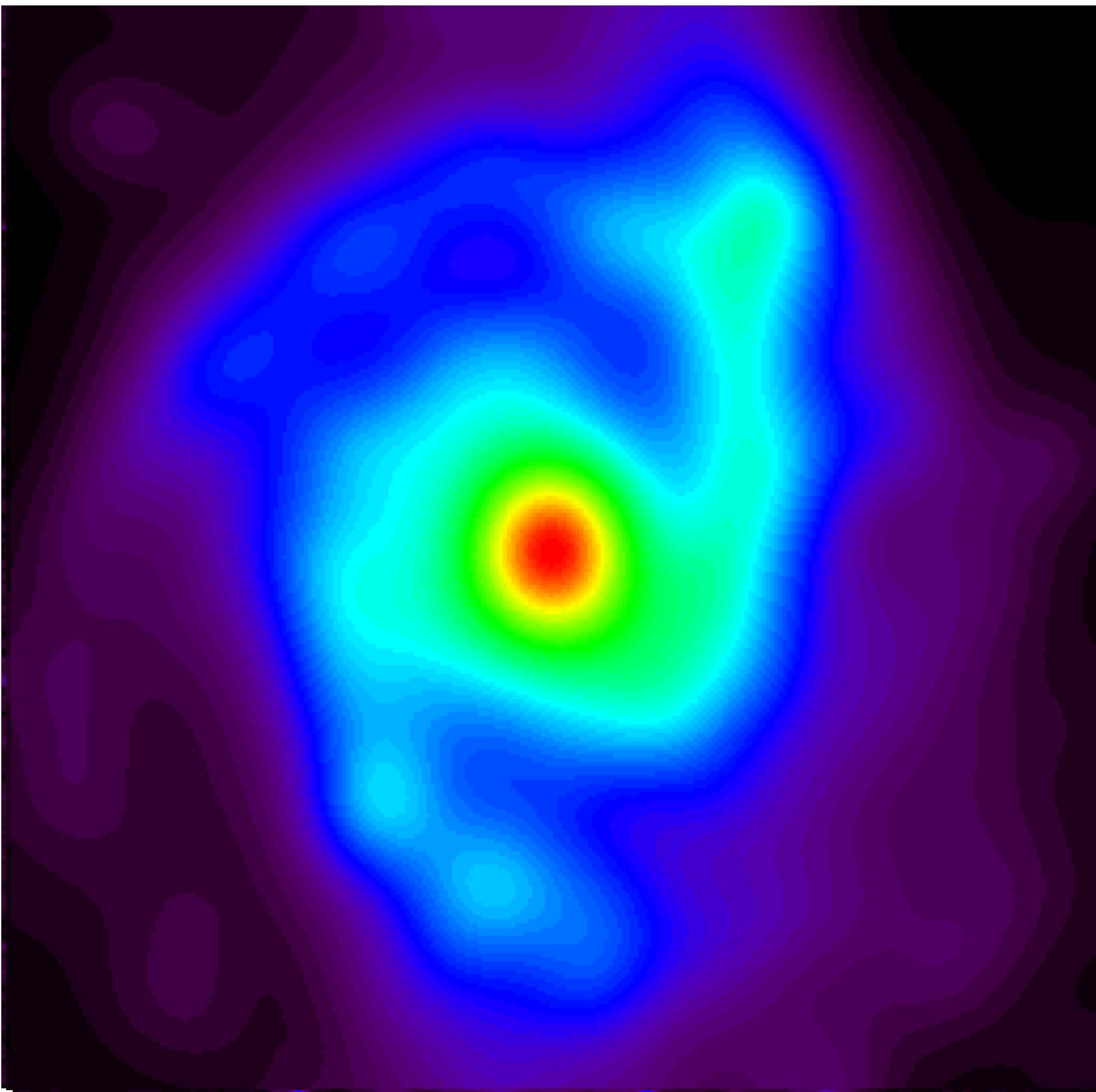
NGC2997



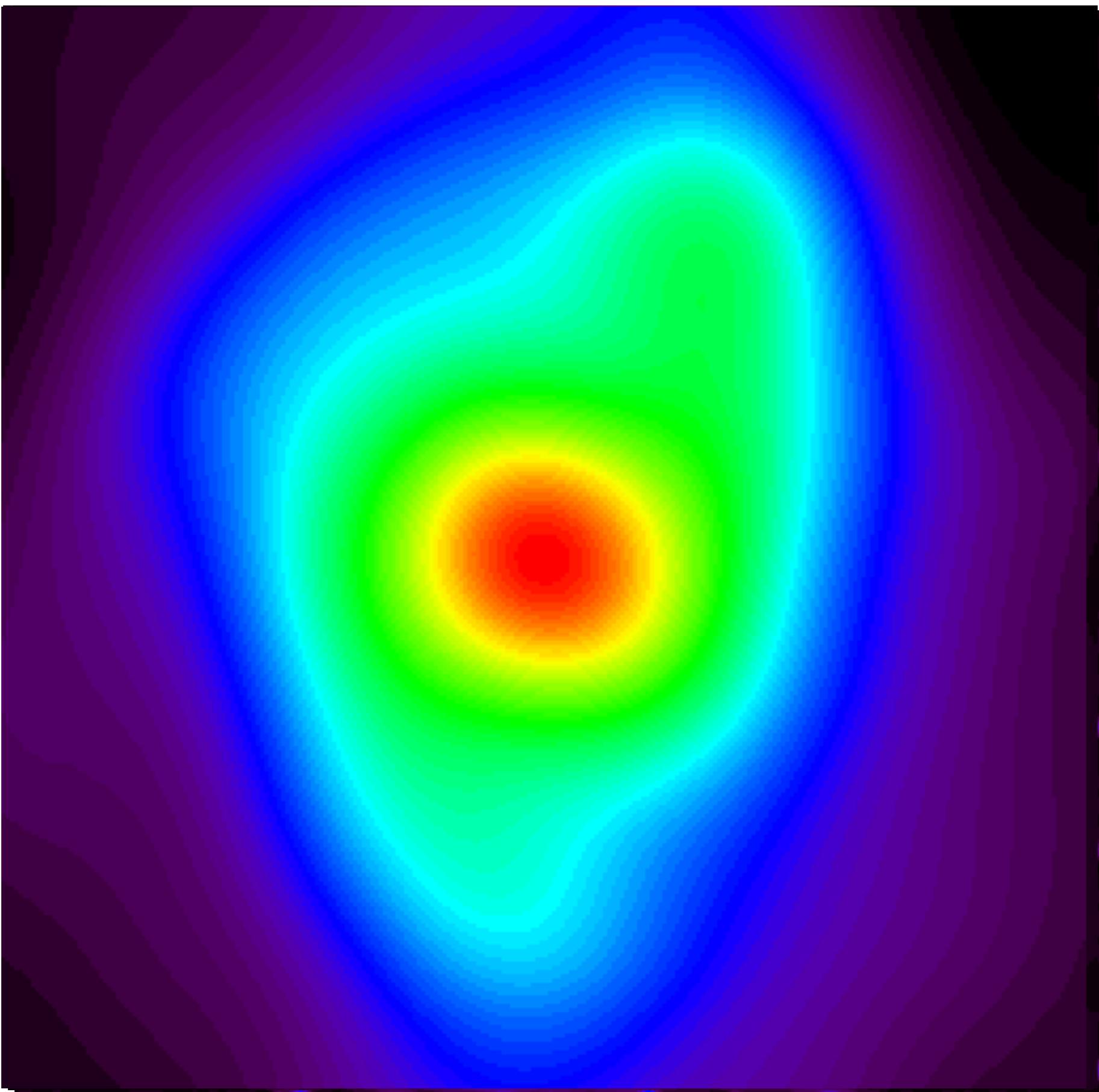
NGC2997



NGC2997



NGC2997



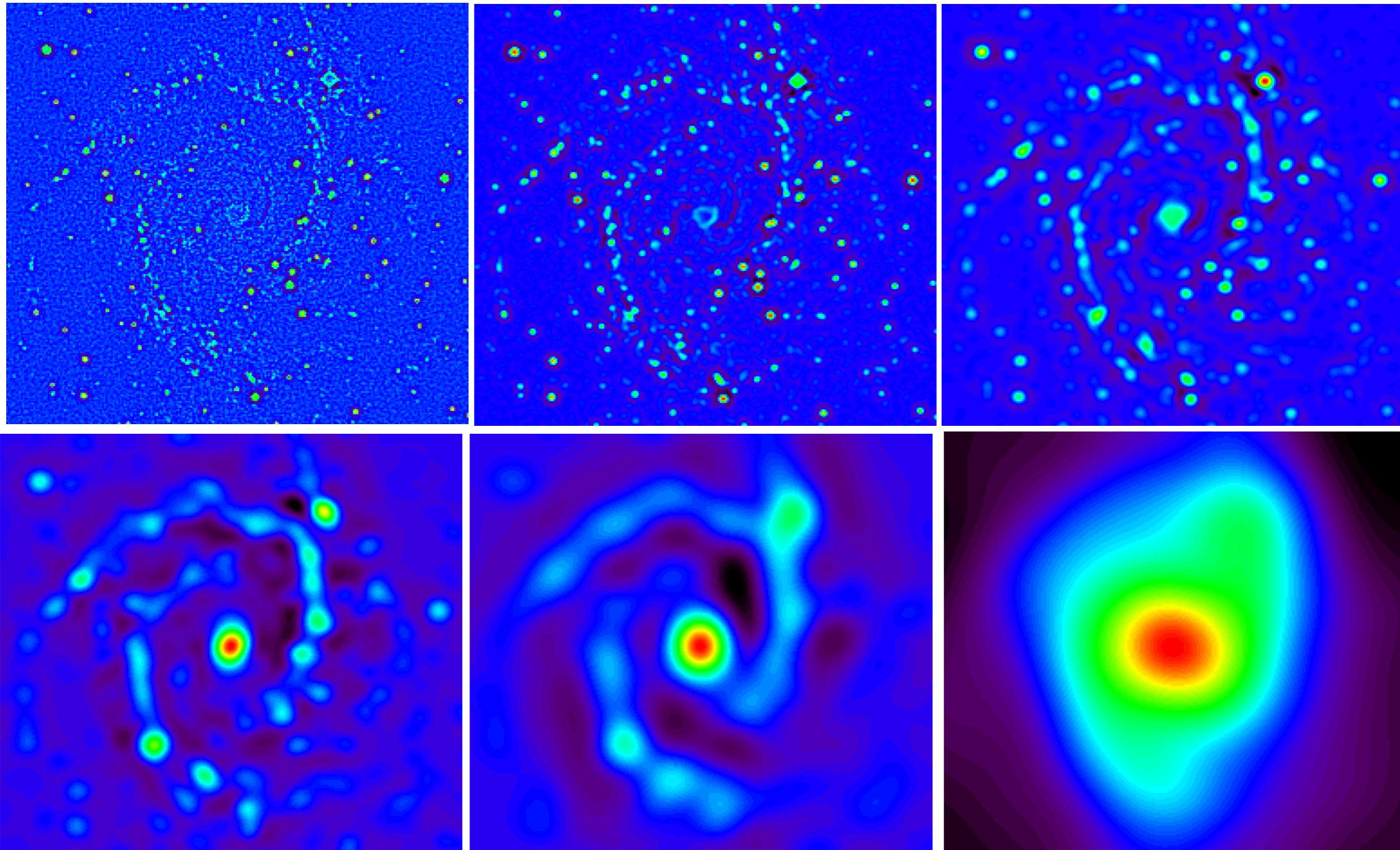
The STARLET Transform

Isotropic Undecimated Wavelet Transform (a trous algorithm)

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

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

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

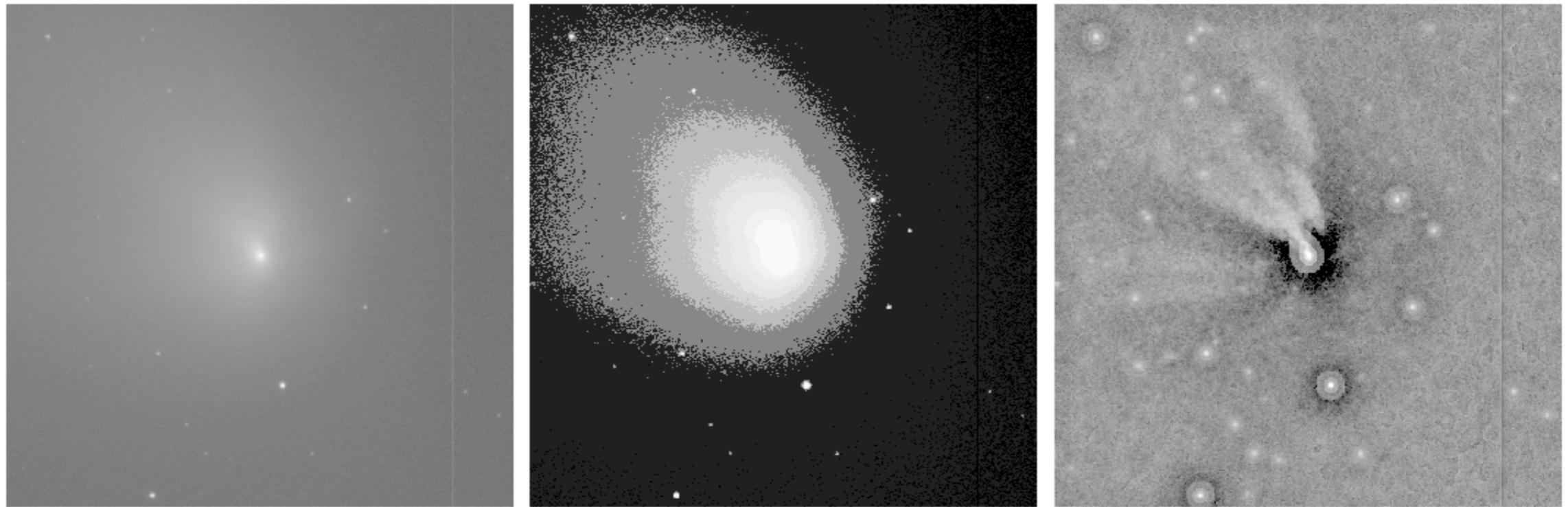


Dynamic Range Compression

Images with a high dynamic range are also difficult to analyze. For example, astronomers generally visualize their images using a logarithmic look-up-table conversion.

Wavelet can also be used to compress the dynamic range at all scales, and therefore allows us to clearly see some very faint features. For instance, the wavelet-log representations consists in replacing $w_{j,k,l}$ by $\log(|w_{j,k,l}|)$, leading to the alternative image

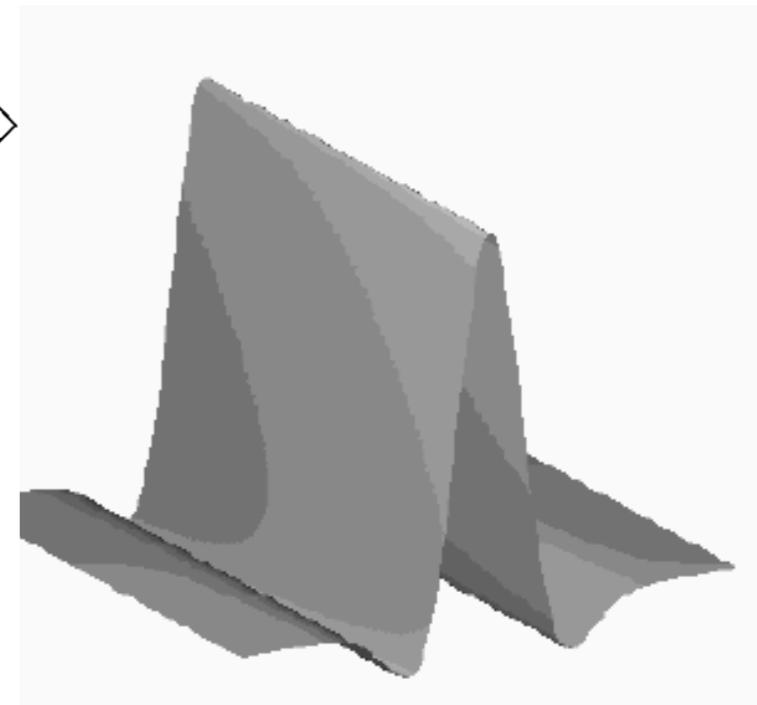
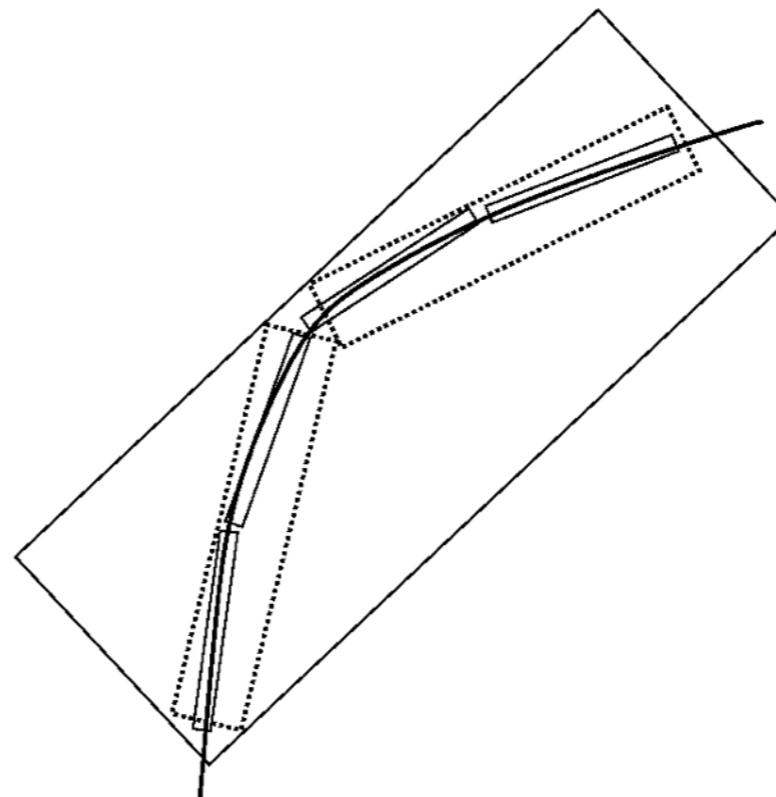
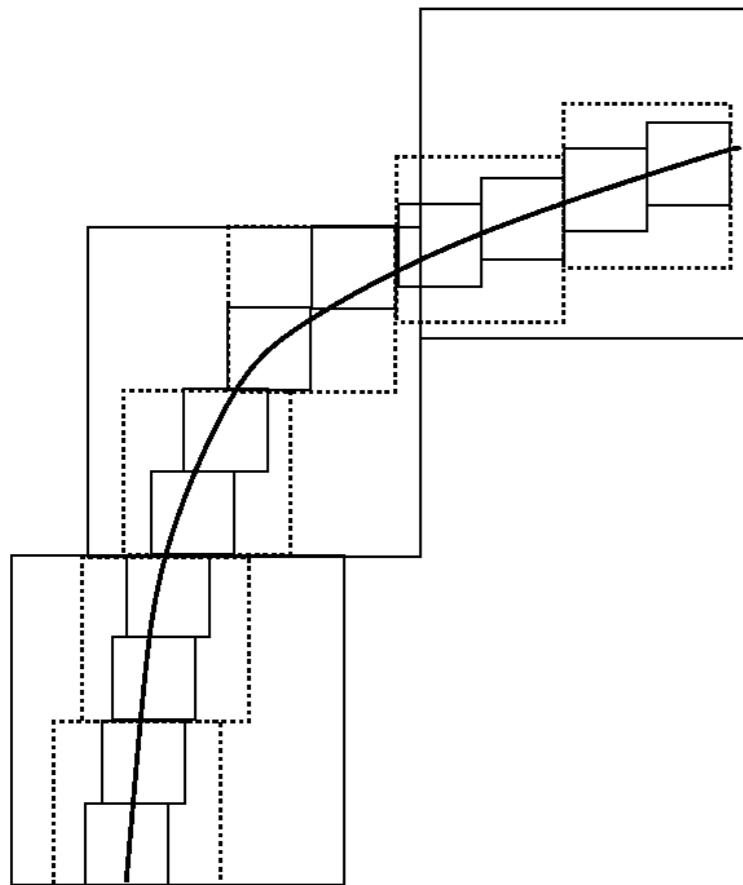
$$I_{k,l} = \log(c_{J,k,l}) + \sum_{j=1}^J \operatorname{sgn}(w_{j,k,l}) \log(|w_{j,k,l}| + \epsilon)$$



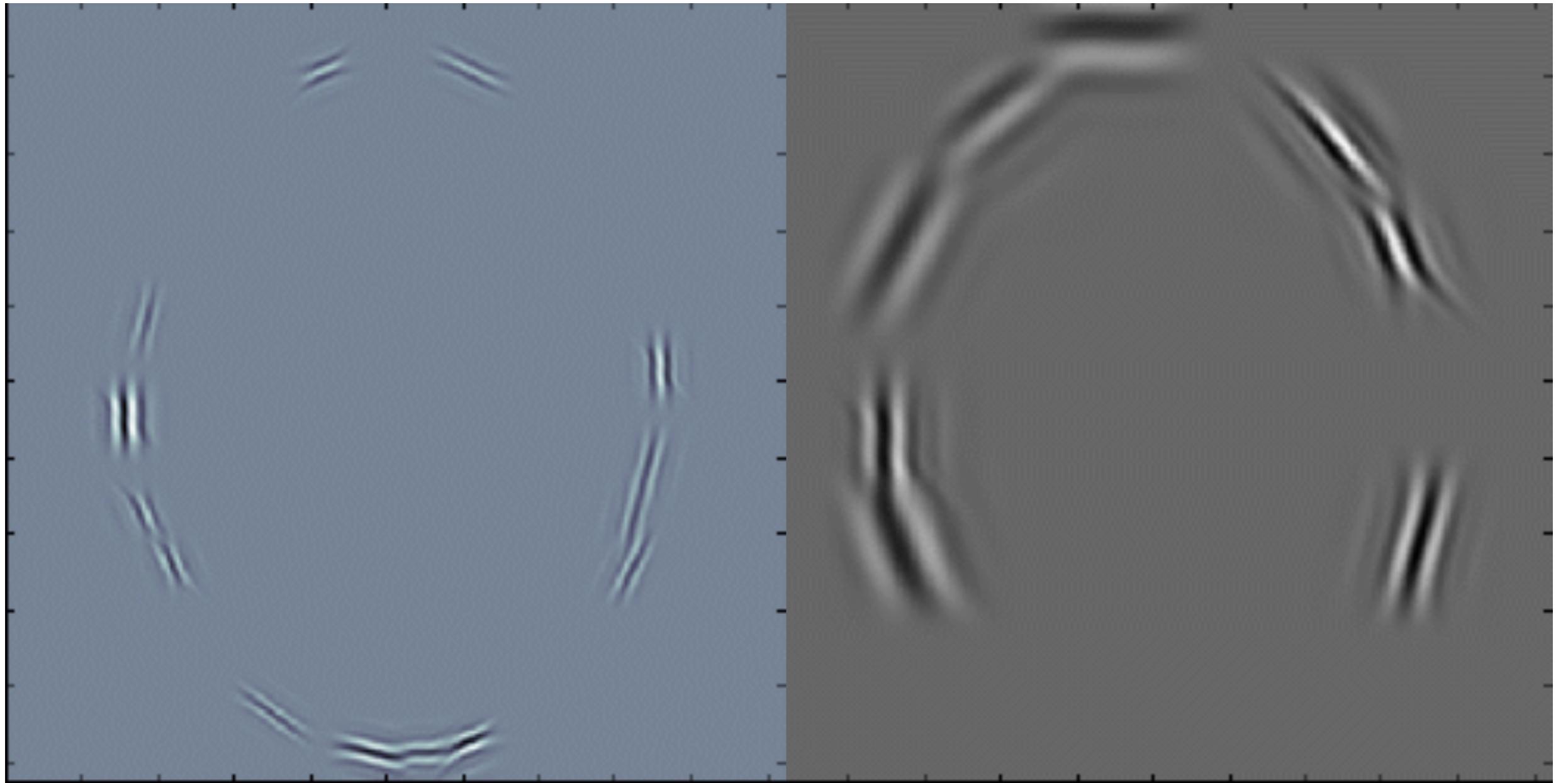
Left - Hale-Bopp Comet image. Middle - histogram equalization results, Right - wavelet-log representations.

Wavelets and edges

- many wavelet coefficients are needed to account for edges i.e. singularities along lines or curves :
- need dictionaries of strongly anisotropic atoms :



ridgelets, curvelets, contourlets, bandelettes, etc.



- J.L. Starck, E. Candes, and D.L. Donoho, "**The Curvelet Transform for Image Denoising**", IEEE Transactions on Image Processing , 11, 6, pp 670 –684, 2002.
- J.-L. Starck, M.K. Nguyen and F. Murtagh, "Wavelets and Curvelets for Image Deconvolution: a Combined Approach", Signal Processing, 83, 10, pp 2279–2283, 2003.
- J.-L. Starck, E. Candes, and D.L. Donoho, "**Astronomical Image Representation by the Curvelet Transform**", Astronomy and Astrophysics, 398, 785–800, 2003.
- J.-L. Starck, F. Murtagh, E. Candes, and D.L. Donoho, "**Gray and Color Image Contrast Enhancement by the Curvelet Transform**", IEEE Transaction on Image Processing, 12, 6, pp 706–717, 2003.

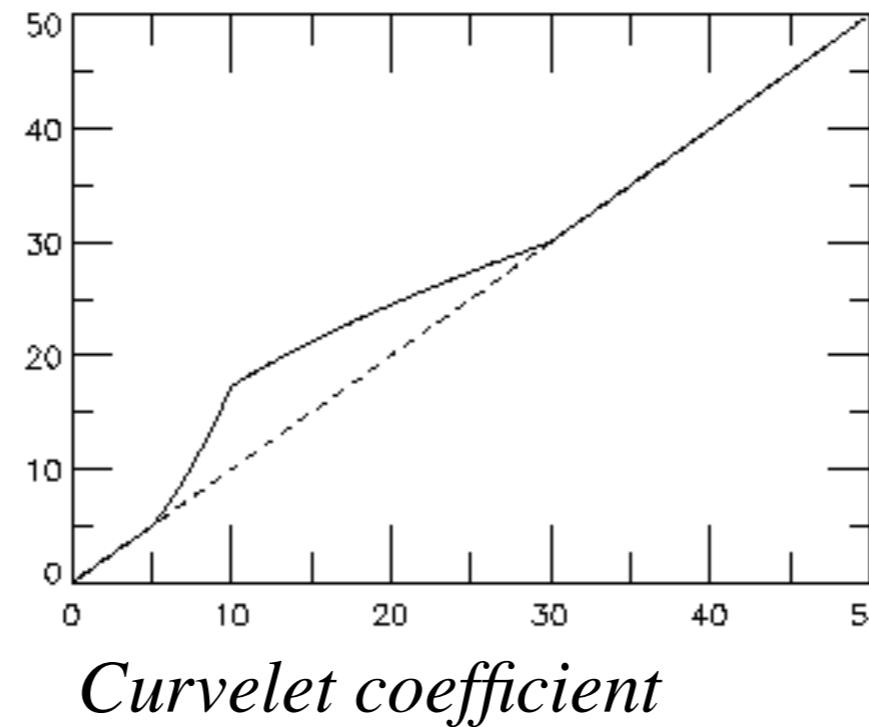
CONTRAST ENHANCEMENT USING THE CURVELET TRANSFORM

J.-L Starck, F. Murtagh, E. Candes and D.L. Donoho, “Gray and Color Image Contrast Enhancement by the Curvelet Transform”,
IEEE Transaction on Image Processing, 12, 6, 2003.

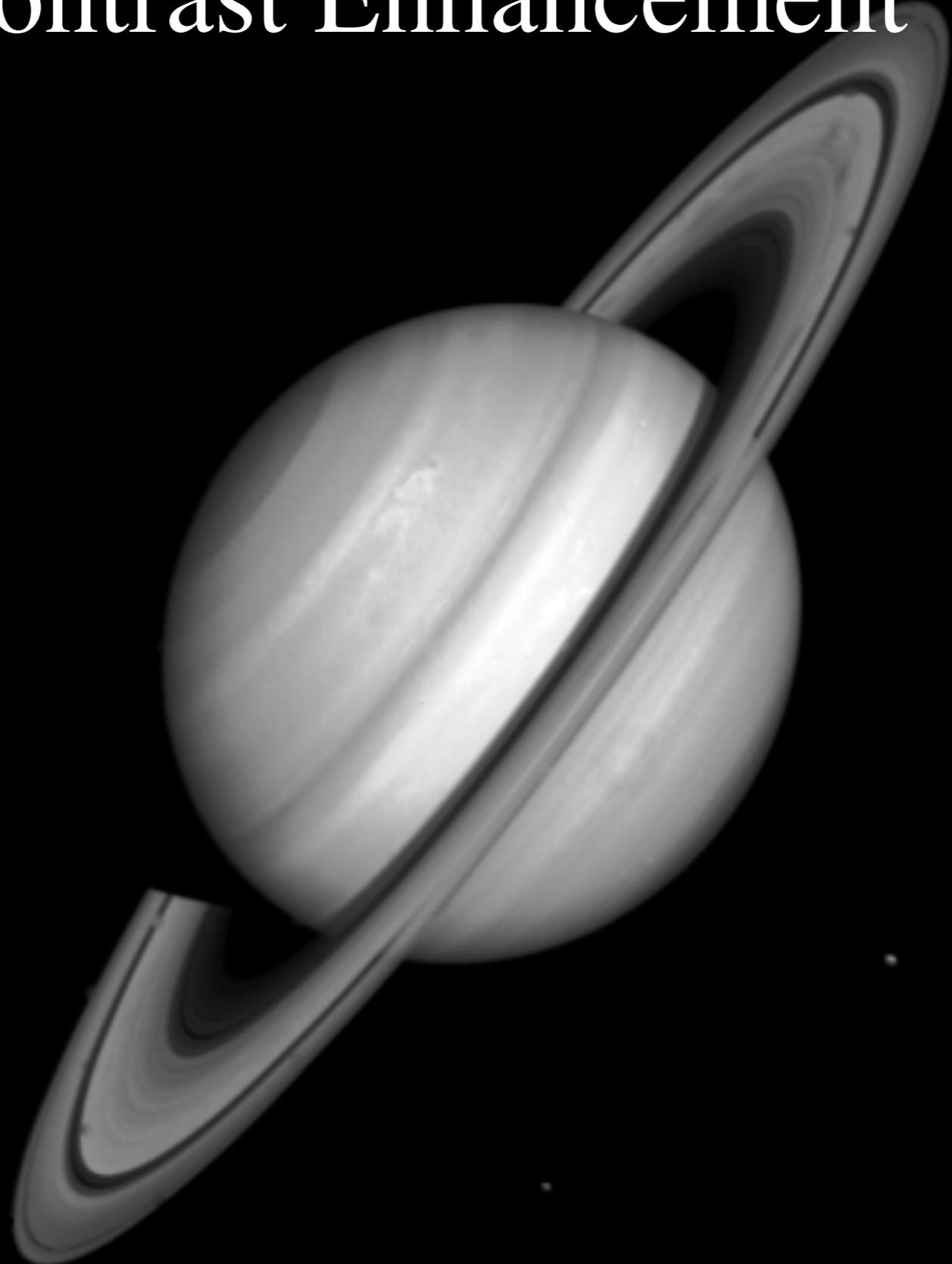
$$\tilde{I} = C_R(y_c(C_T I))$$

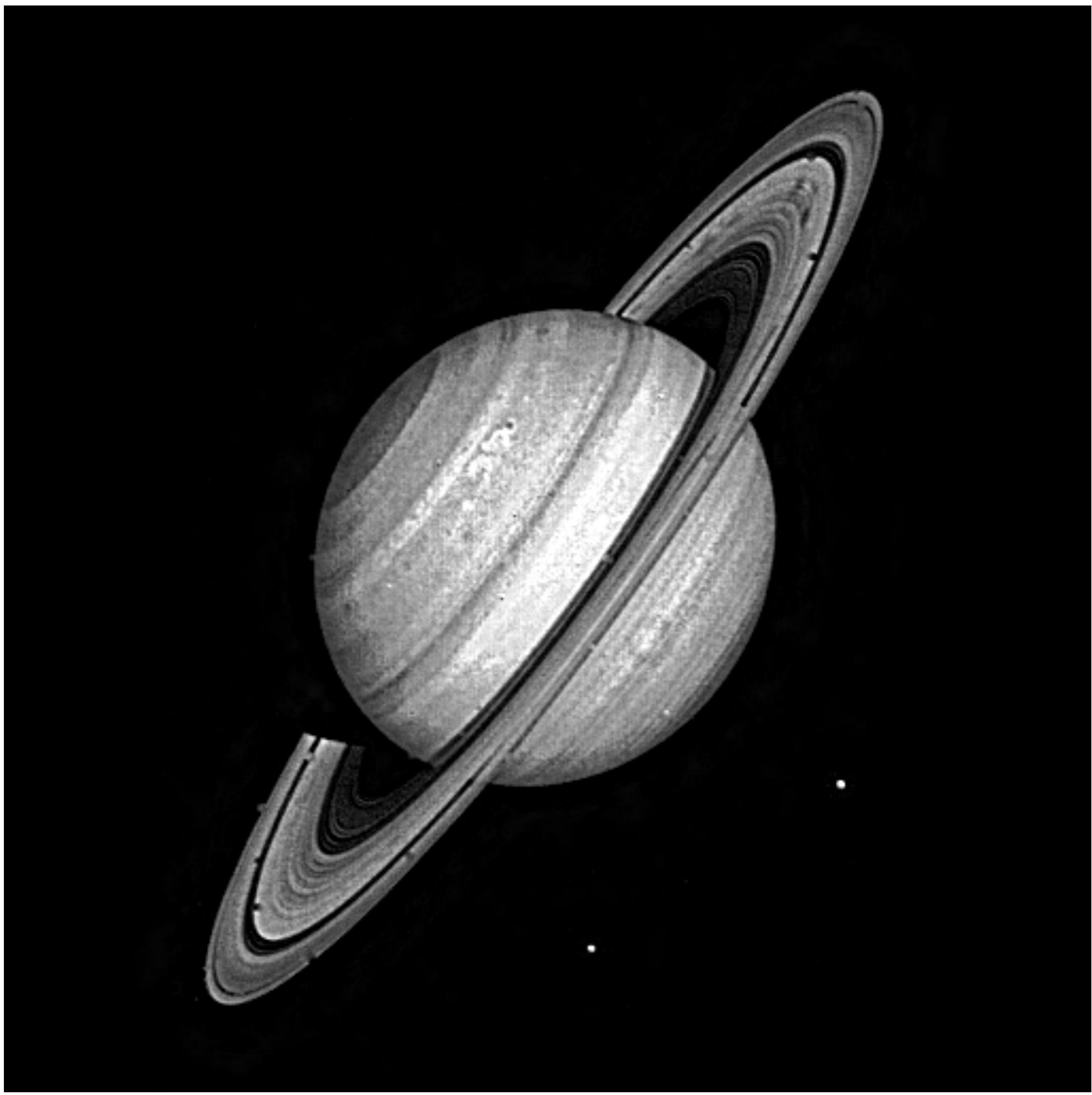
$$\left\{ \begin{array}{ll} y_c(x,\sigma) = 1 & \text{if } x < c\sigma \\ y_c(x,\sigma) = \frac{x - c\sigma}{c\sigma} \left(\frac{m}{c\sigma} \right)^p + \frac{2c\sigma - x}{c\sigma} & \text{if } x < 2c\sigma \\ y_c(x,\sigma) = \left(\frac{m}{x} \right)^p & \text{if } 2c\sigma \leq x < m \\ y_c(x,\sigma) = \left(\frac{m}{x} \right)^s & \text{if } x > m \end{array} \right.$$

*Modified
curvelet
coefficient*



Contrast Enhancement





Multiscale Transforms

Critical Sampling

(bi-) Orthogonal WT

Lifting scheme construction
Wavelet Packets
Mirror Basis

Redundant Transforms

Pyramidal decomposition (Burt and Adelson)
Undecimated Wavelet Transform
Isotropic Undecimated Wavelet Transform
Complex Wavelet Transform
Steerable Wavelet Transform
Dyadic Wavelet Transform
Nonlinear Pyramidal decomposition (Median)

Multiscale Geometric Analysis

Contourlet

Bandelet
Finite Ridgelet Transform
Platelet
(W-)Edgelet
Adaptive Wavelet

Ridgelet

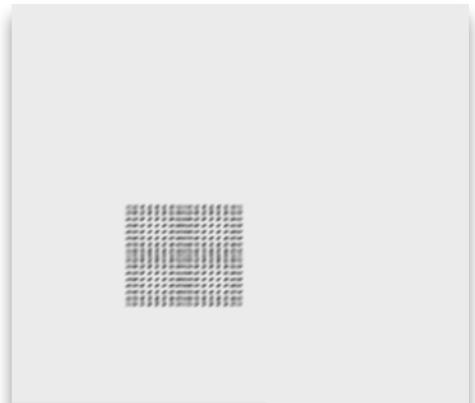
Curvelet (Several implementations)
Wave Atom
Grouplet
[adaptive] **Shearlet** (Several implementations)

Sparsity Model : we consider a dictionary which has a fast transform/reconstruction operator:

$$X = \Phi \alpha$$

Local DCT

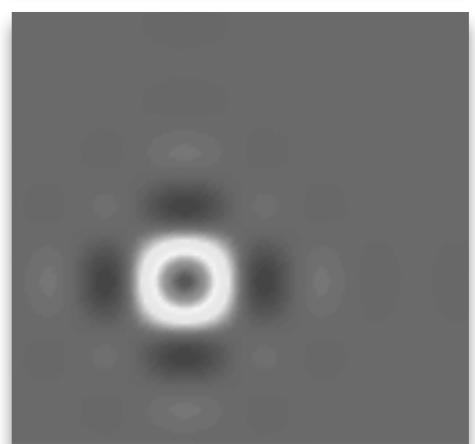
Stationary textures



Locally oscillatory

Wavelet transform

Piecewise smooth



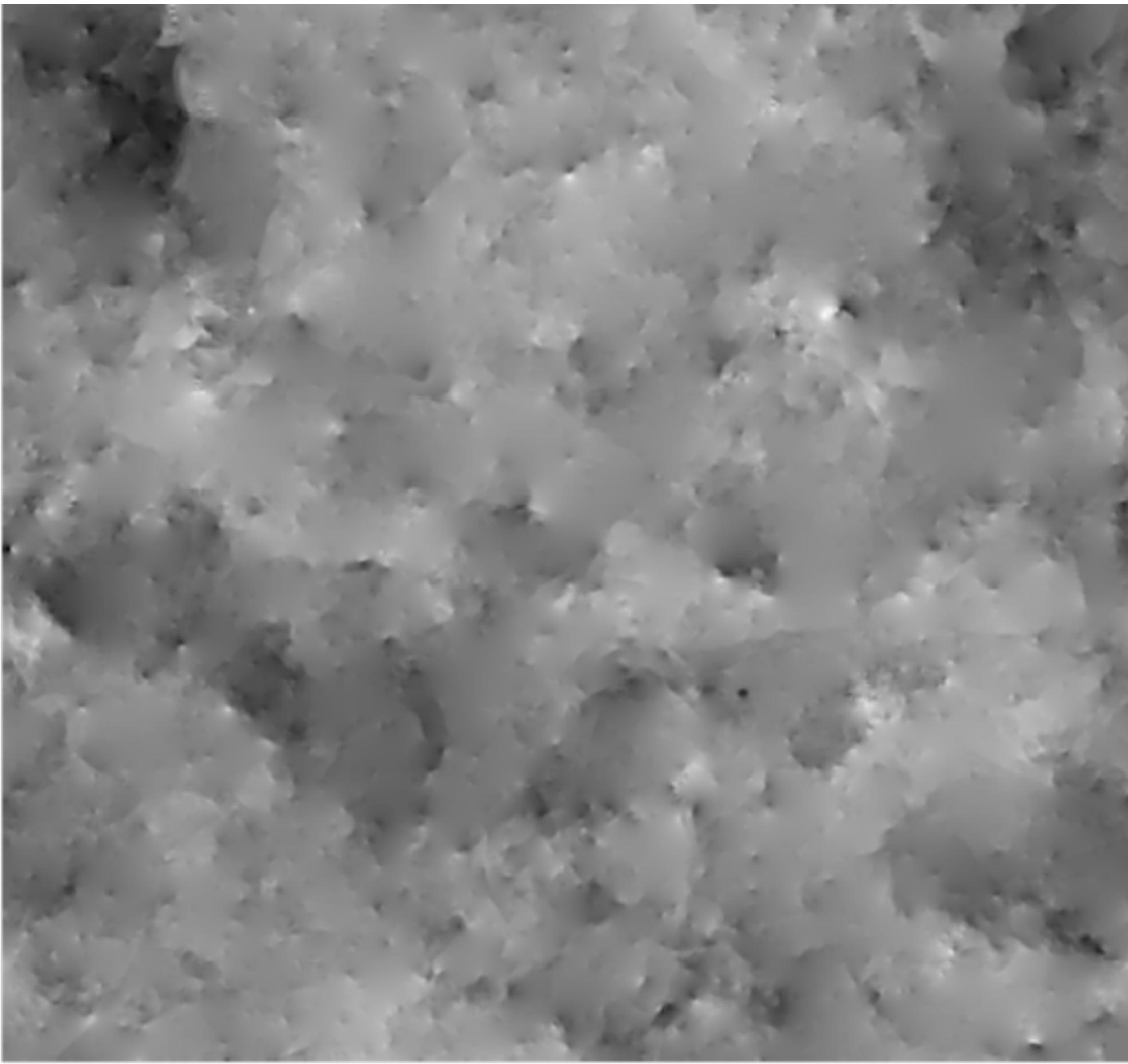
Isotropic structures

Curvelet transform

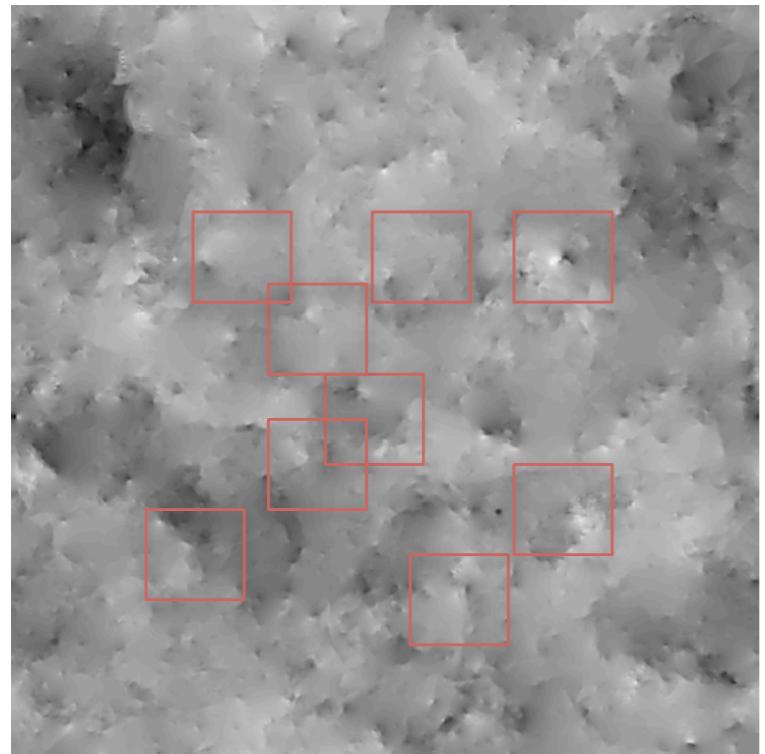
Piecewise smooth, edge



Simulated Cosmic String Map



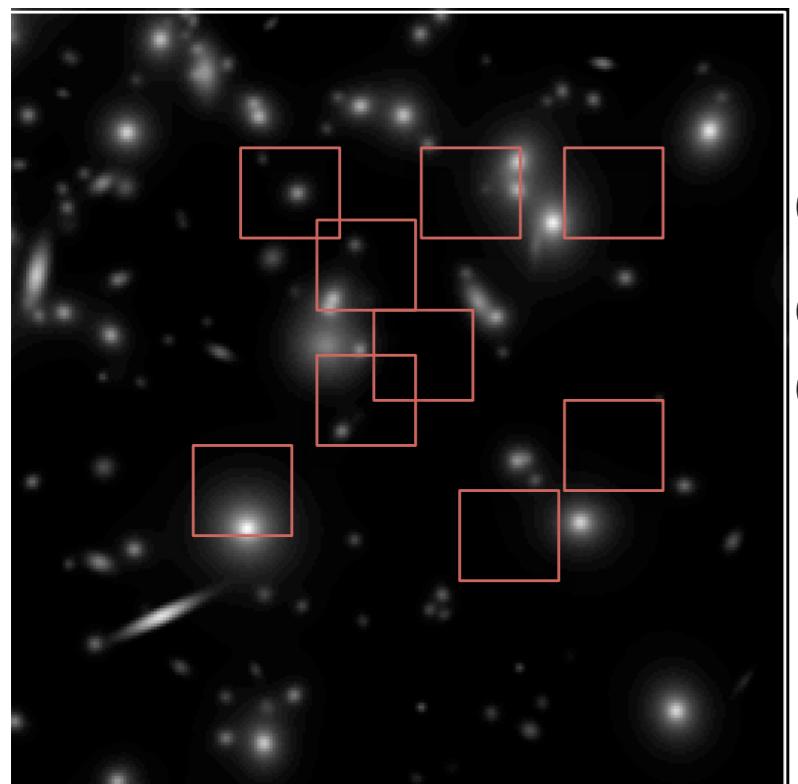
Dictionary Learning



Training basis.



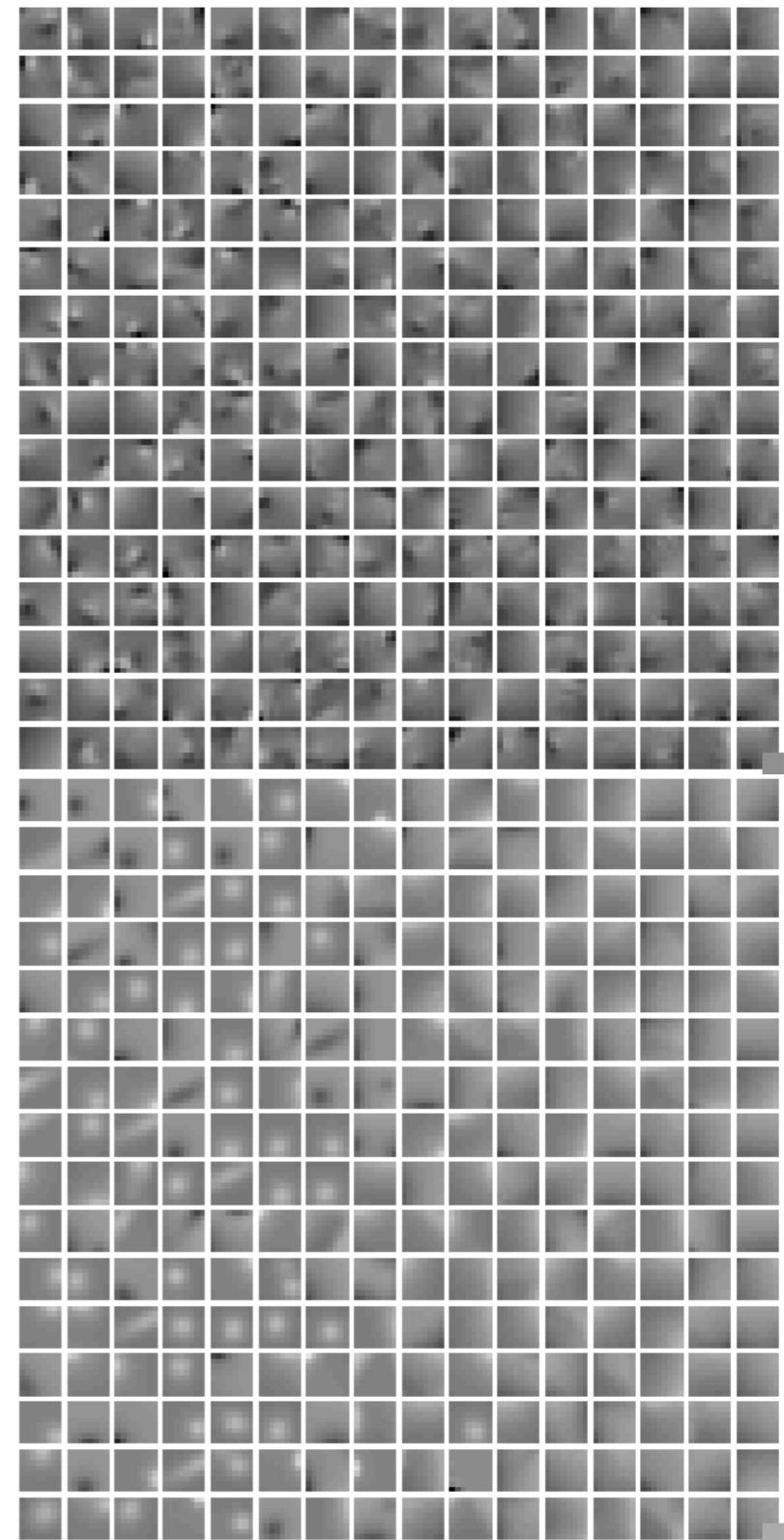
$$(\hat{D}, \hat{A}) = \arg \min_{\substack{D \in C_1 \\ A \in C_2}} (Y = DA)$$



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 ?

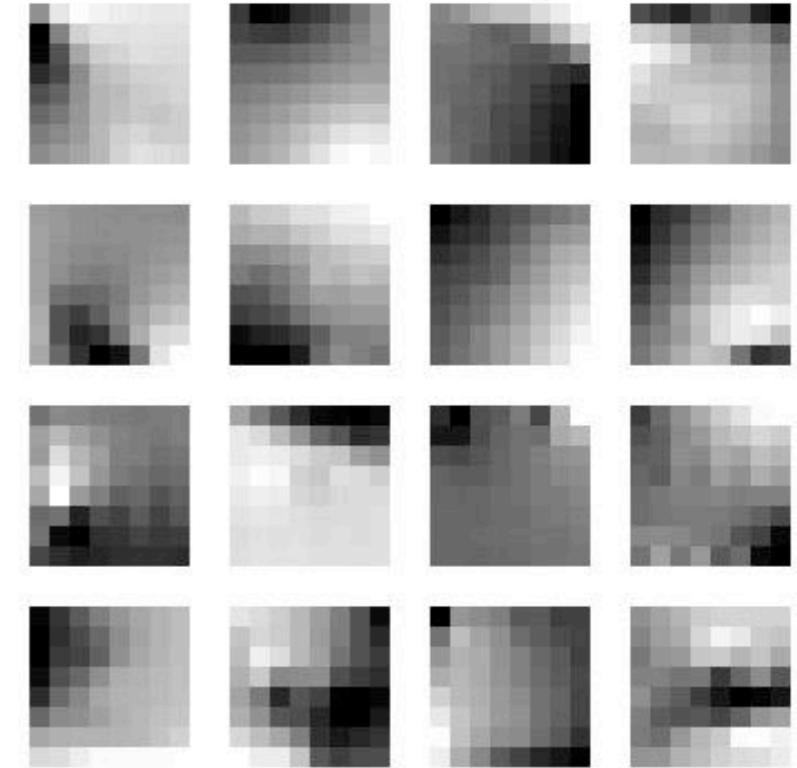
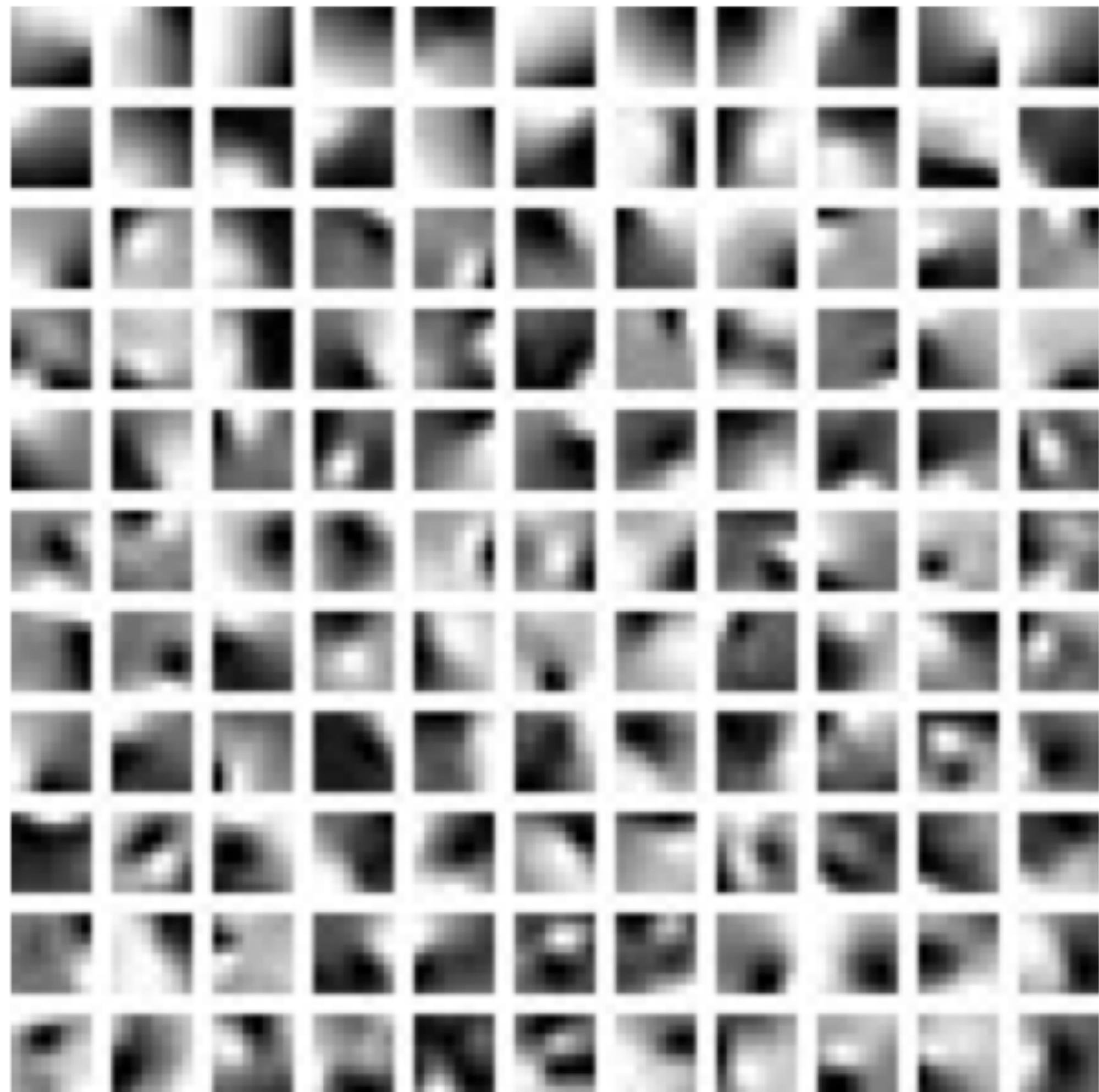


Image to learn from

N training patches \mathbf{Y}

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



Advantages of fixed dictionary : extremely fast.

Advantages of dictionary learning:

atoms can be obtained which are well adapted to the data, and which could never be obtained with a fixed dictionary.

Drawback:

We pay the price of dictionary learning by being less sensitive to detect very faint features.

Complexity: Computation time, parameters, etc

- Part I: Introduction to Inverse Problems in Astrophysics
- Part II: Sparsity & Dictionaries
- Part III: Sparse Regularization**
- Part IV: Unmixing
- Part V: Sparsity for Planck and Euclid Space Missions

Example: DENOISING

Denoising using a sparsity model ($H=Id$)

$$Y = X + N$$

$$\tilde{\alpha} \in \arg \min_{\alpha} \frac{1}{2} \| Y - \Phi \alpha \|^2 + \lambda \| \alpha \|_p^p, \quad 0 \leq p \leq 1.$$

$$X = \Phi \alpha$$

$$Y = HX + N \quad \text{with} \quad X = \Phi\alpha$$

$$\|X\|_p = \left(\sum_i |X_i|^p \right)^{\frac{1}{p}}$$

1 - Which Norm ? P in $[0,1]$

2 - Constraint versus Lagrangian formulation

Constraint formulation: $\min_{\alpha} \|\alpha\|_p^p$ subject to $\|Y - H\Phi\alpha\|^2 \leq \epsilon$

Lagrangian formulation: $\min_{\alpha} \|Y - H\Phi\alpha\|^2 + \lambda \|\alpha\|_p^p$

3 - Analysis versus Synthesis ?

Synthesis form: $\min_{\alpha} \|Y - H\Phi\alpha\|^2 + \lambda \|\alpha\|_p^p$

Analysis form: $\min_X \|\alpha\|_p^p \|Y - HX\|^2 + \lambda \|\Phi^t X\|_p^p$

4 - Which dictionary ?

5 - Which noise model ?

6 - Which minimization method ?

7 - How to fix the regularization parameter ?

Minimization: P=0

$$\tilde{\alpha} \in \arg \min_{\alpha} \frac{1}{2} \| Y - \Phi \alpha \|^2 + \lambda \| \alpha \|_0$$

\implies Solution via Iterative **Hard** Thresholding

$$\tilde{\alpha}^{(t+1)} = \text{HardThresh}_{\mu t}(\tilde{\alpha}^{(t)} + \mu \Phi^T (Y - \Phi \tilde{\alpha}^{(t)})), \mu = 1/\|\Phi\|^2.$$

$$\tilde{\alpha}_{j,k} = \text{HardThresh}_t(\alpha_{j,k}) = \begin{cases} \alpha_{j,k} & \text{if } |\alpha_{j,k}| \geq t, \\ 0 & \text{otherwise.} \end{cases}$$

$$\lambda = \frac{t^2}{2}$$

1st iteration solution:

$$\tilde{X} = \Phi \text{HardThresh}_t(\Phi^T Y) = \Delta_{\Phi,t}(Y)$$

Exact for Φ orthonormal.

Minimization: P=1

$$\tilde{\alpha} \in \arg \min_{\alpha} \frac{1}{2} \| Y - \Phi \alpha \|^2 + \lambda \| \alpha \|_0$$

\implies Solution via iterative **Soft Thresholding**

$$\tilde{\alpha}^{(t+1)} = \text{SoftThresh}_{\mu t}(\tilde{\alpha}^{(t)} + \mu \Phi^T(Y - \Phi \tilde{\alpha}^{(t)})), \mu \in (0, 2/\|\Phi\|^2).$$

$$\text{SoftThresh}_t(x) = \text{sgn}(x) \max(0, |x| - t)$$

1st iteration solution:

$$\tilde{X} = \Phi \text{ SoftThresh}_t(\Phi^T Y) = \Delta_{\Phi,t}(Y)$$

Exact for Φ orthonormal.

DETECTION of Active Coefficients:

For a positive coefficient:

$$P = \text{Prob}(x > \alpha_i)$$

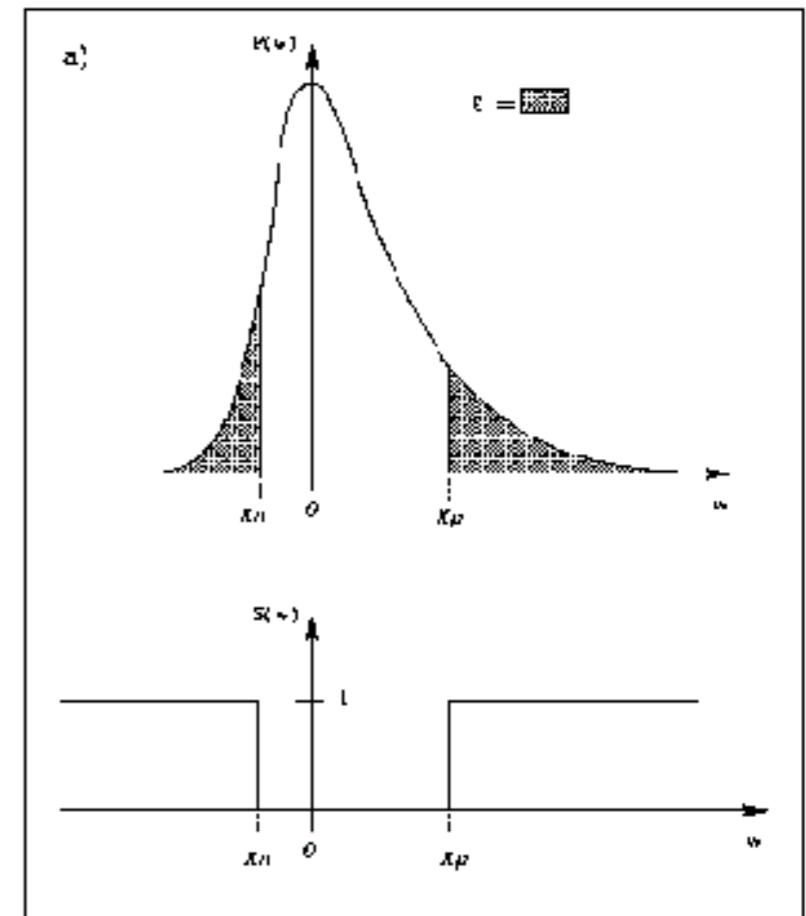
For a negative coefficient: $P = \text{Prob}(x < \alpha_i)$

Hypothesis testing: given a threshold e :

H_0 : if $P > e$, the coefficient could be due to the noise.

H_1 : if $P < e$, the coefficient cannot be due to the noise, and a **significant coefficient** is detected.

Gaussian Noise: $p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$



It suffices to compare α_i to $t = k\sigma$, (with $\lambda = \frac{t^2}{2}$) :

if $|\alpha_i| \geq k\sigma$ then α_i is **significant** or **active**.

if $|\alpha_i| < k\sigma$ then α_i is **NOT significant**.

More **sophisticated tests** exist: False Discovery Rate (FDR), block sparsity, etc

DETECTION of Active Coefficients:

For a positive coefficient:

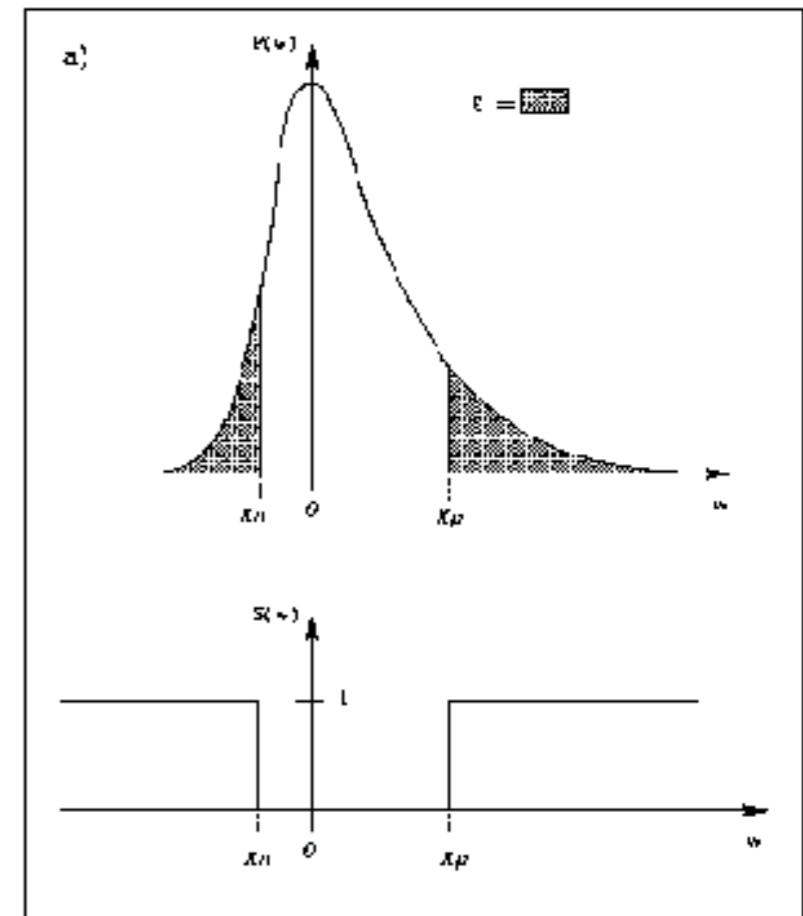
$$P = \text{Prob}(x > \alpha_i)$$

For a negative coefficient: $P = \text{Prob}(x < \alpha_i)$

Hypothesis testing: given a threshold e :

H_0 : if $P > e$, the coefficient could be due to the noise.

H_1 : if $P < e$, the coefficient cannot be due to the noise, and a **significant coefficient** is detected.



Gaussian Noise: $p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$

It suffices to compare α_i to $t = k\sigma$, (with $\lambda = \frac{t^2}{2}$) :

if $|\alpha_i| \geq k\sigma$ then α_i is **significant** or **active**.

if $|\alpha_i| < k\sigma$ then α_i is **NOT significant**.

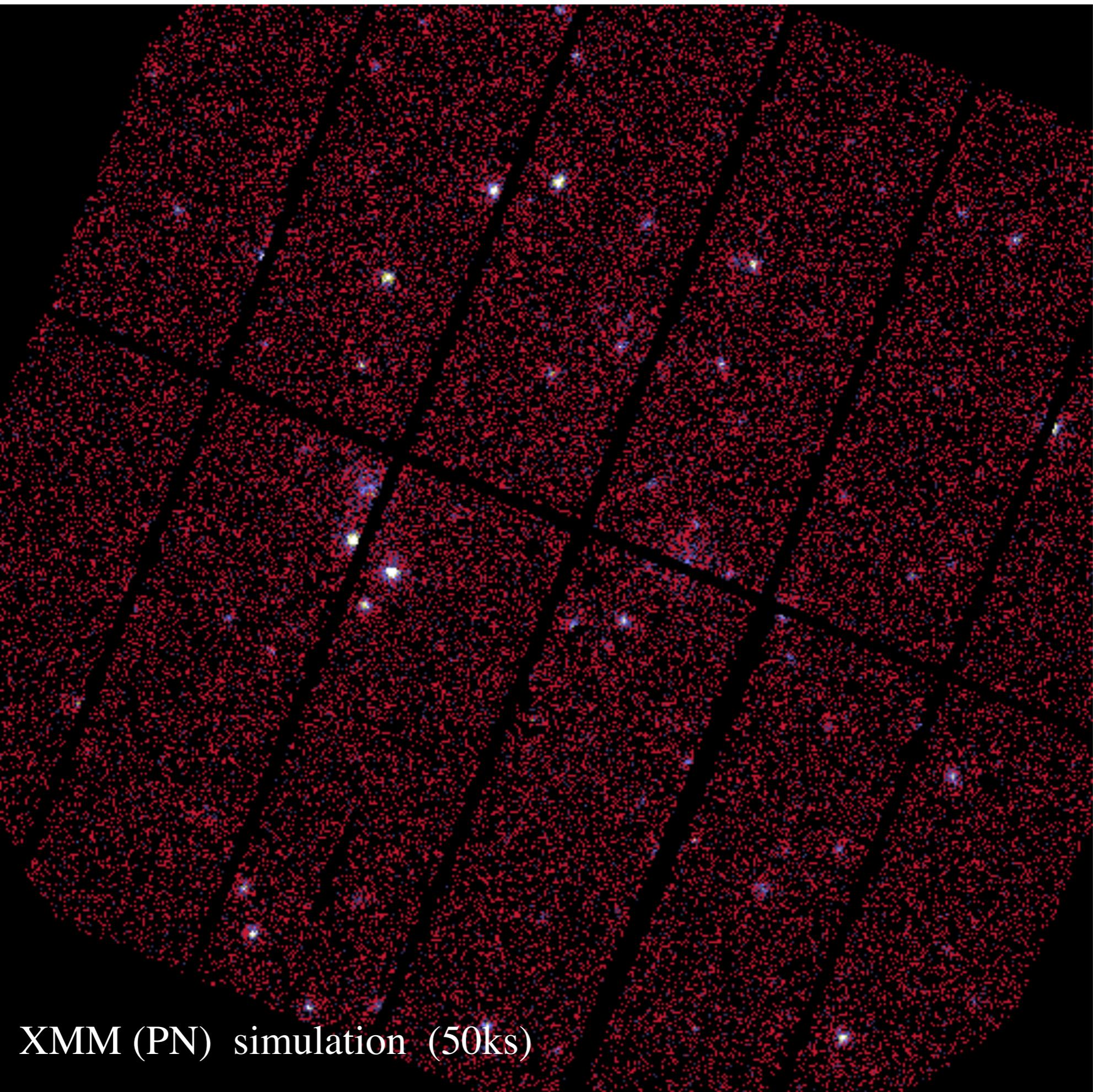
More **sophisticated tests** exist: False Discovery Rate (FDR), block sparsity, etc

Regularization = DETECTION STRATEGY

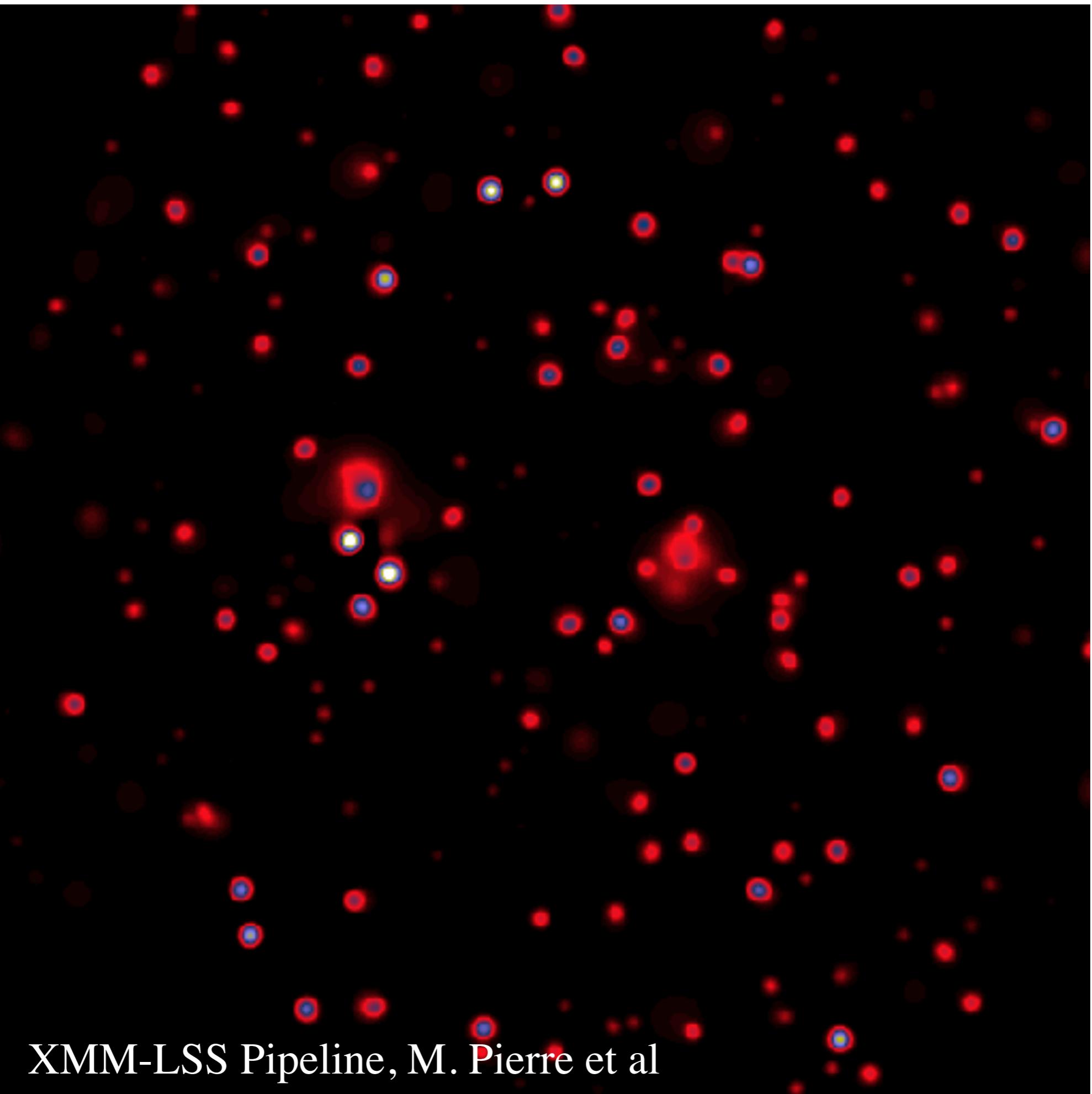
Non Gaussian Noise

The noise in the data follows a distribution law which can be:

- a White Gaussian Noise
- Correlated Noise
- a Poisson Noise
- a Poisson + Gaussian distribution (noise in the CCD)
- Poisson noise with few events (Galaxies counting, X ray images, ...)
- Speckle noise
- Root Mean Square map: we have a noise standard deviation of each data value.



XMM (PN) simulation (50ks)



XMM-LSS Pipeline, M. Pierre et al

Iterative thresholding methods were proposed initially in [F. and Nowak, 2003], [Daubechies, Defrise, De Mol, 2003], [Starck, Candès, Donoho, 2003]:

Synthesis:

$$\begin{aligned}\alpha^{(n+1)} &= \text{HardThresh}_\lambda \left(\alpha^{(n)} + \Phi^T H^T \left(Y - H\Phi\alpha^{(n)} \right) \right) \\ \alpha^{(n+1)} &= \text{SoftThresh}_\lambda \left(\alpha^{(n)} + \Phi^T H^T \left(Y - H\Phi\alpha^{(n)} \right) \right)\end{aligned}$$

Analysis:

$$\begin{aligned}X^{(n+1)} &= \Phi \text{HardThresh}_\lambda \left[\Phi^T \left(X^{(n)} + H^T \left(Y - HX^{(n)} \right) \right) \right] \\ X^{(n+1)} &= \Phi \text{SoftThresh}_\lambda \left[\Phi^T \left(X^{(n)} + H^T \left(Y - HX^{(n)} \right) \right) \right]\end{aligned}$$

IST can be seen as a generalization of projected gradient descent In the framework of **proximal theory** [Moreau 62], we have [Combettes and Wajs, 2005]:

$$\alpha^{n+1} = \text{prox}_{\mathcal{C}}(\alpha^n + \mu \Phi^t H^t (Y - H\Phi\alpha^n)).$$

$$X^{(n+1)} = \Phi \text{SoftThresh}_{\lambda} \left[\Phi^T \left(X^{(n)} + H^T \left(Y - HX^{(n)} \right) \right) \right]$$

At each iteration, the back-projected residual is added to the solution:

$$R^{(n)} = H^T \left(Y - HX^{(n)} \right)$$

In most cases, the noise in $R^{(n)}$ is not white Gaussian, and be quite complex:

The use of a **single** hyper parameter does not allow us to properly take into account the signal and noise behavior in different bands:

The use of a **single** hyper parameter does not allow us to properly take into account the signal and noise behavior in different bands:

Analysis: $\min_X \|Y - HX\|^2 + \lambda \|\Phi^t X\|_p^p$



$$\min_X \|Y - HX\|^2 + \sum_j \lambda_j \|\Phi_j^t X\|_p^p$$

Synthesis: $\min_{\alpha} \|Y - H\Phi\alpha\|^2 + \lambda \|\alpha\|_p^p$



$$\min_{\alpha} \|Y - H\Phi\alpha\|^2 + \sum_j \lambda_j \|\alpha_j\|_p^p$$

Signal Driven Strategy

Study the statistical distribution of the coefficient of a class of signal in the different bands (amplitude, decay, etc).

- Noise driven strategy from MC noise realizations

$$\alpha_j^{N^{(i)}} = \Phi_j^t H^T N^{(i)} \longrightarrow \lambda_j = k\sigma(\alpha_j^{N^{(i)}})$$

- Spatially variant noise

$$\alpha_{j,l}^{N^{(i)}} = \left(\Phi_j^t H^T N^{(i)} \right)_l \longrightarrow \lambda_{j,l} = k\sigma(\alpha_{j,l}^{N^{(i)}})$$

- Noise driven strategy from the residual

$$\alpha^{R^{(n)}} = \Phi^t H^T \left(Y - Hx^{(n)} \right) \longrightarrow \lambda_j = k\sigma(\alpha_j^{R^{(n)}})$$

Condat-Vu Primal-Dual Algorithm

IST can be seen as a generalization of projected gradient descent In the framework of **proximal theory** [Moreau 62], we have [Combettes and Wajs, 2005]:

$$\alpha^{n+1} = \text{prox}_{\mathcal{C}}(\alpha^n + \mu \Phi^t H^t (Y - H \Phi \alpha^n)).$$

Drawback: slow convergence, $O(1/n)$

MANY NEW ALGORITHMS: Vonesch et al, 2007; Elad et al 2008; Wright et al., 2008; Nesterov, 2008 and Beck-Teboulle, 2009; Blumensath, 2008; Maleki et Donoho, 2009, Starck et al, 2010, Raguet, Fadili, and Peyre, 2012; **Condat-Vu , 2013** ; etc.

$$\min_X \frac{1}{2} F(X) + \lambda \| \Phi^t X \|_1 \quad \text{with} \quad F(X) = \frac{1}{2} \| Y - HX \|_2^2$$

Primal-dual splitting: Convergence: $O(\frac{1}{n^2})$

$$\begin{cases} X^{(n+1)} &= X^{(n)} + \tau (\nabla F(X^{(n)}) + \Phi \alpha^{(n)}) \\ \alpha^{(n+1)} &= (\text{Id} - \text{ST}_\lambda) (\alpha^{(n+1)} + \Phi^t (2X^{(n+1)} - X^{(n)})) \end{cases}$$

adapted from Vu (2013)

But which NORM ?

Which Norm ? p in [0,1]

- ℓ_1 norm introduces a bias, and leads to very bad solution for astronomers.
- But ℓ_1 algorithms are more stable, with proof of convergence, etc

ℓ_1 reweighting scheme (Candes, 2007)

ℓ_0 can be approximated by solving a sequence of relaxed problems using the re-weighted ℓ_1 minimisation technique (Candes et al., 2008).
=> solving a sequence of weighted ℓ_1 problems of the form:

$$\min_X \frac{1}{2} \| Y - HX \|_2^2 + \lambda \| \mathbf{W}^\ell \Phi^t X \|_1$$

$$\alpha^{(\ell)} = \Phi^t X^{(\ell)}$$

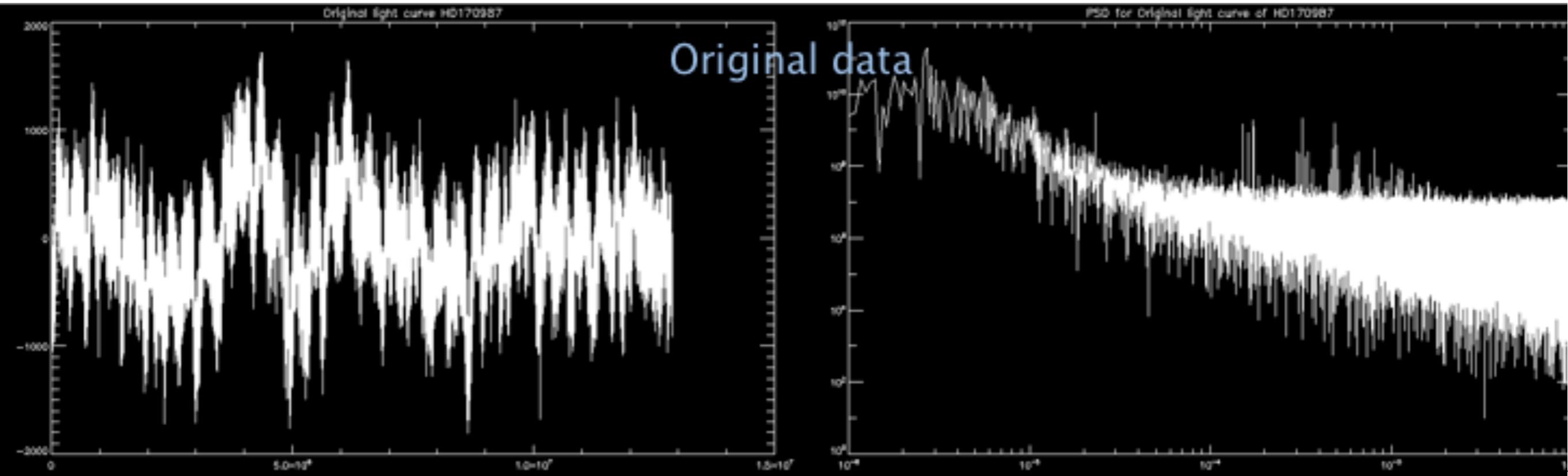
$$w_i^{(\ell+1)} = \begin{cases} \frac{w_i^{(0)}}{|\alpha_i^{(\ell)}|/\lambda w_i^{(0)}} & \text{if } |\alpha_i^{(\ell)}| \geq \lambda w_i^{(0)} \\ w_i^{(0)} & \text{if } |\alpha_i^{(\ell)}| < \lambda w_i^{(0)} \end{cases},$$

- 1 - Which Norm ? **ℓ_0 through re-weighted ℓ_1**
- 2 - Constraint versus Lagrangian formulation ? **Lagrangian**
- 3 - Analysis versus Synthesis ? **Analysis**
- 4 - Which dictionary ? **this is a prior, it depends on the problem**
- 5 - Which noise model ? **this depends on the experiment**
- 6 - Which minimization method ? **Condat-Vu algorithm is very efficient.**
- 7 - How to fix the regularization parameter ? **Physical interpretation of the regularization parameter, through the noise modeling => fully automatic approach.**

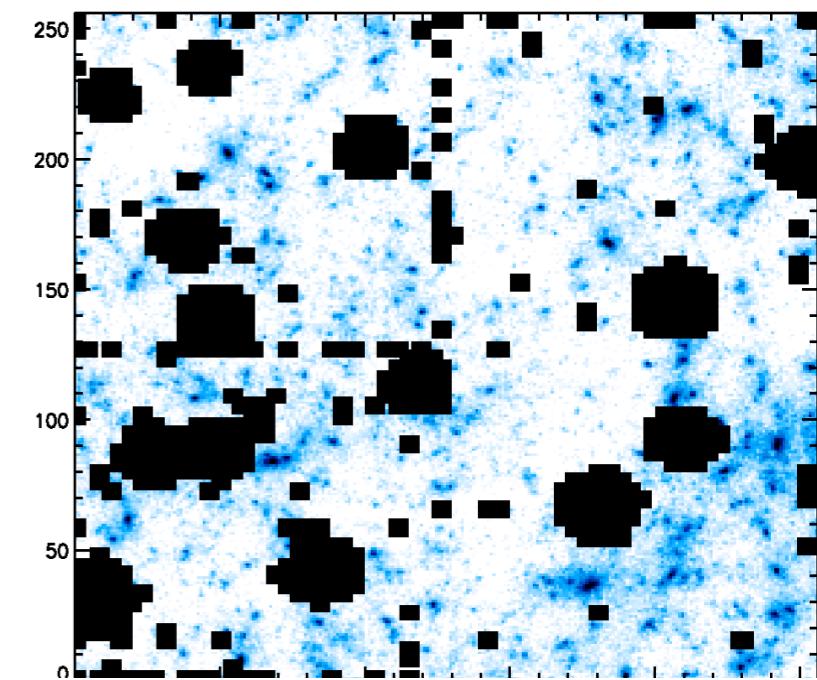
Missing Data

- Period detection in temporal series

COROT: HD170987



- Bad pixels, cosmic rays,
point sources in 2D images, ...





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", The Computer Journal, 52, 1, pp 64-79, 2009.*

$$\Theta_\Lambda = \text{Id}_\Lambda$$

$$\min_{\alpha} \|\alpha\|_{\ell_0} \text{ s.t. } y = Mx$$

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

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

Iterative Hard Thresholding with a decreasing threshold.

MCAlab available at: <http://www.greyc.ensicaen.fr/~jfadili>

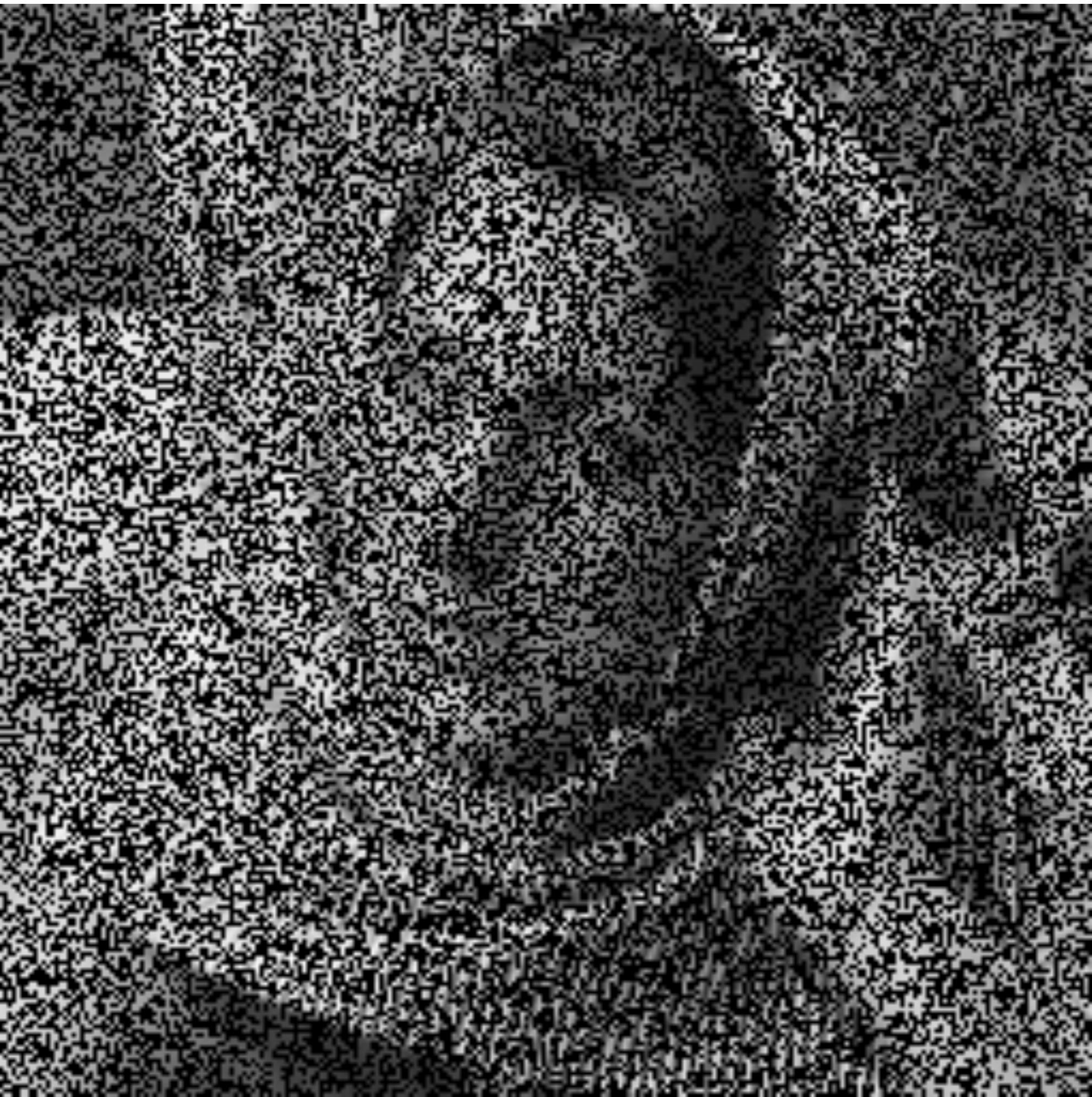
20%



20%



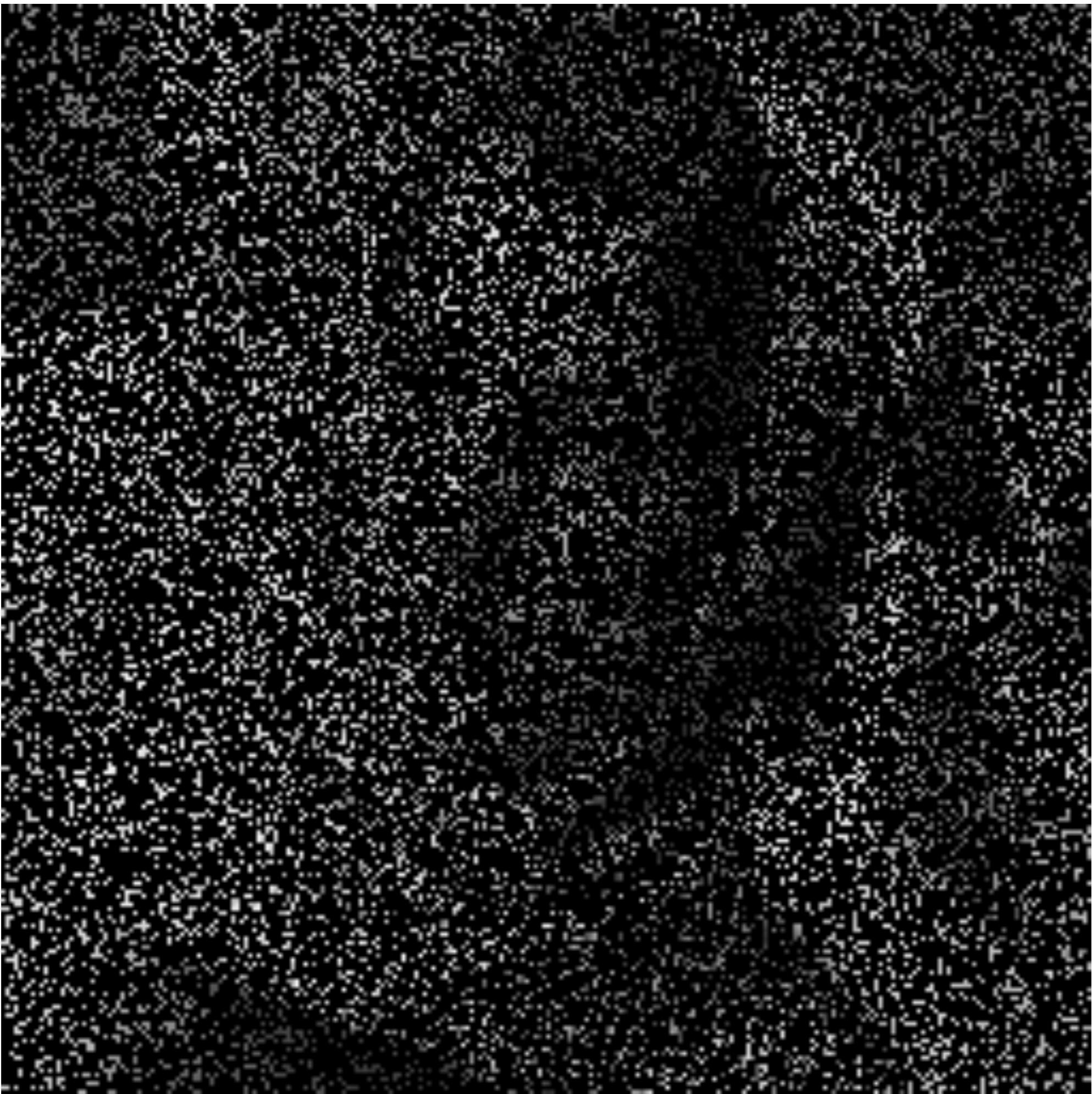
50%



50%



80%



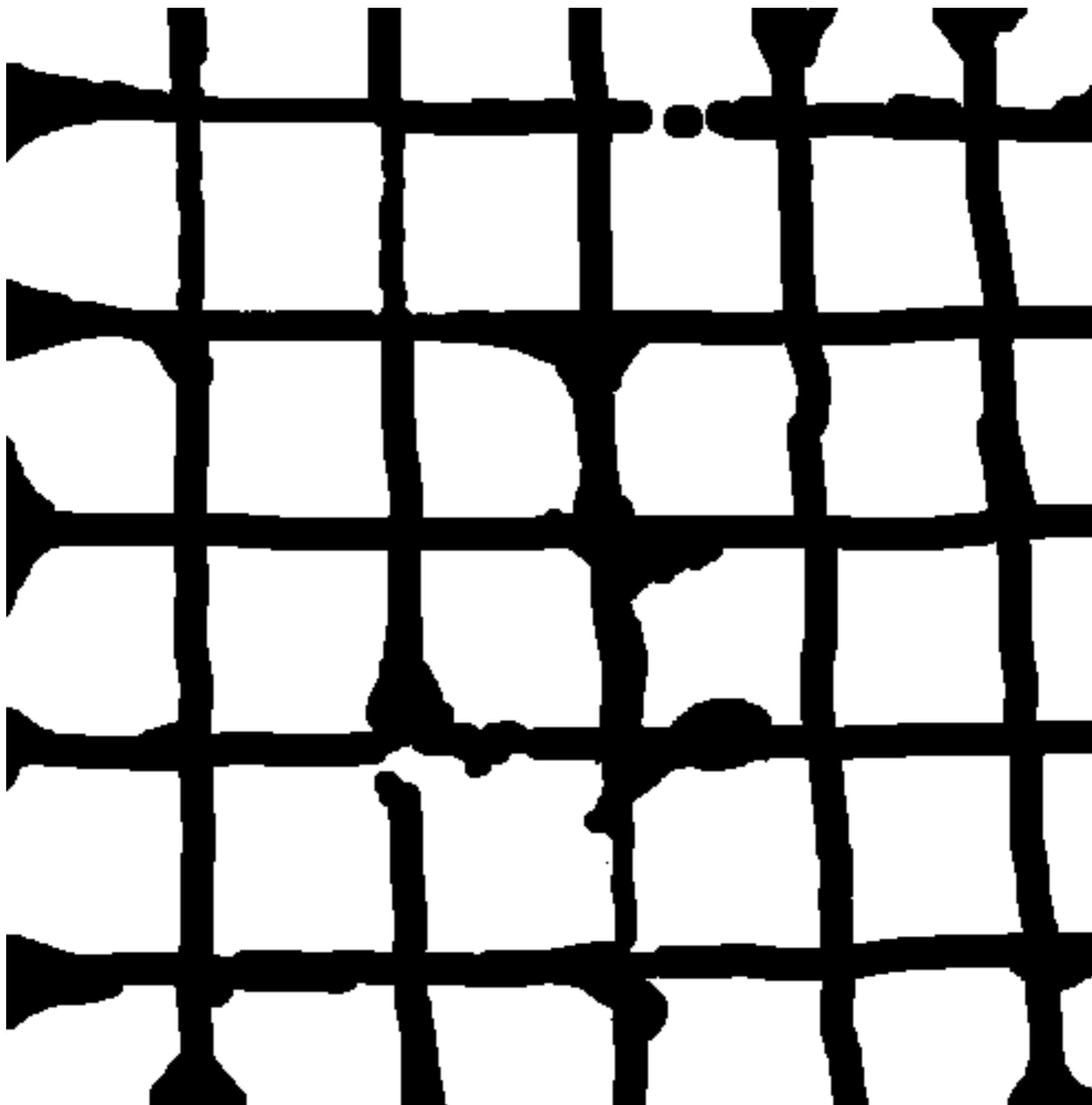
80%



Image inpainting [2, 10, 20, 38] is the process of restoring data in a designated region of a still or moving image. Applications range from removing objects from such as patching damaged paintings and photographs to produce a revised image in which the repair is seamlessly merged into the image in a way that is not detectable by a typical viewer. Traditionally, this has been done by professional artists.⁷ For photographs, inpainting is used to revert deterioration of photographs or scratches and dust spots in film. It can remove elements (e.g., removal of stamped date from photographs, the infamous “airbrushed enemies” [20]). A current active area of research is













Lenna masked (80% data missing)

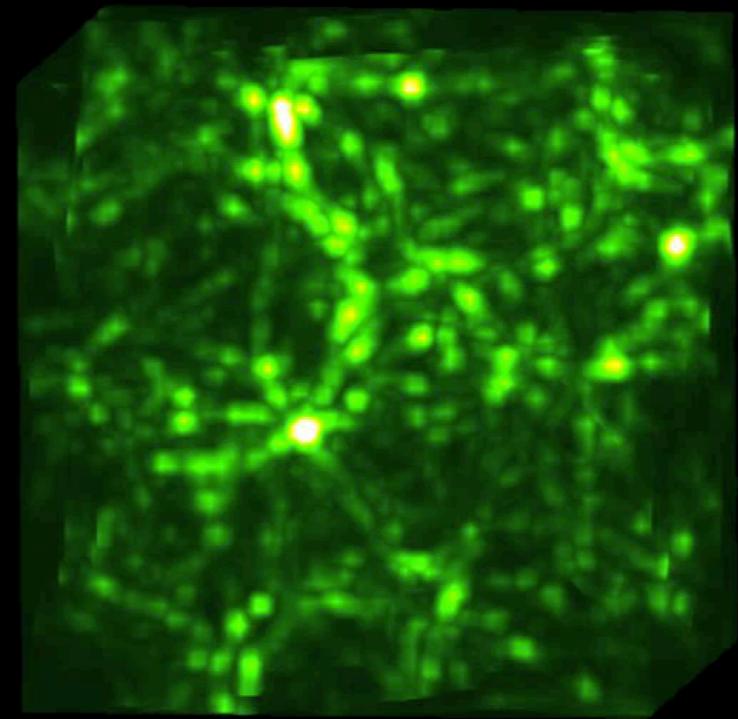
Inpainting is a typical restoration problem where the missing data must be estimated from incompletely observed ones. Here the EM algorithm is used as a general estimation framework with incomplete data. This example shows how it performs when a text must be removed from a useful informative part.

Inpainting is a typical restoration problem where the missing data must be estimated from incompletely observed ones. Here the EM algorithm is used as a general estimation framework with incomplete data. This example shows how it performs when a text must be recovered from a useful informative part.

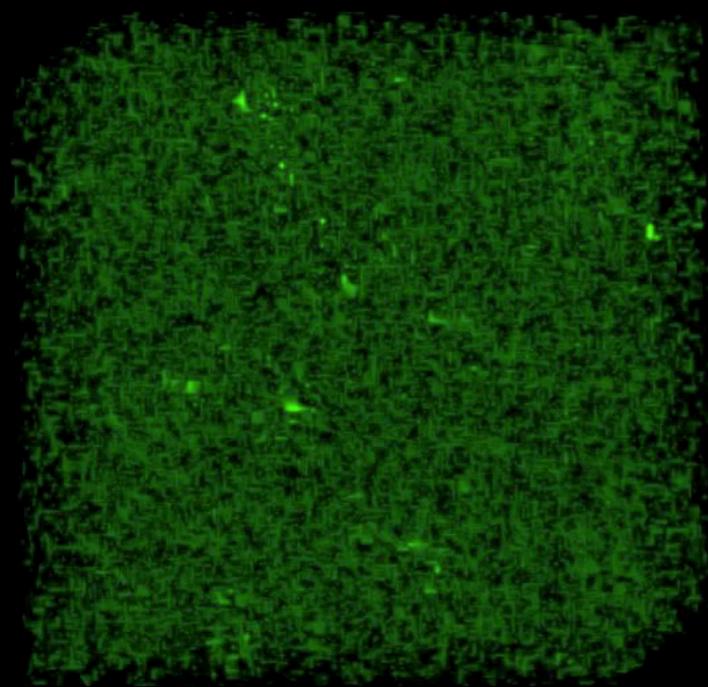
Inpainted with the curvelet dictionary (80% data missing)



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

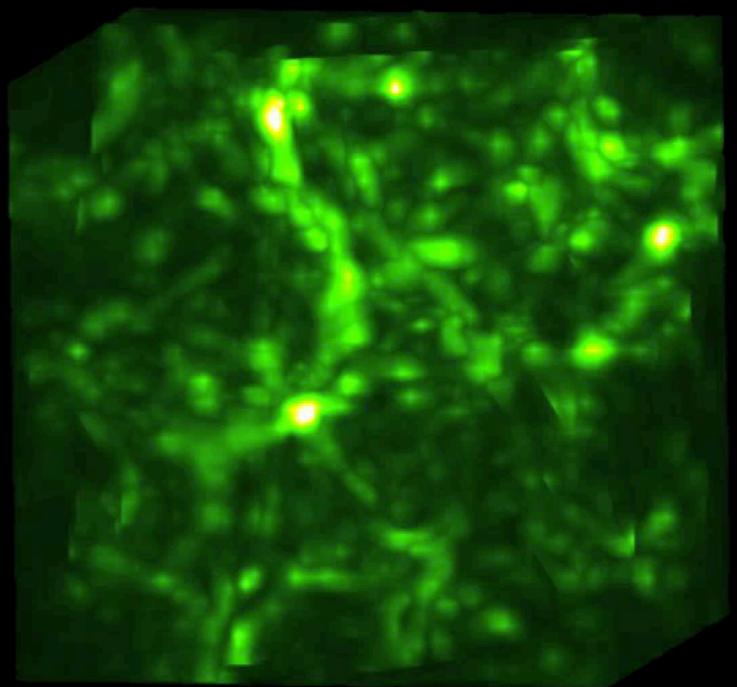


Original

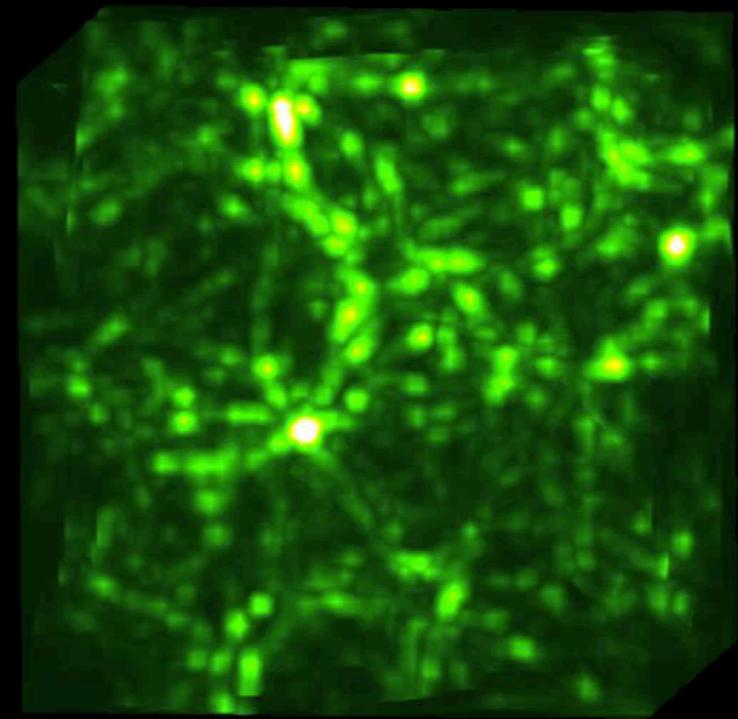


Mask

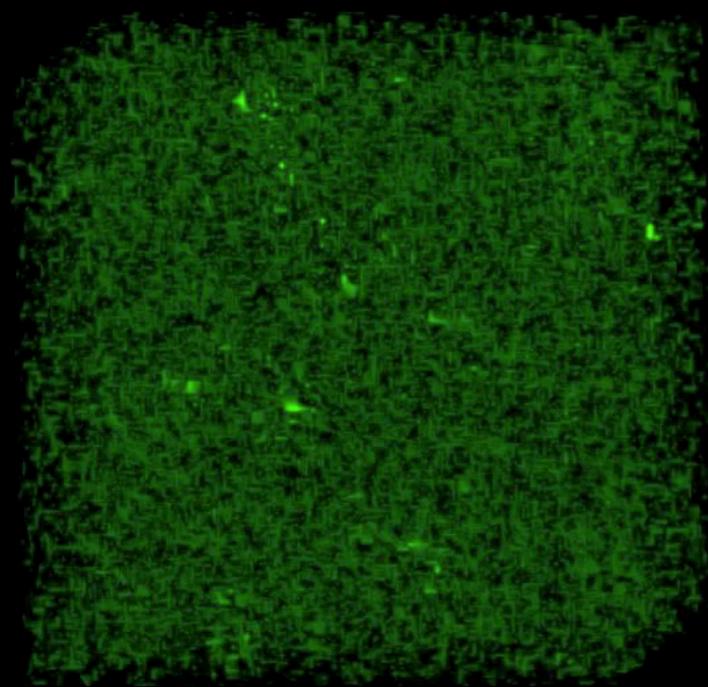
Dictionary
BeamCurvelets



Inpainted

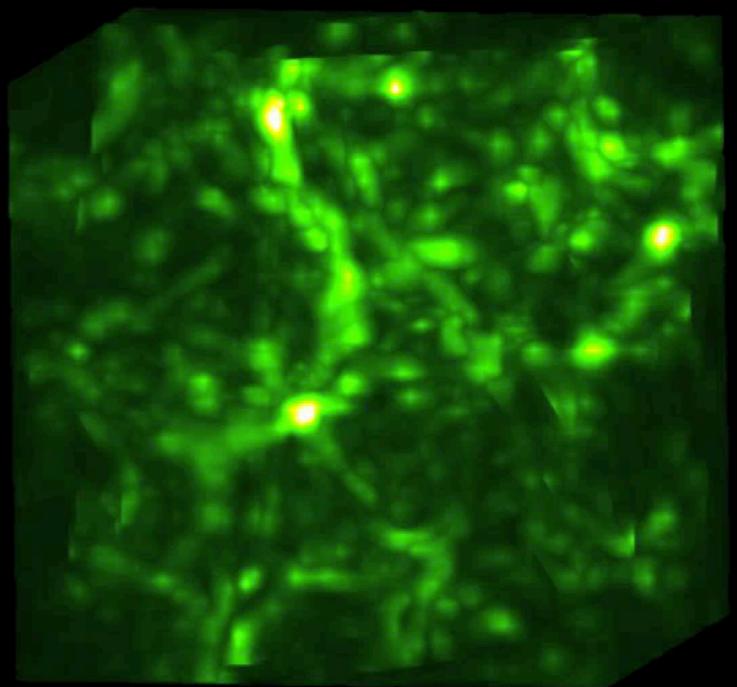


Original



Mask

Dictionary
BeamCurvelets



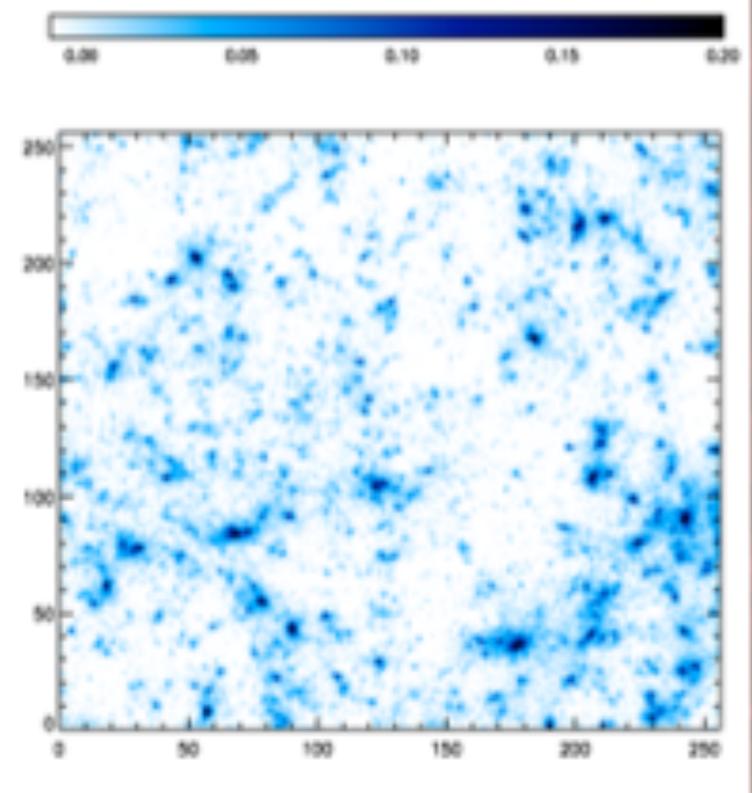
Inpainted

Inpainting :

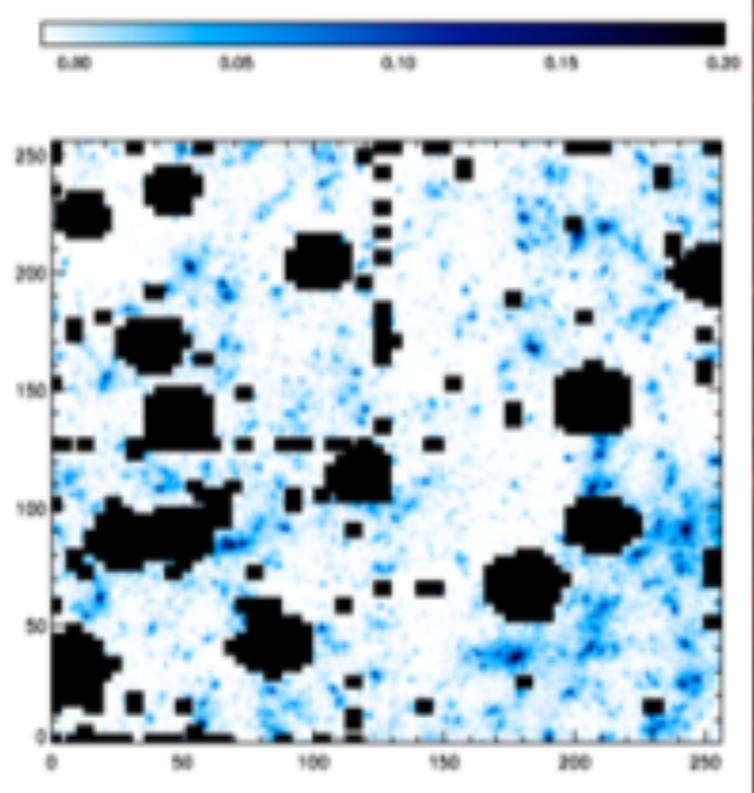
S. Pires, J.-L. Starck, A. Amara, R. Teyssier, A. Refregier and J. Fadili, "FASTLens (FAst STatistics for weak Lensing) : Fast method for Weak Lensing Statistics and map making", MNRAS, 395, 3, pp. 1265-1279, 2009.



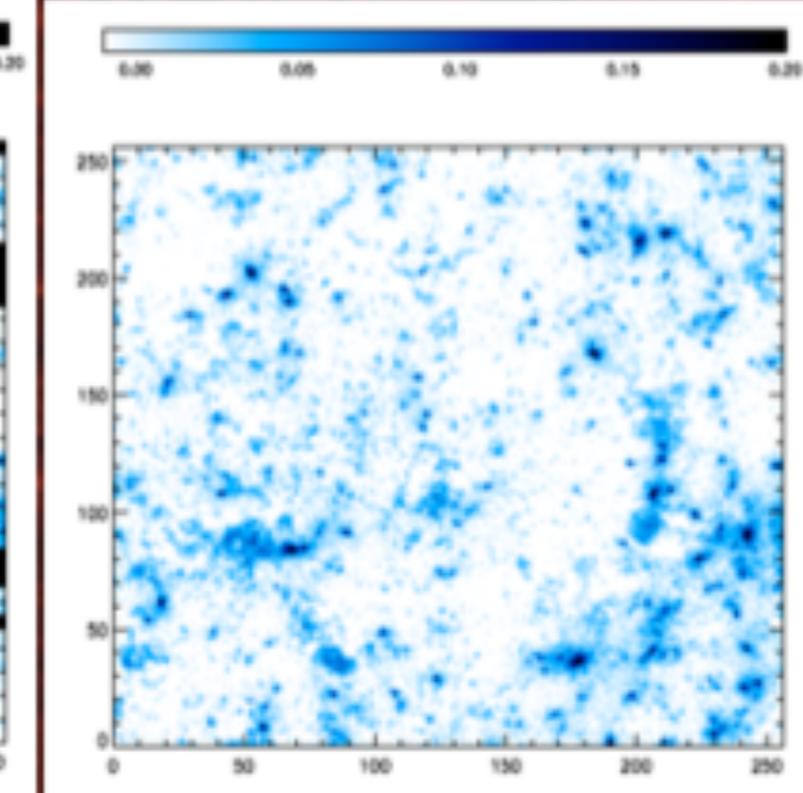
Original map



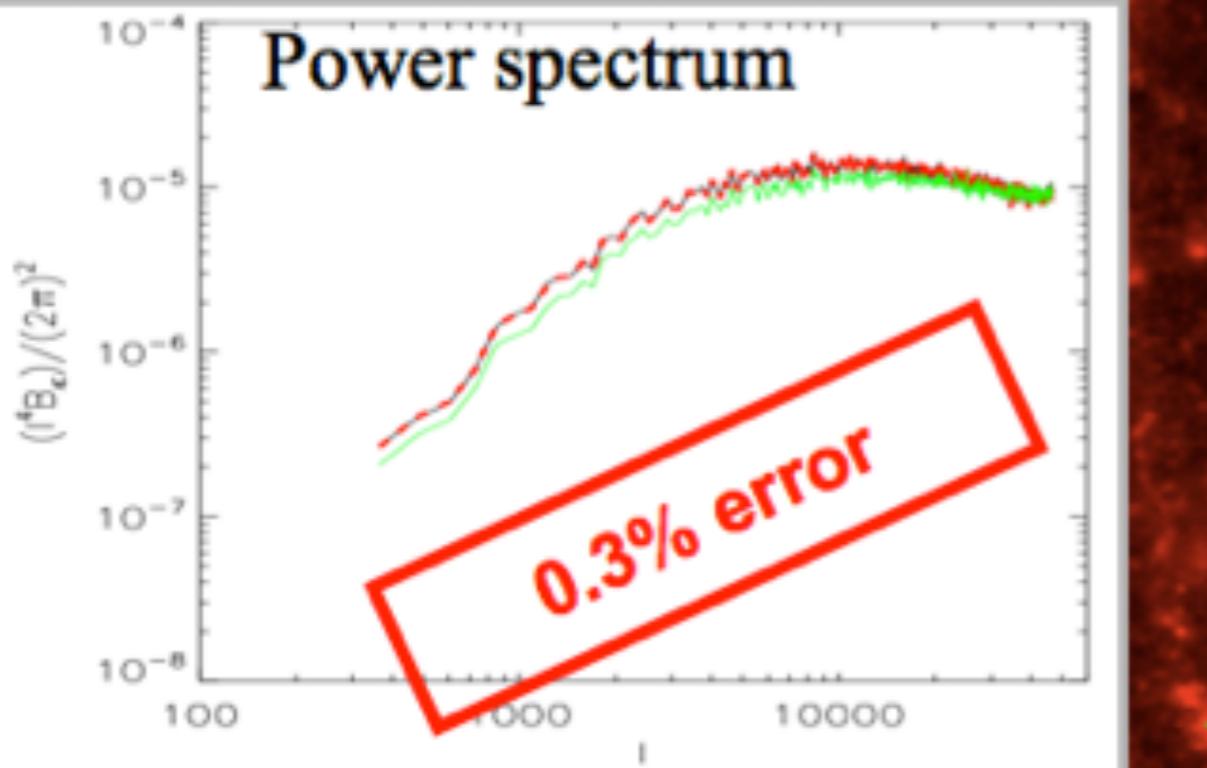
Masked map



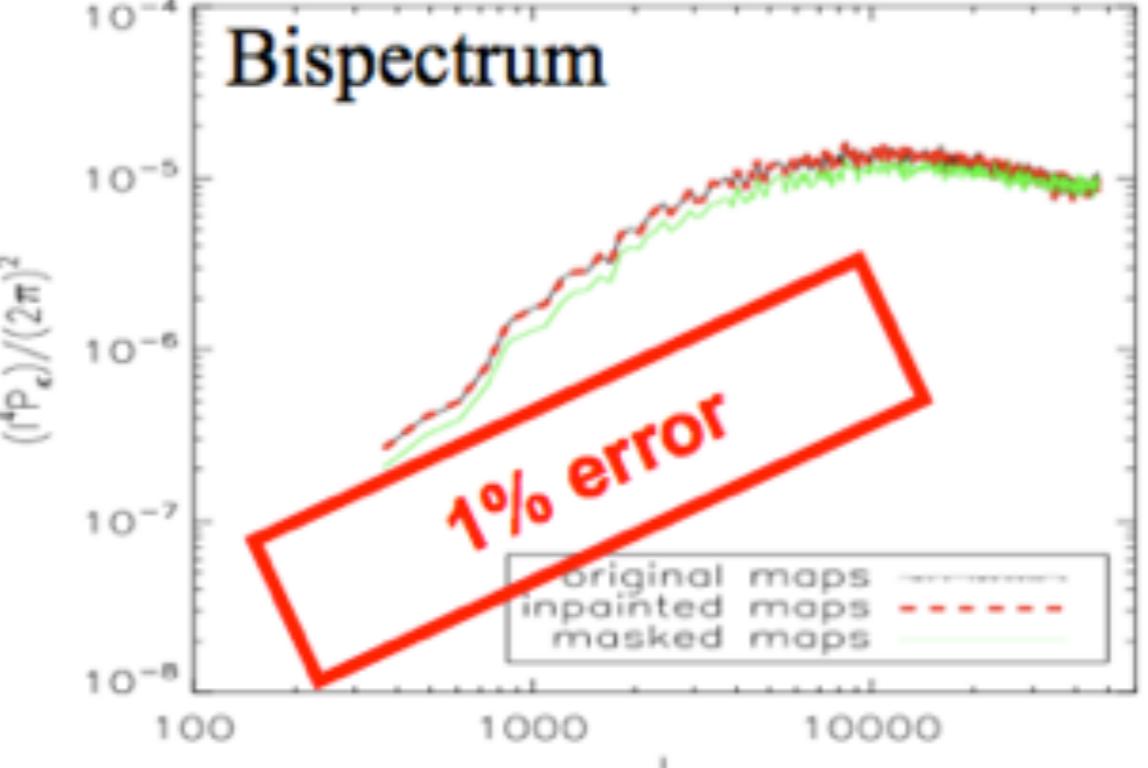
Inpainted map



Power spectrum

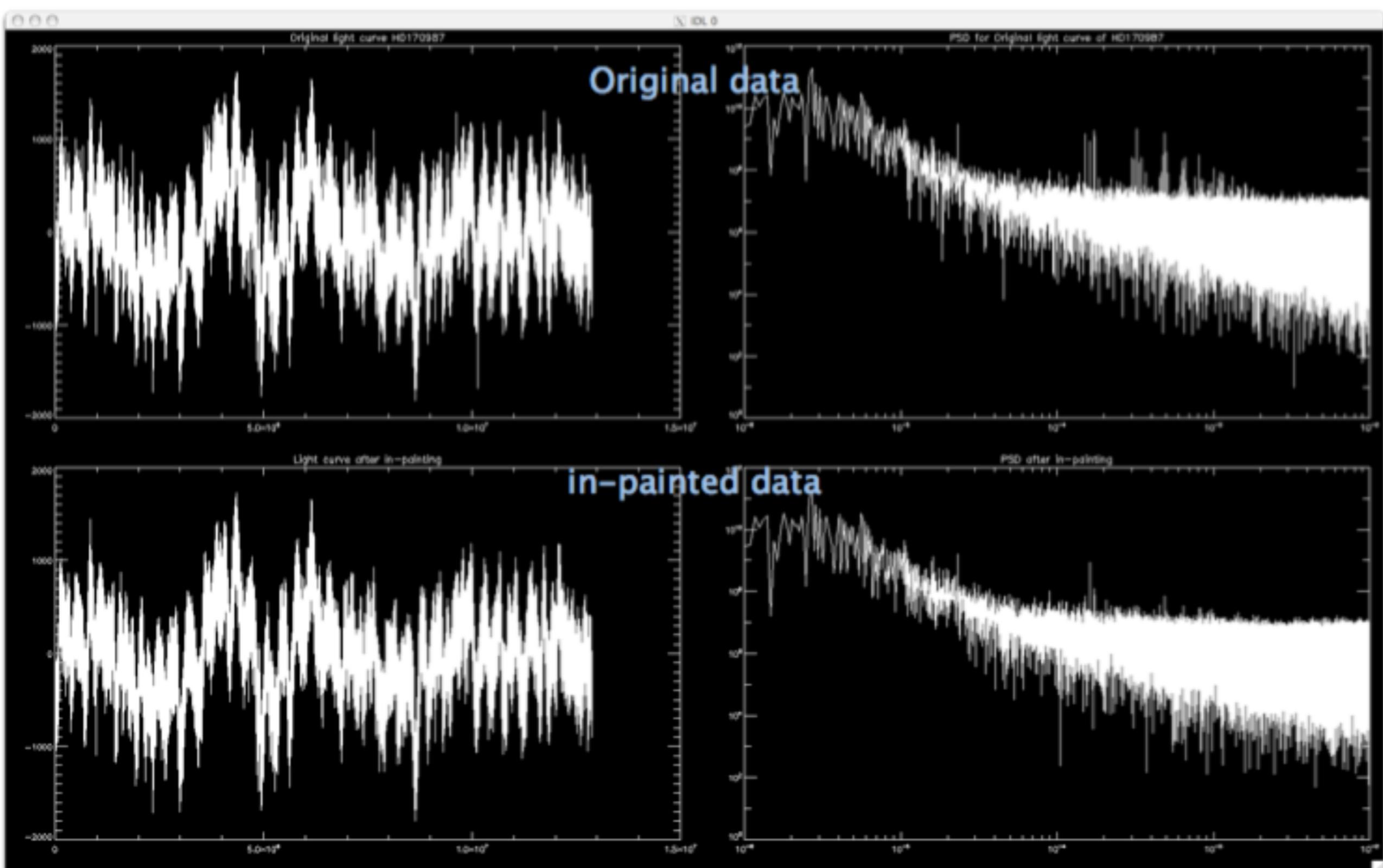


Bispectrum



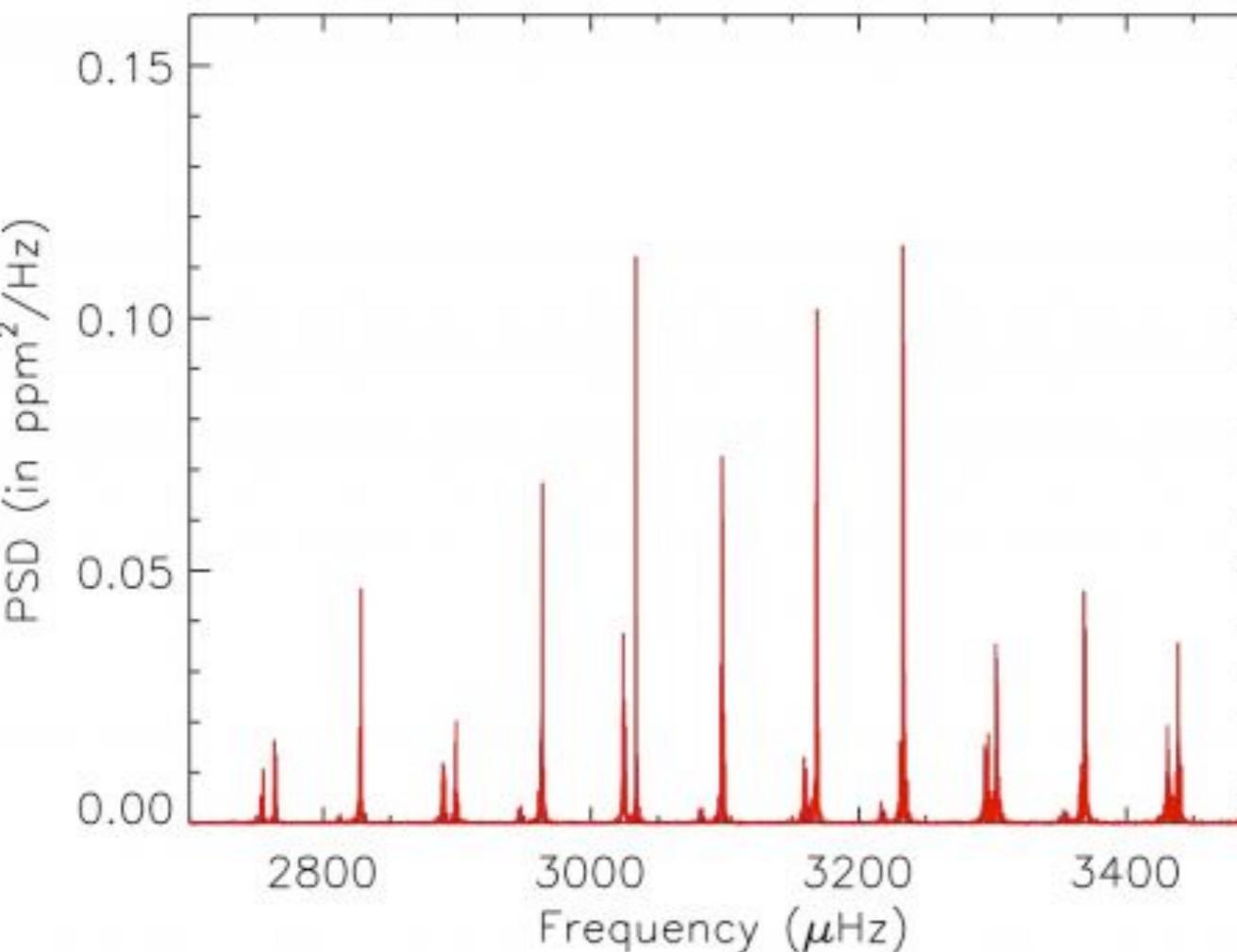
COROT: HD170987 with in-painting

[arXiv:1003.5178](https://arxiv.org/abs/1003.5178)



Sparse inpainting & asteroseismology

Gap interpolation by Inpainting methods: Application to Ground and Space-based data, S. Pires, S. Mathur, R.A. Garcia, J. Ballot, D. Stello and K. Sato, Astronomy and Astrophysics, in press.



CoRoT: sparse inpainting is in the official pipeline.
Kepler: 18.000 stars have been processed.
GOLF: ongoing tests

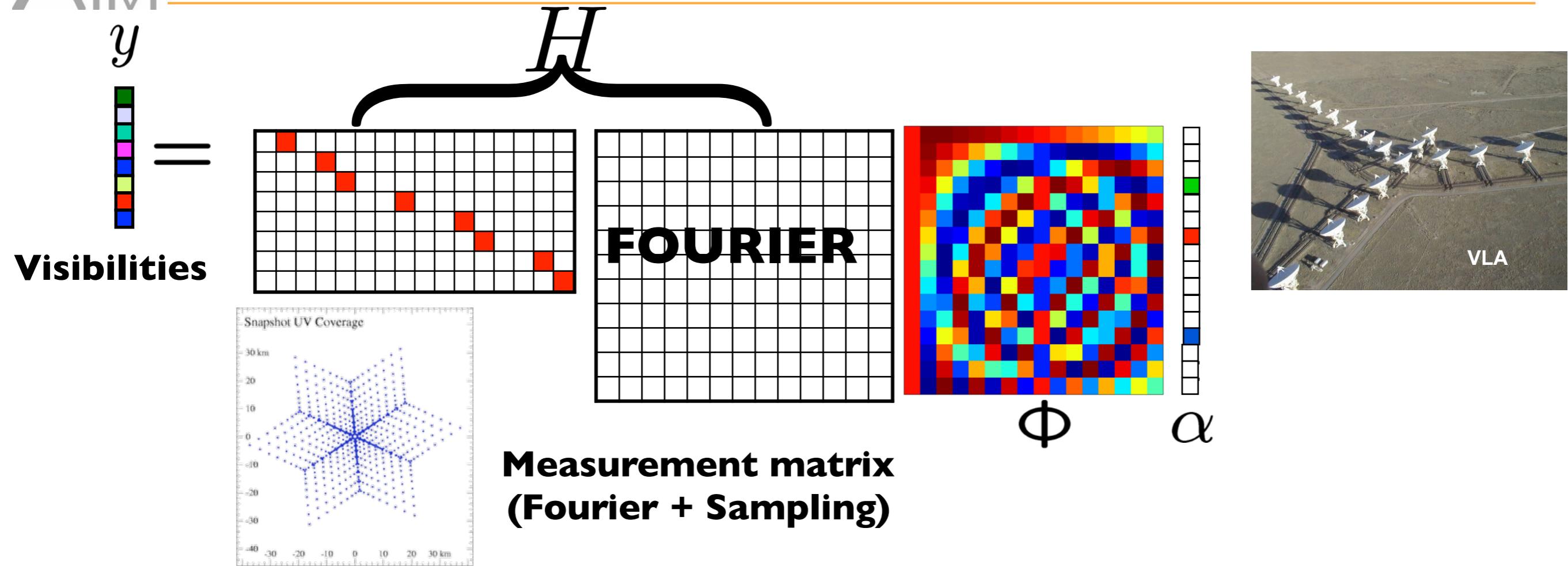


SOFTWARE K-INPAINTING : INPAINTING FOR KEPLER

[S. Pires](#), [R. A. Garcia](#), [S. Mathur](#), [J. Ballot](#)

www.cosmostat.org/software.html

Radio-Interferometry Sparse Recovery



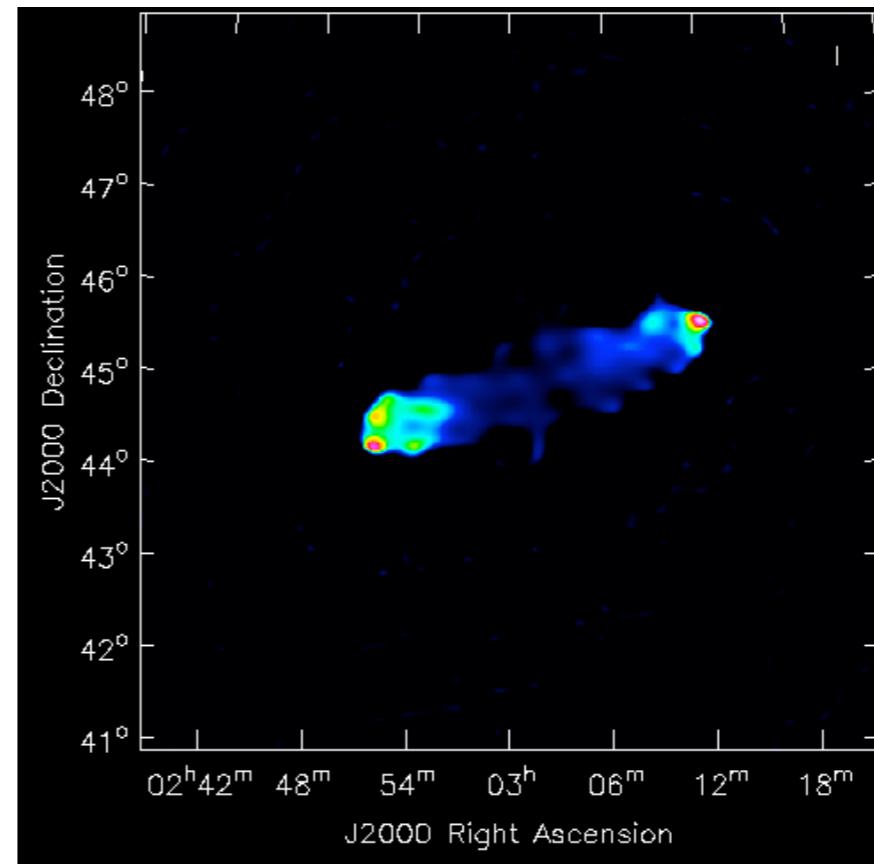
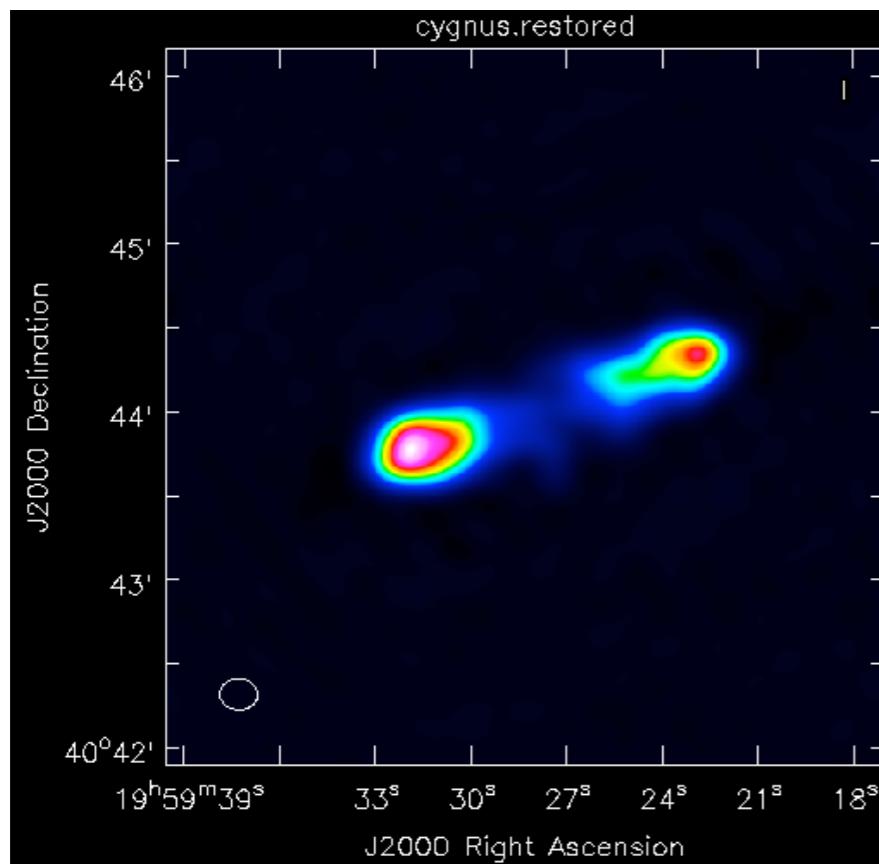
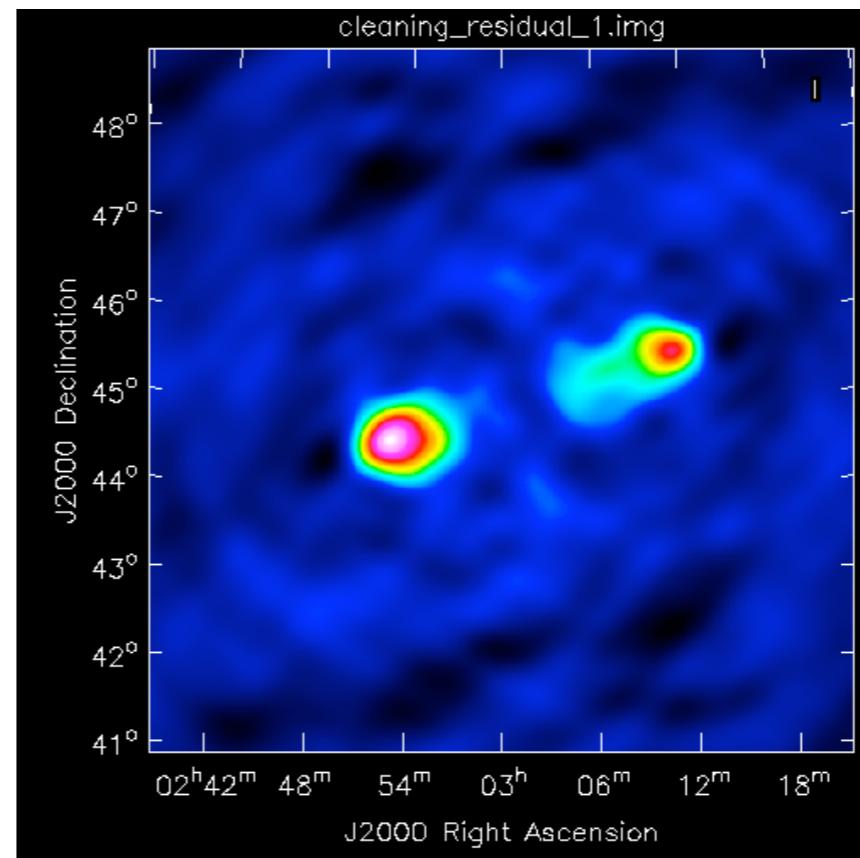
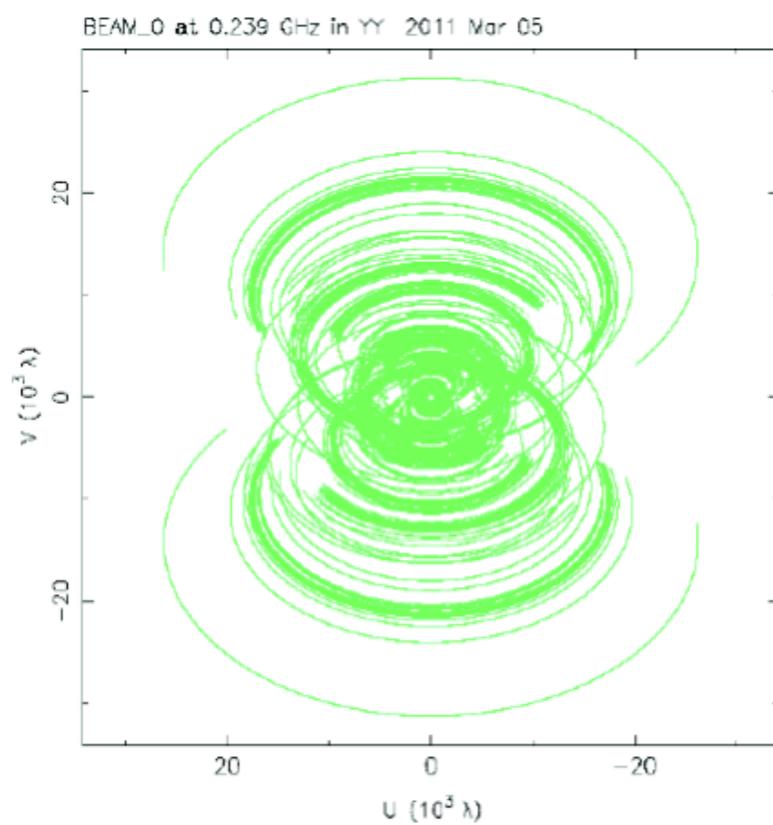
$$Y = HX + N$$

$$X = \Phi\alpha$$

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

- Photometry: similar to CLEAN on point sources.
- Resolution: improved by a factor larger than 2 for $\text{SNR} > 10$.
- Extended objects reconstruction much better than CLEAN and Multiscale CLEAN.
- Improved image quality (RMS better by a factor of 10 compared to CLEAN)

Compressed Sensing & LOFAR Cygnus A Data





J. Girard



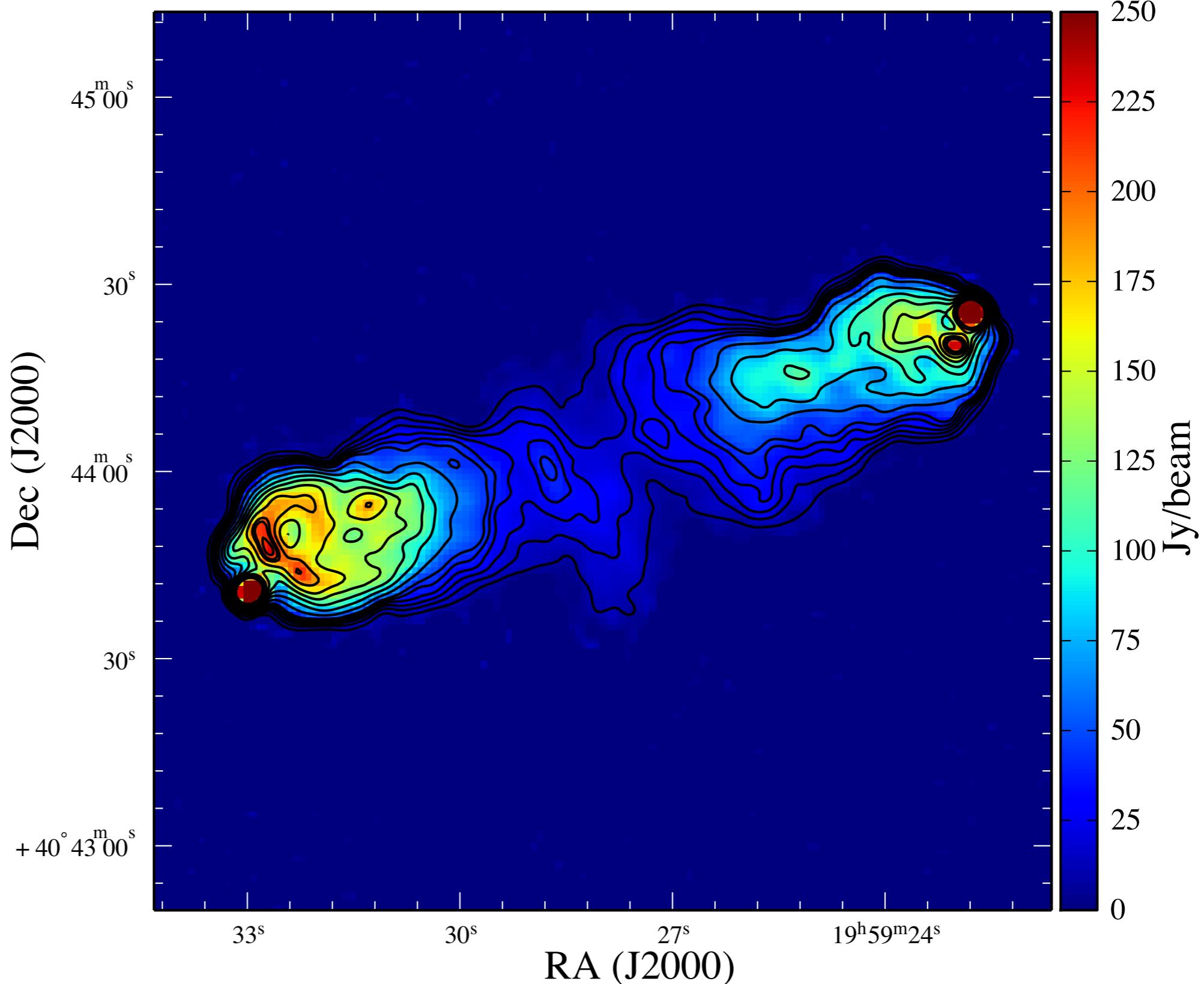
H. Garsden



S. Corbel



C. Tasse

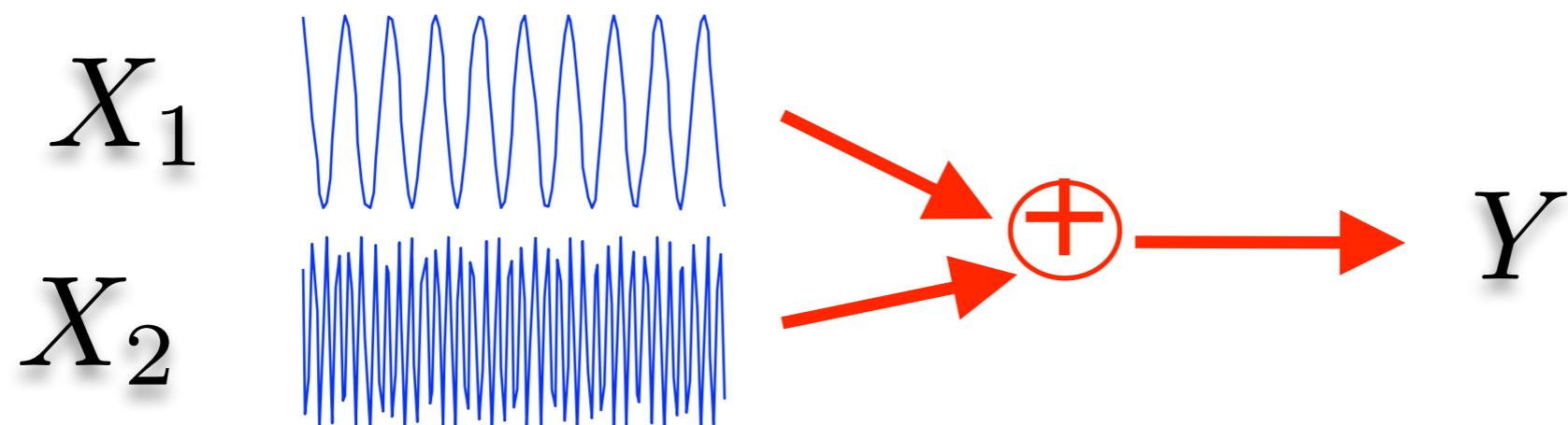


Colorscale: reconstructed 512x512 image of Cygnus A at 151 MHz (with resolution 2.8" and a pixel size of 1"). Contours levels are [1,2,3,4,5,6,9,13,17,21,25,30,35,37,40] Jy/Beam from a 327.5 MHz Cyg A VLA image (Project AK570) at 2.5" angular resolution and a pixel size of 0.5".
Recovered features in the CS image correspond to real structures observed at higher frequencies.

- Part I: Introduction to Inverse Problems in Astrophysics
- Part II: Sparsity & Dictionaries
- Part III: Sparse Regularization
- Part IV: Unmixing**
- Part V: Sparsity for Planck and Euclid Space Missions

- Mono-channel mixture:

$$Y = X_1 + X_2 + N$$

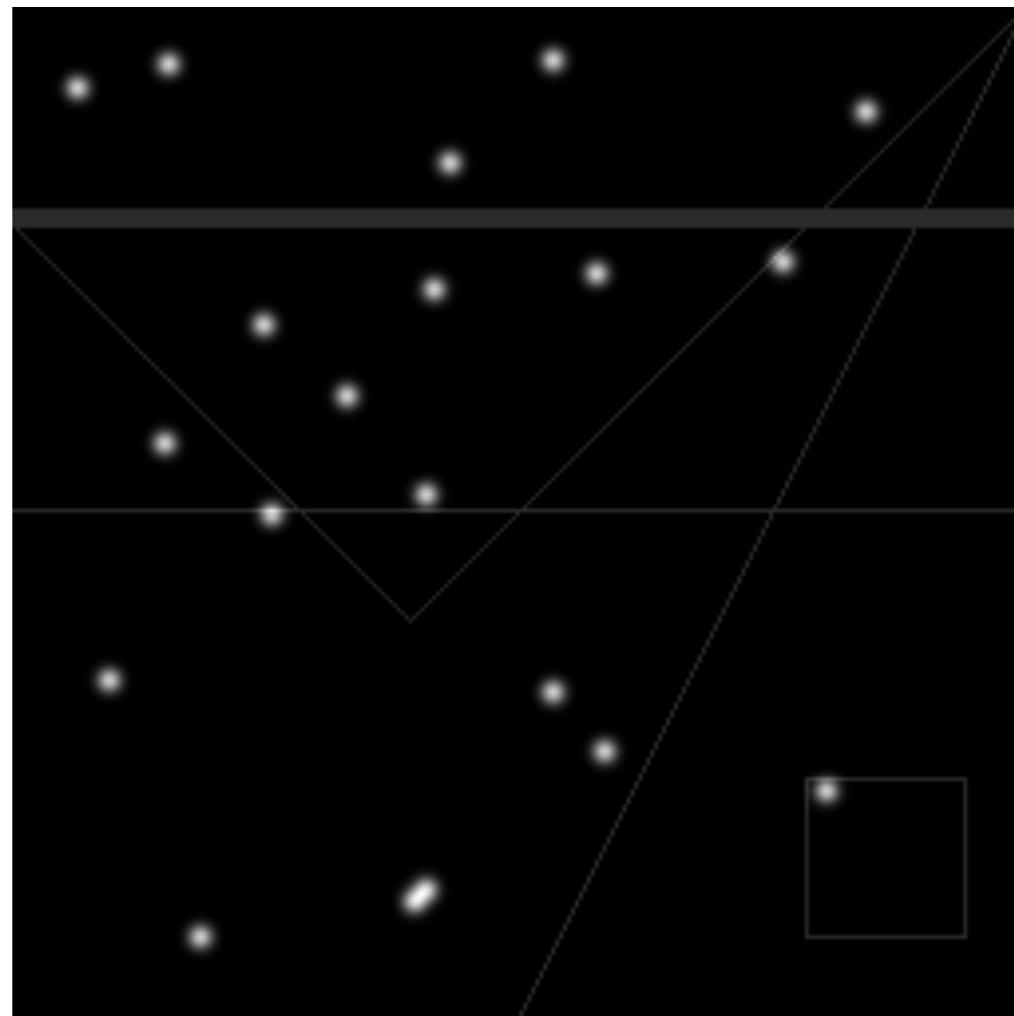


- Hyper/Multispectral mixture:

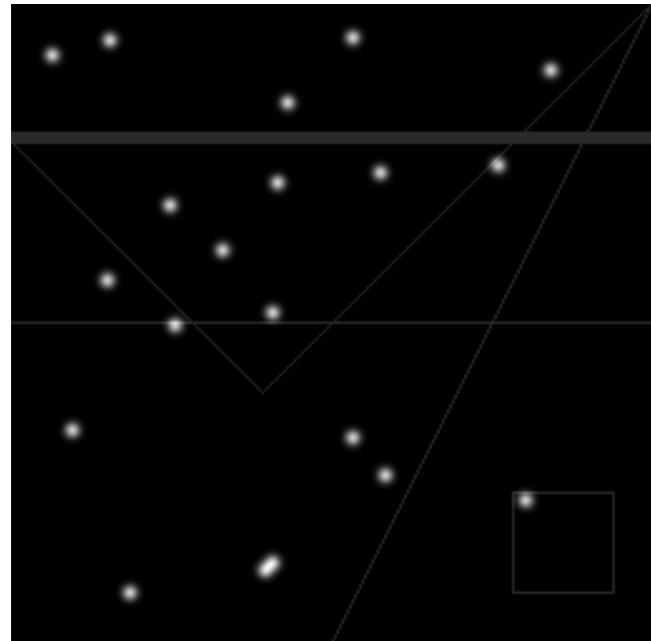
$$Y_i = H_i * \sum_{s=1}^S a_{i,s} X_s + N$$

A difficult issue

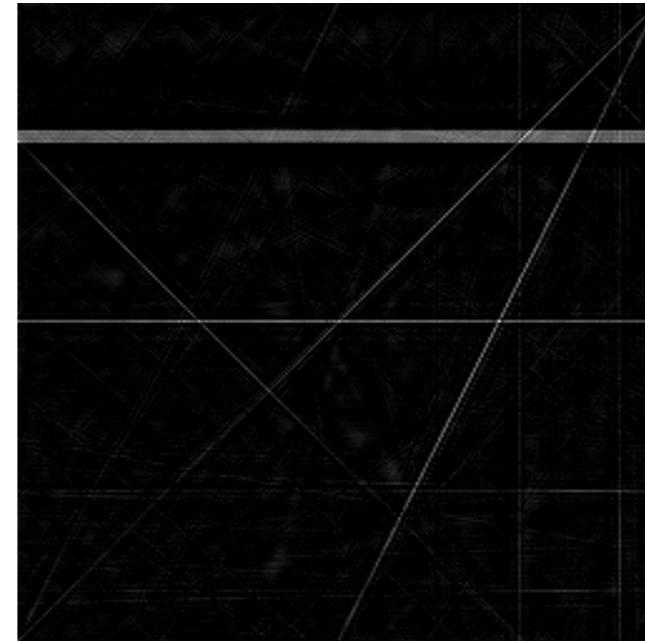
Is there any representation that well represents the following image ?



Going further

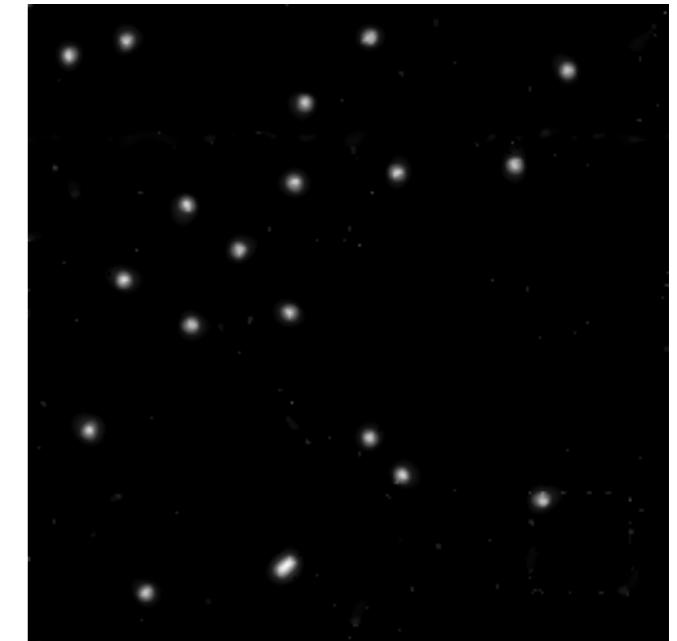


=



Lines

+



Gaussians

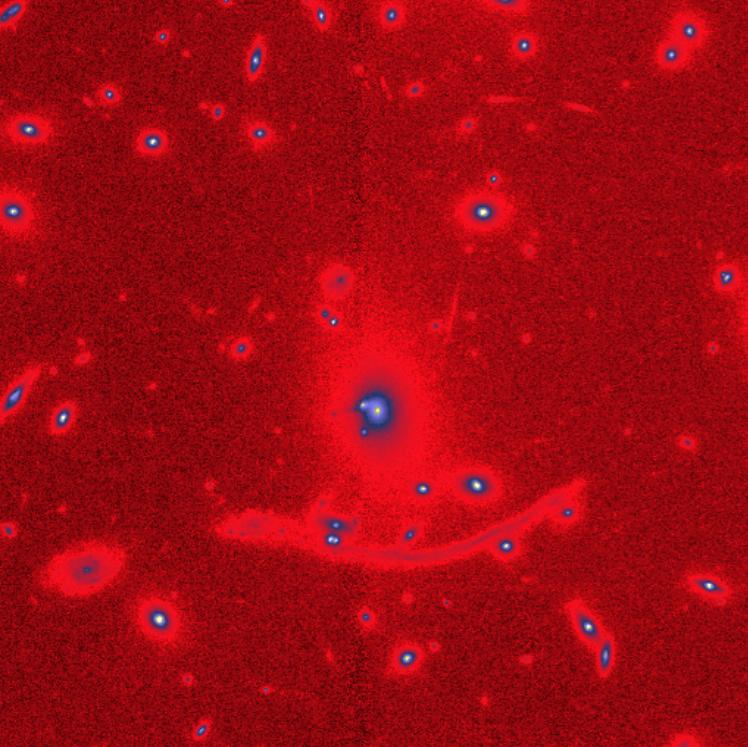


Curvelets



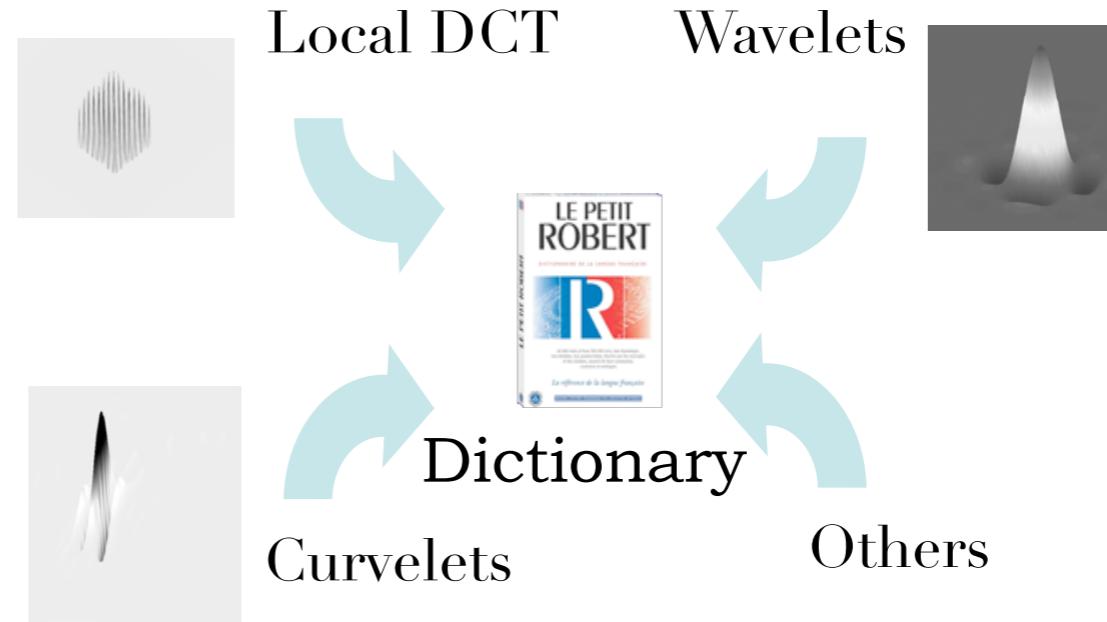
Wavelets

Redundant Representations



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.



$$\phi = [\phi_1, \mathcal{K}, \phi_L], \quad \alpha = \{\alpha_1, \mathcal{K}, \alpha_L\}, \quad s = \phi\alpha = \sum_{k=1}^L \phi_k \alpha_k$$

Model:

$$s = \sum_{k=1}^L s_k + n$$

and s_k ($s_k = \phi_k \alpha_k$) is sparse in ϕ_k .

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 :

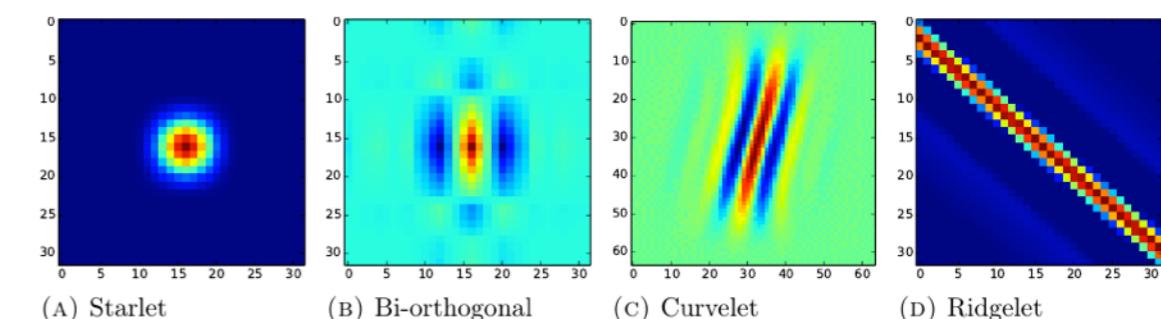
$$Y = X_1 + X_2$$

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.

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

$$C_1(X_1) = \| \Phi_1 X_1 \|_1$$

$$C_2(X_2) = \| \Phi_2 X_2 \|_1$$



$$\min_X \parallel Y - \sum_{k=1}^L x_k \parallel^2 + \lambda \sum_{k=1}^L \parallel \phi_k^* x_k \parallel_p$$

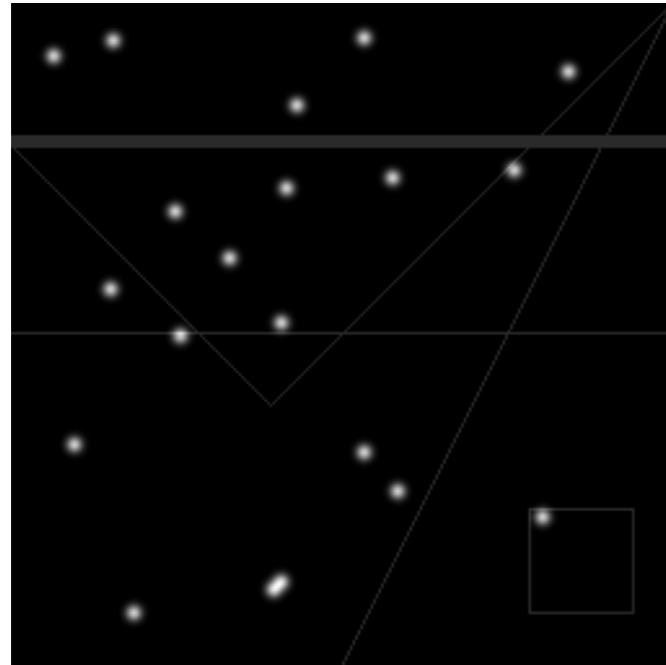
- . Initialize all x_k to zero
- . Iterate $j=1, \dots, N_{\text{iter}}$
 - Iterate $k=1, \dots, L$

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

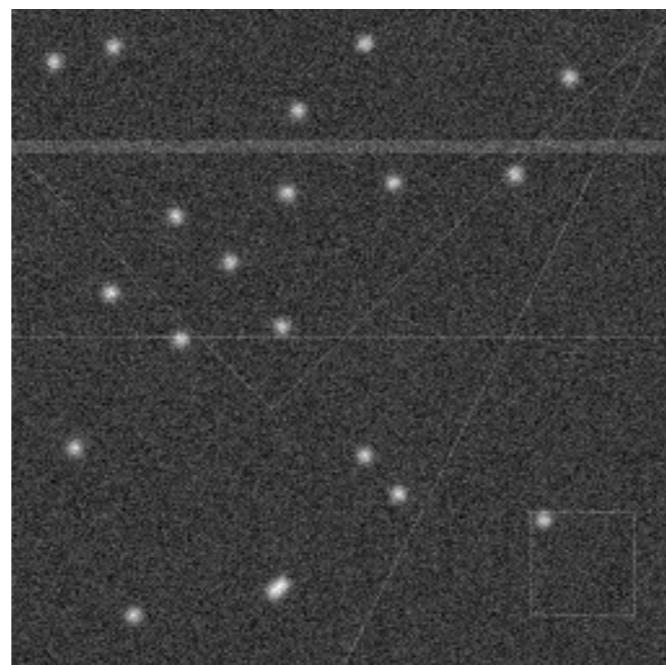
$$\min_{x_k} \parallel Y - \sum_{i=1, i \neq k}^L x_i - x_k \parallel^2 + \lambda \parallel \phi_k^* x_k \parallel_p$$

- Decrease the threshold $\lambda^{(j)}$
Which is obtained by a simple **hard**/soft thresholding of :

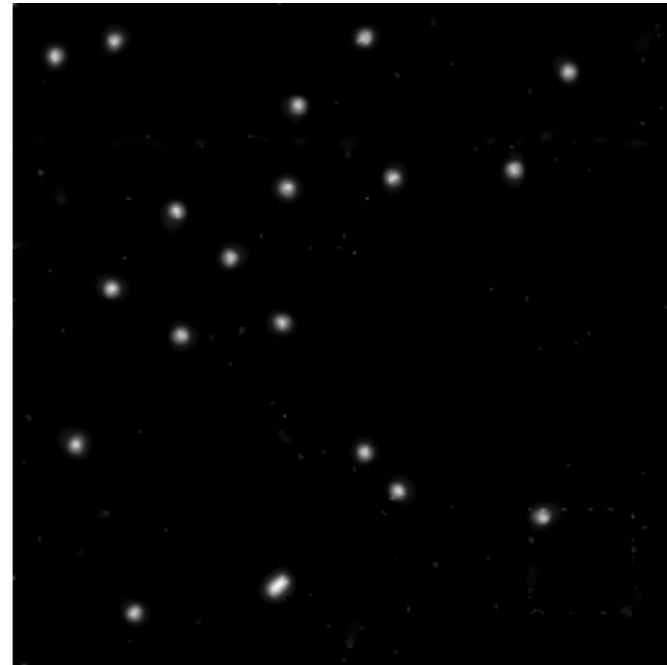
$$MIN_{s_1, s_2} (\|W s_1\|_p + \|C s_2\|_p) \quad \text{subject to} \quad \|s - (s_1 + s_2)\|_2^2 < \varepsilon$$



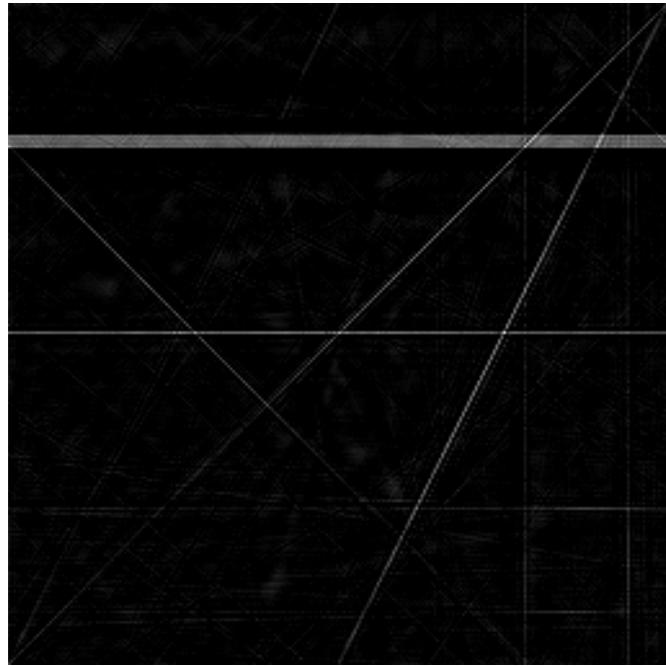
a) Simulated image (gaussians+lines)



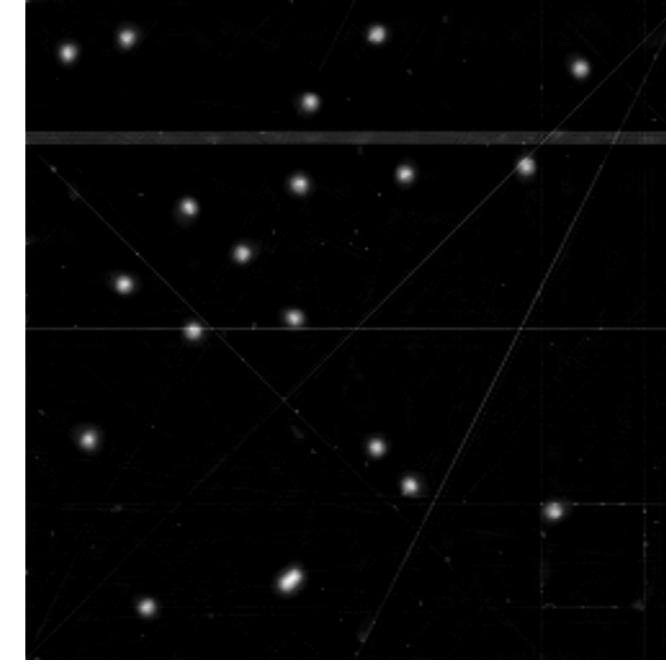
b) Simulated image + noise



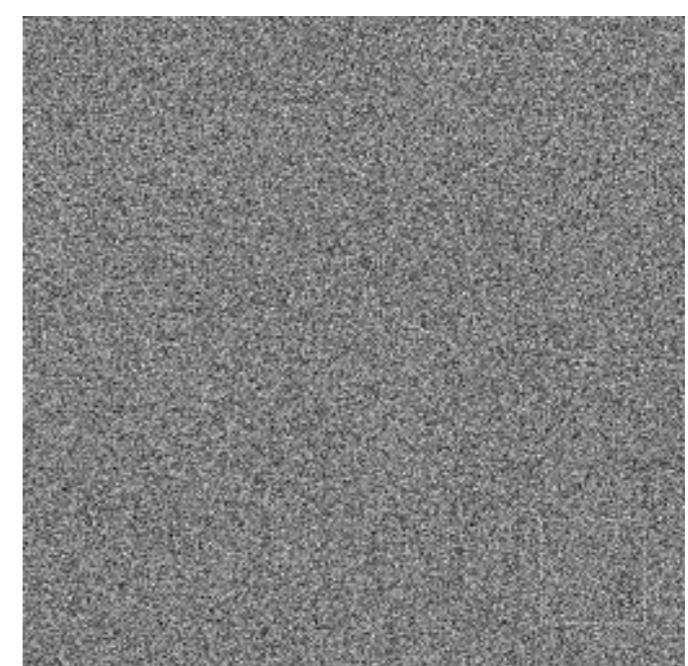
c) A trous algorithm



d) Curvelet transform

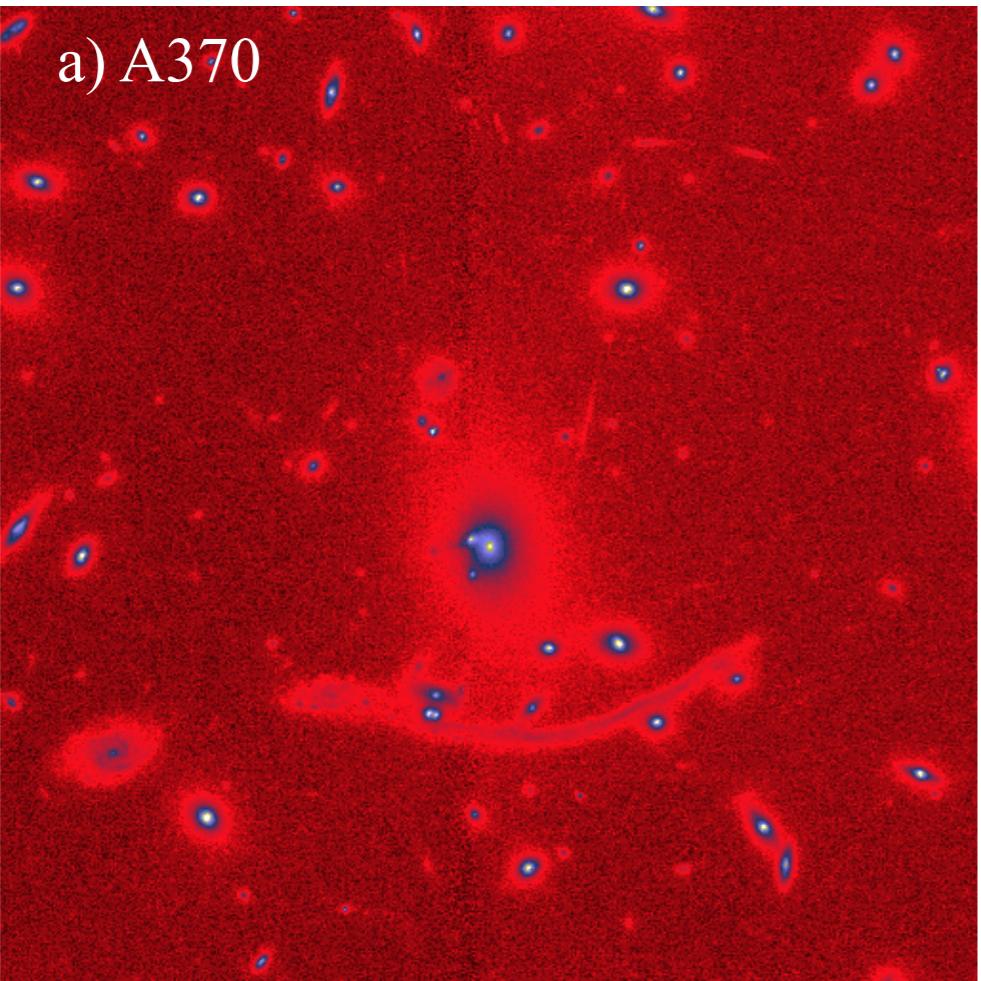


e) coaddition c+d

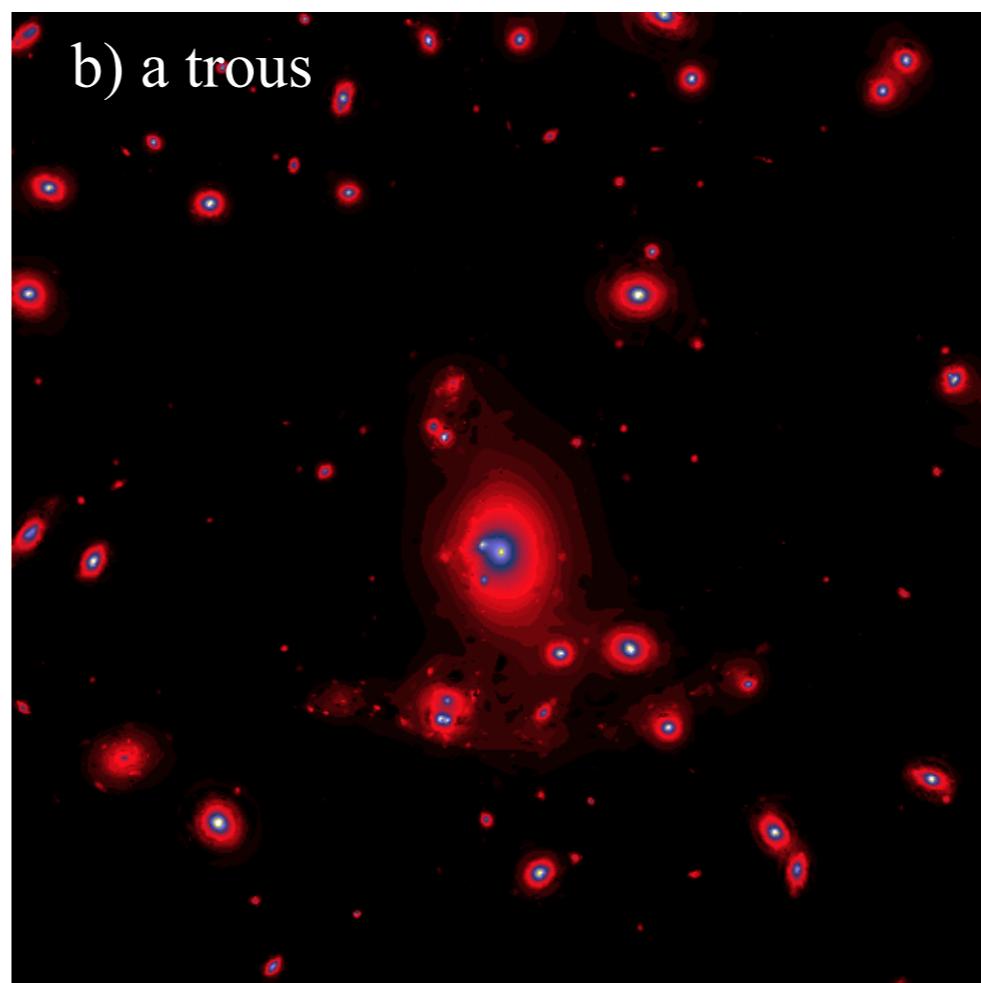


f) residual = e-b

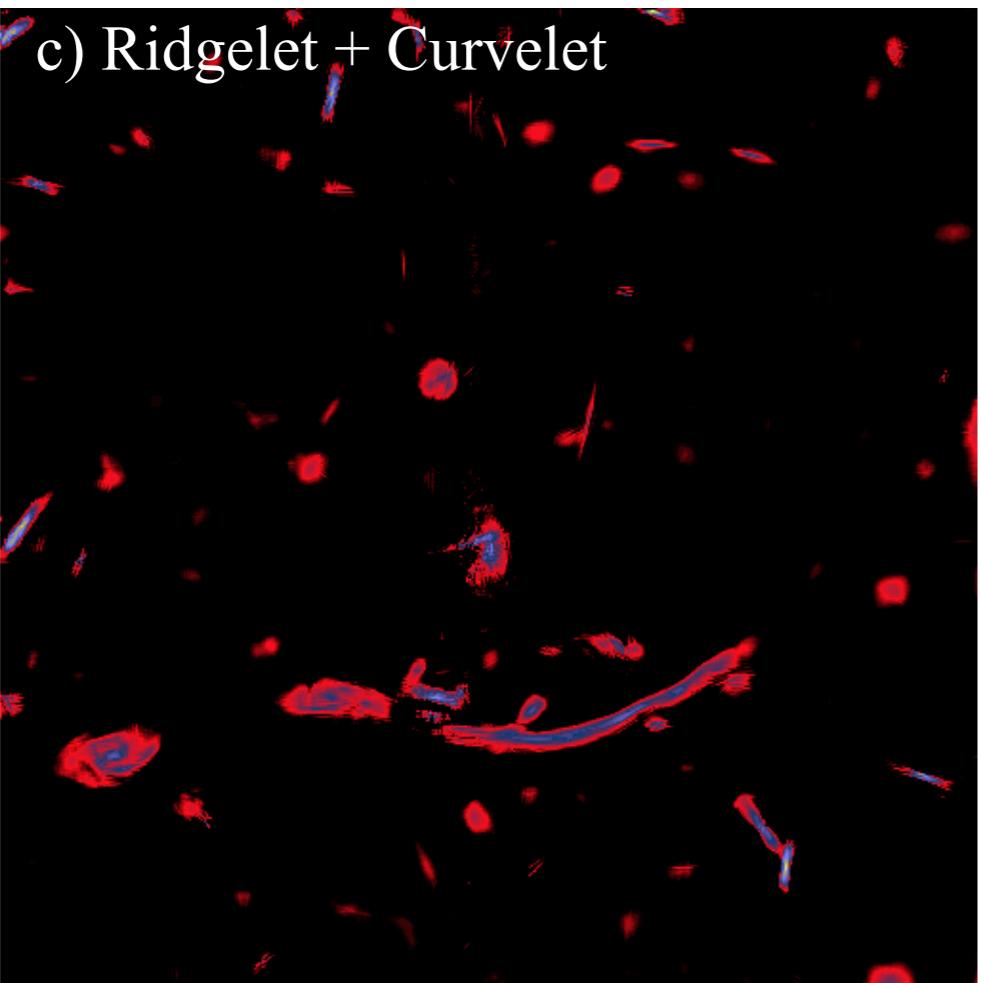
a) A370



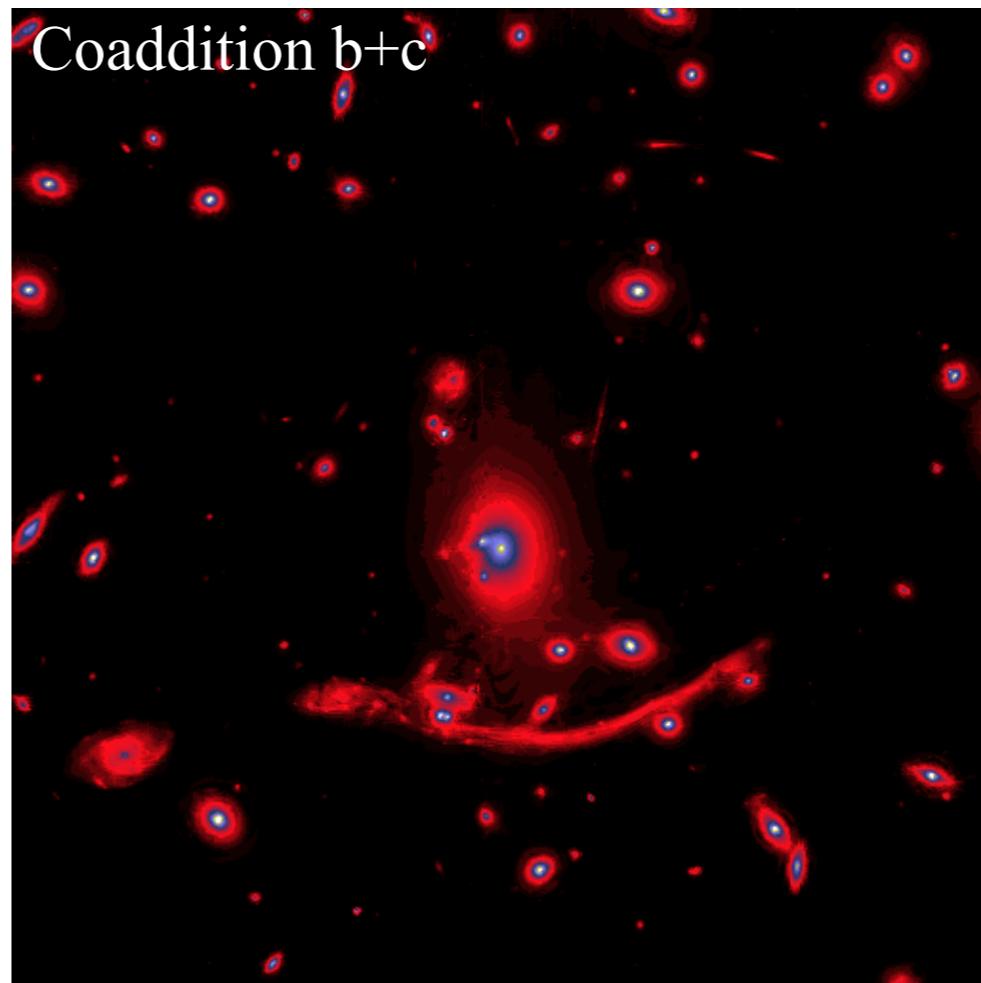
b) a trous



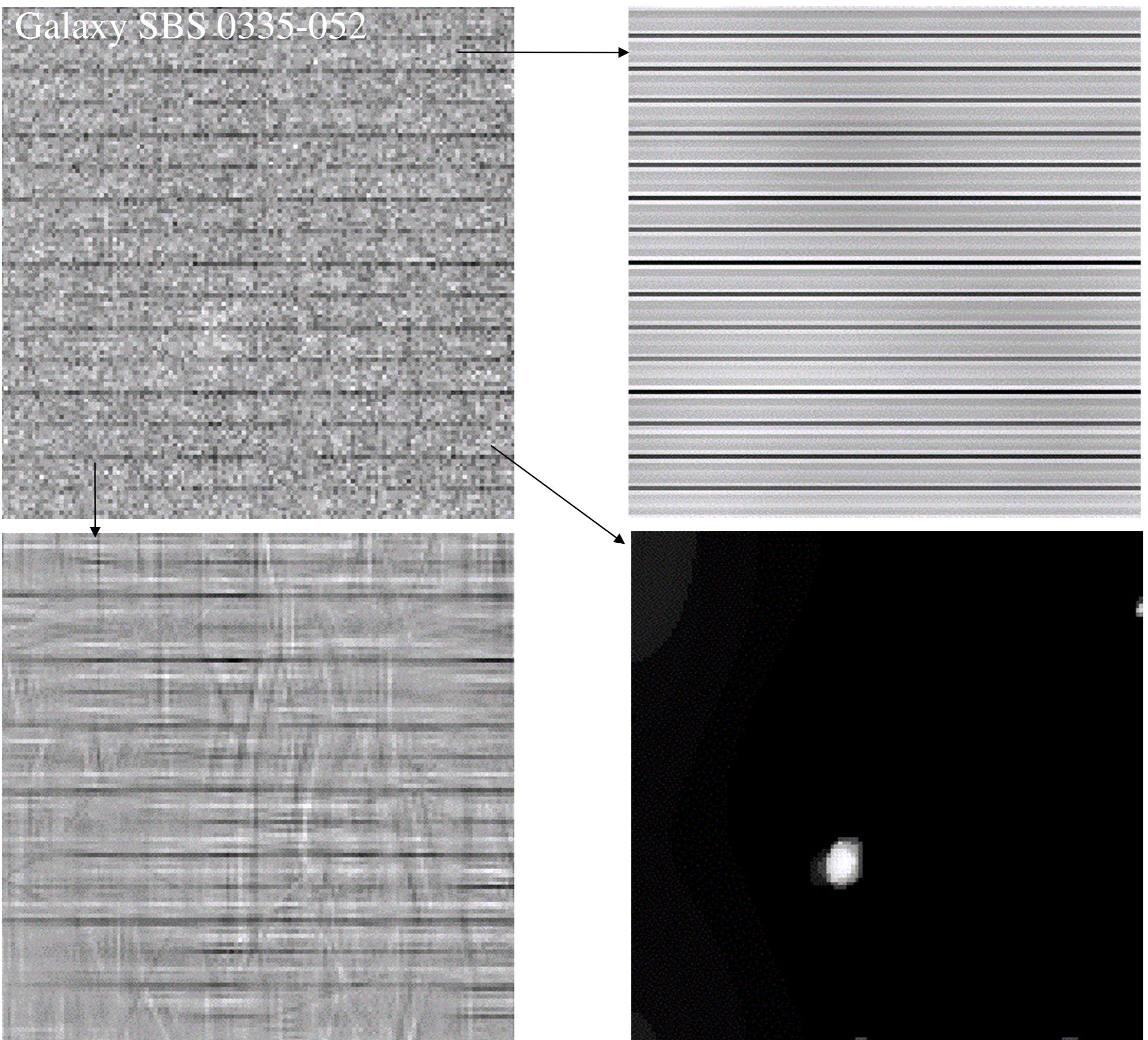
c) Ridgelet + Curvelet



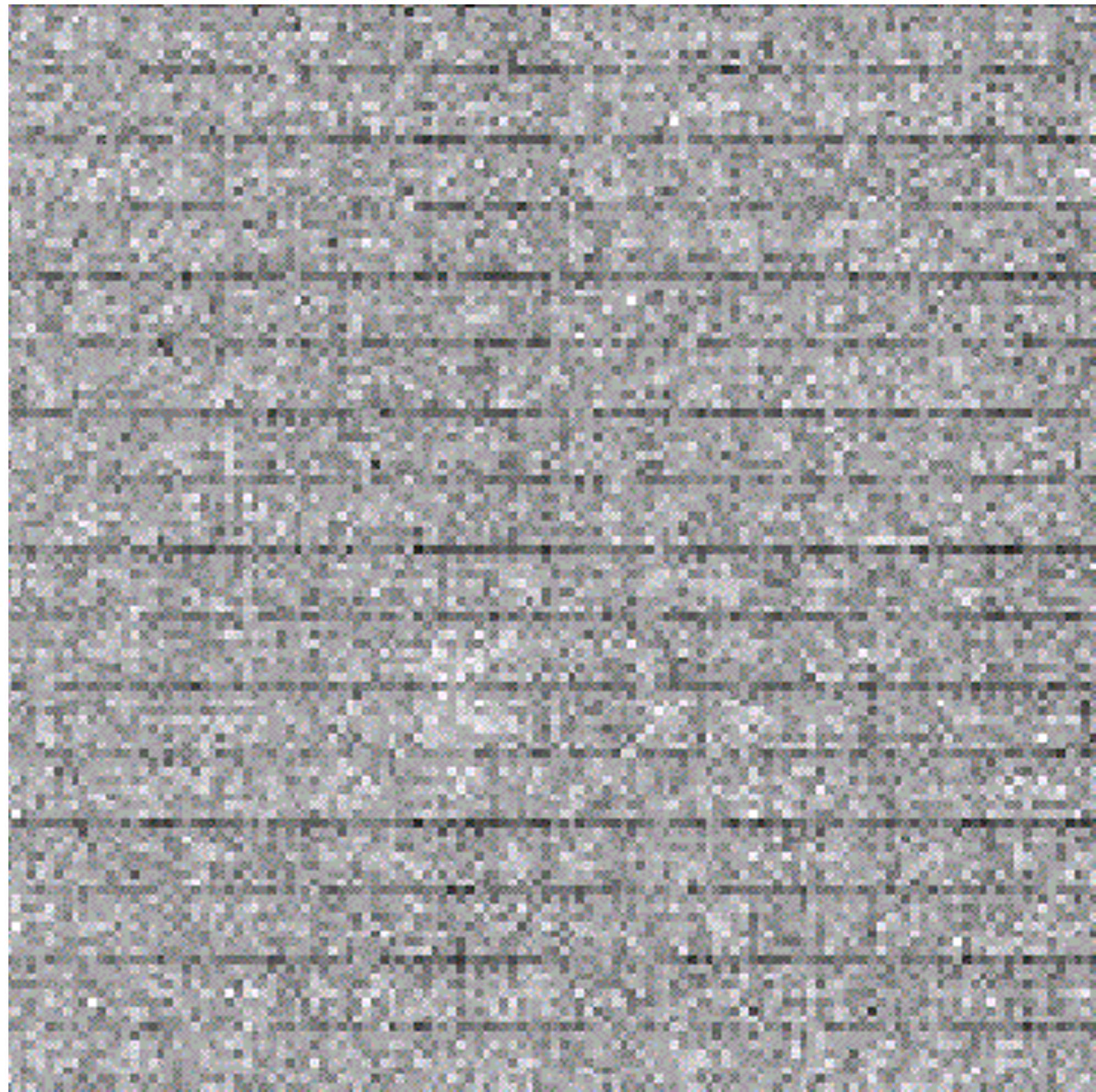
Coaddition b+c



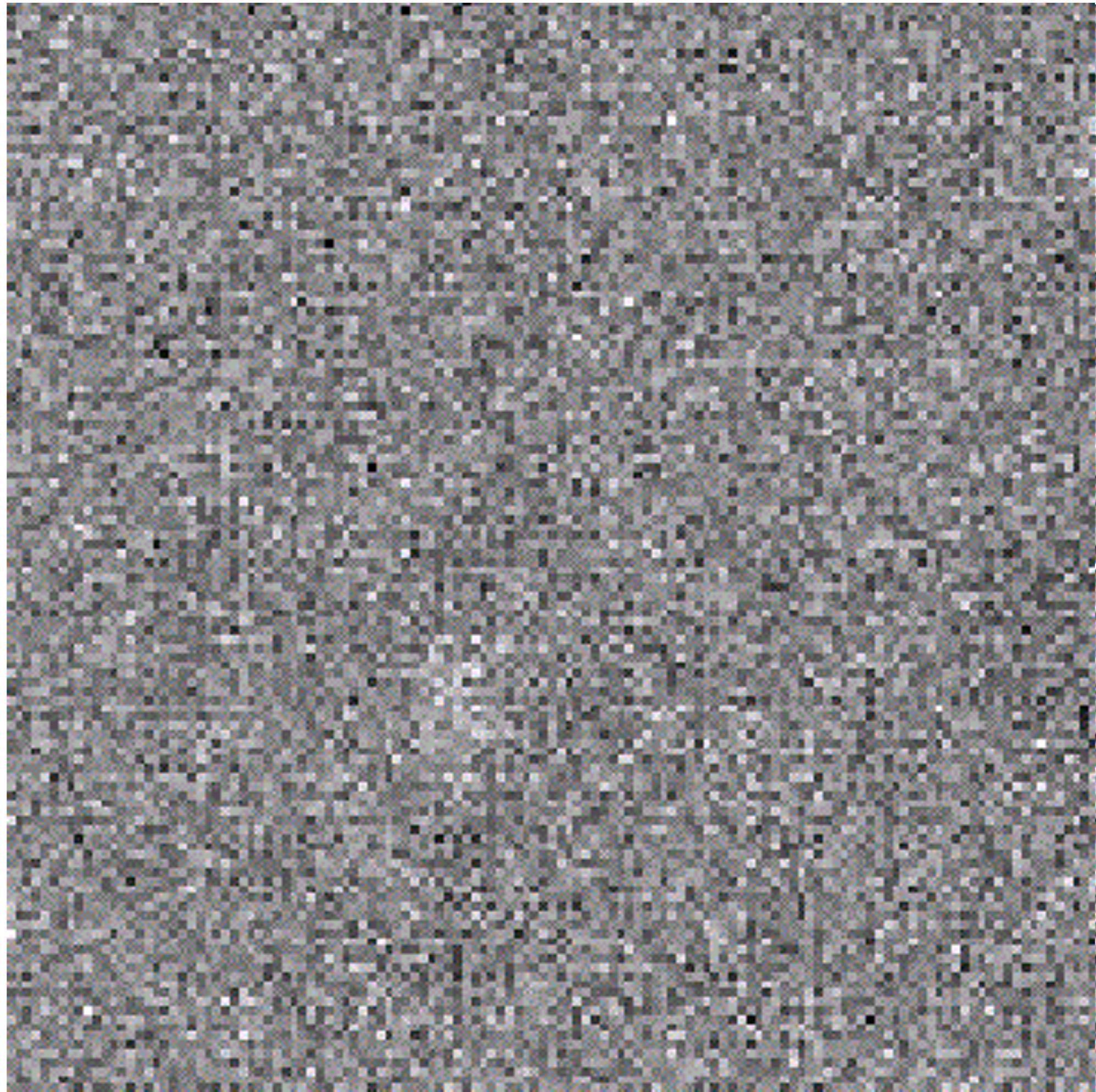
Galaxy SBS 0335-052



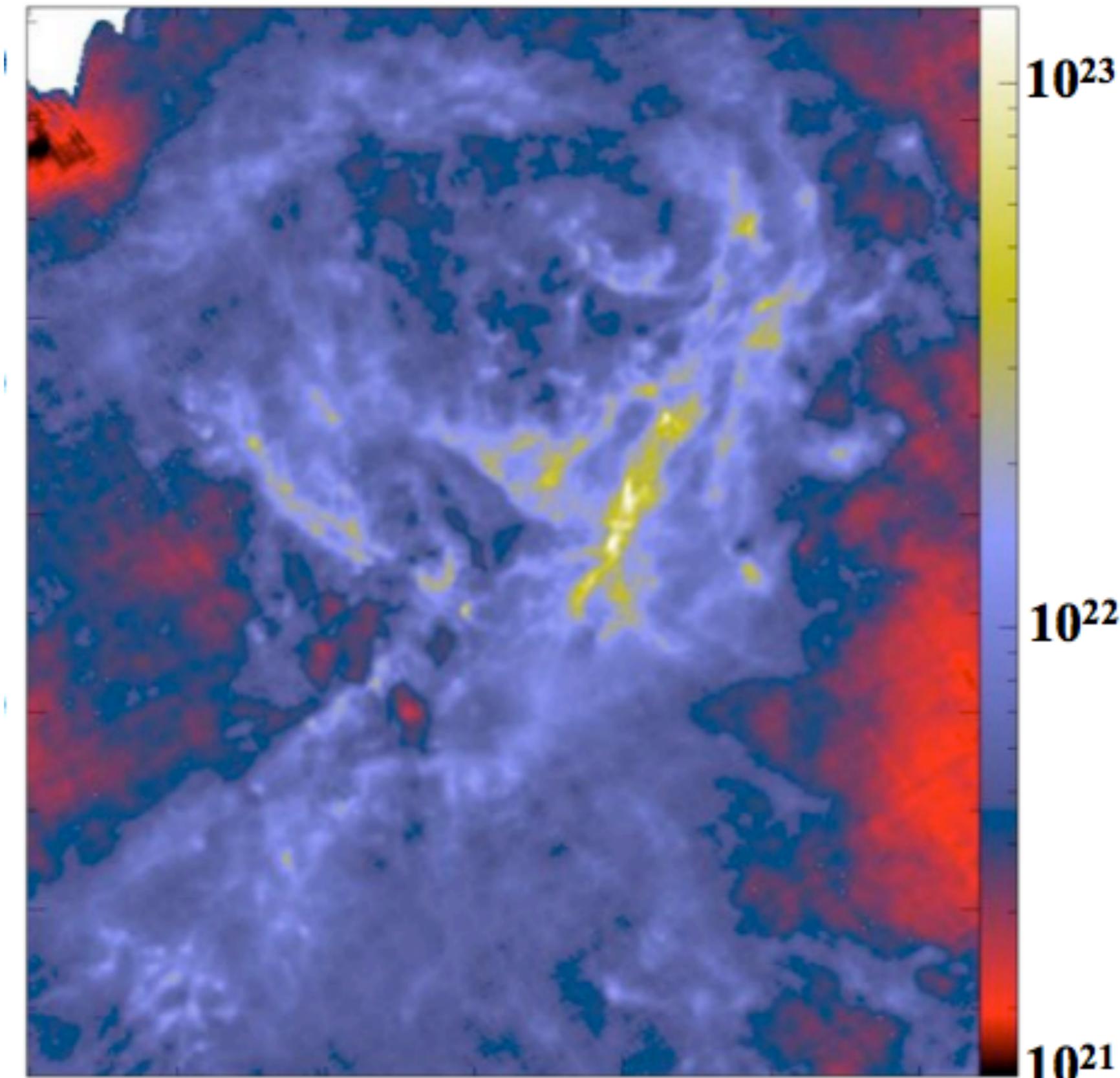
Galaxy SBS 0335-052
10 micron
GEMINI-OSCIR



Galaxy SBS 0335-052
10 micron
GEMINI-OSCIR



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



one of the nearest infrared dark clouds (Aquila Main: $d \sim 260$ pc)

Dense cores form primarily in filaments

Morphological Component Analysis:

Herschel Column density map

Cores

Wavelet component (H_2/cm^2)

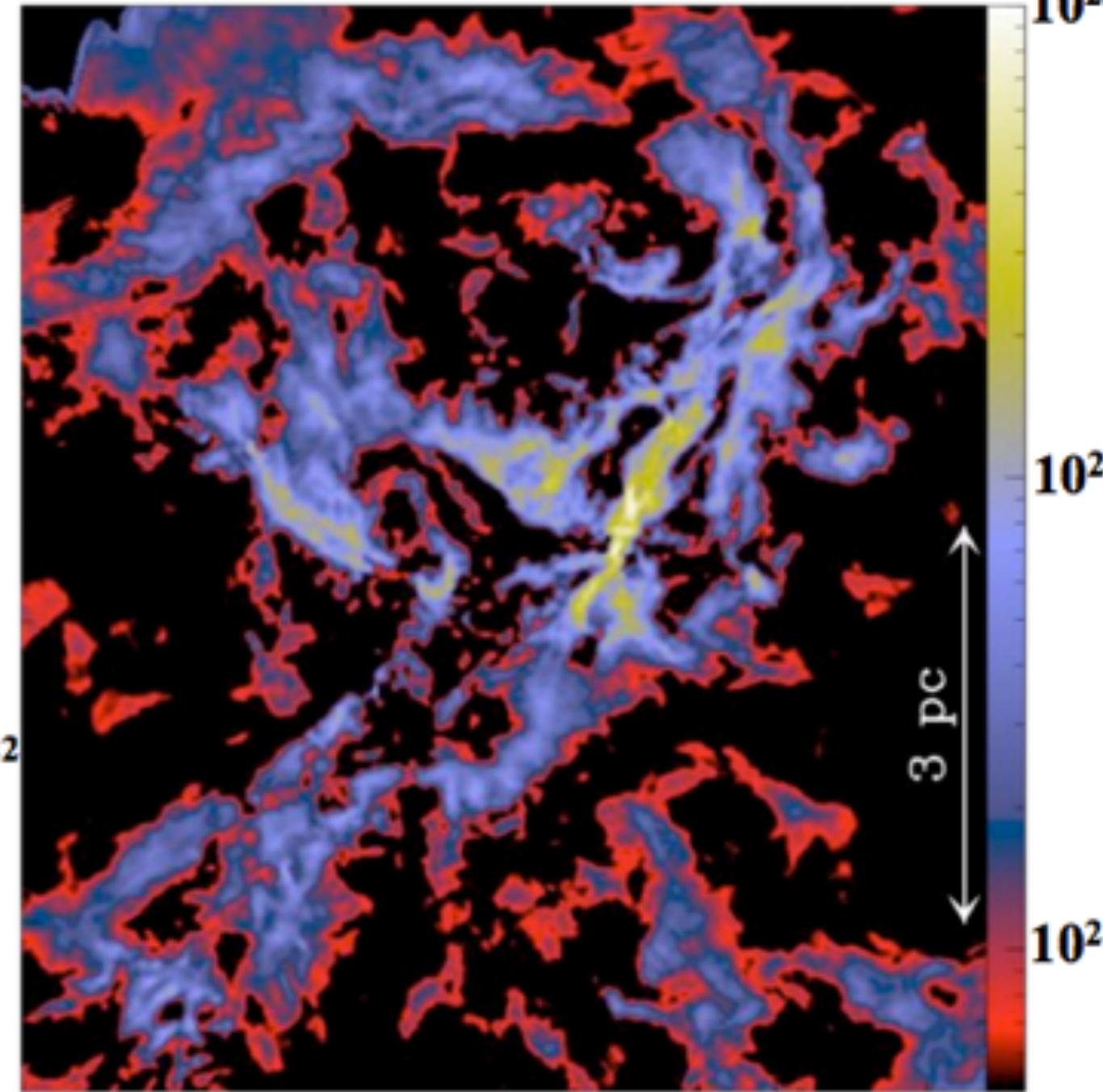
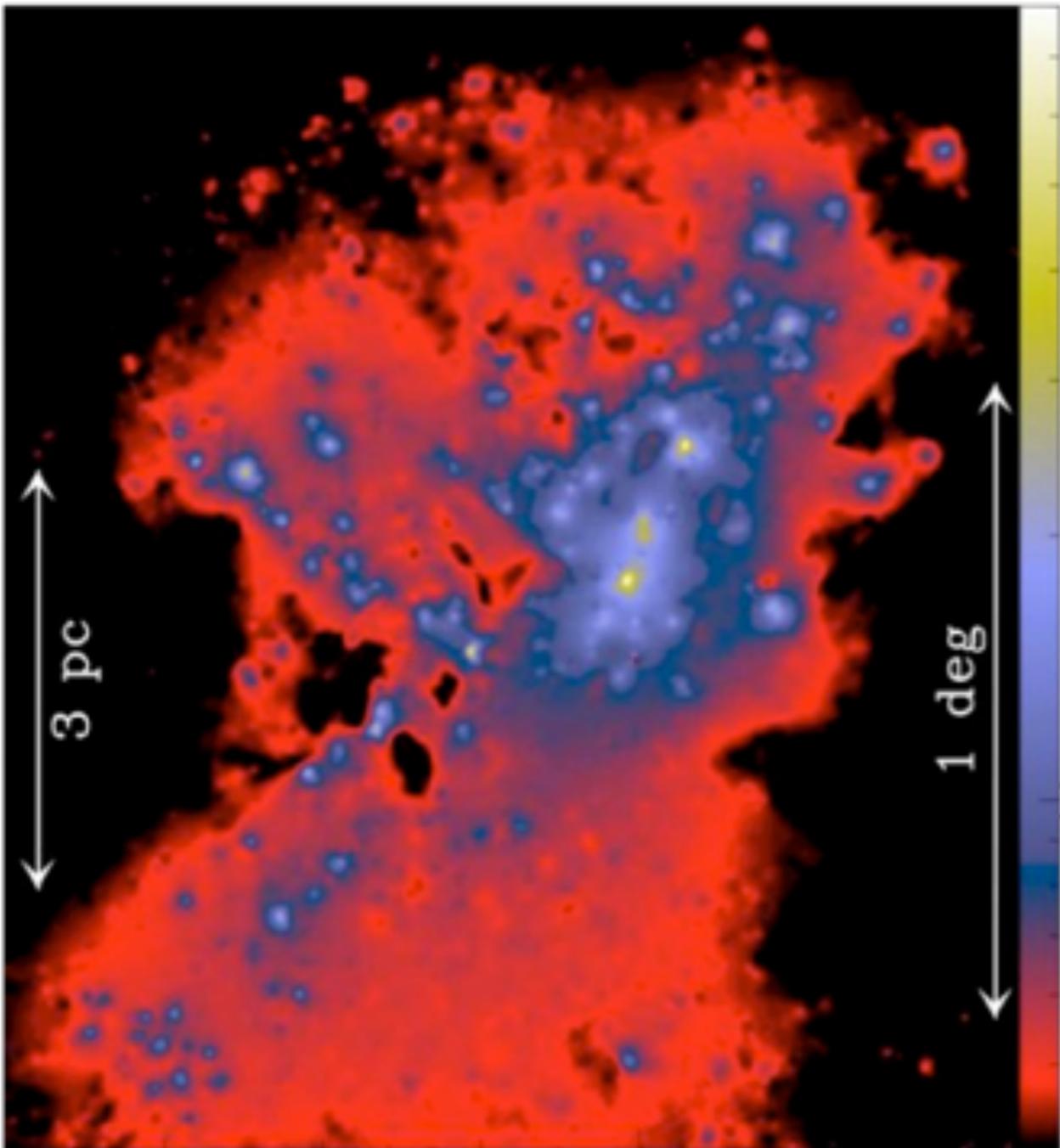
=

Filaments

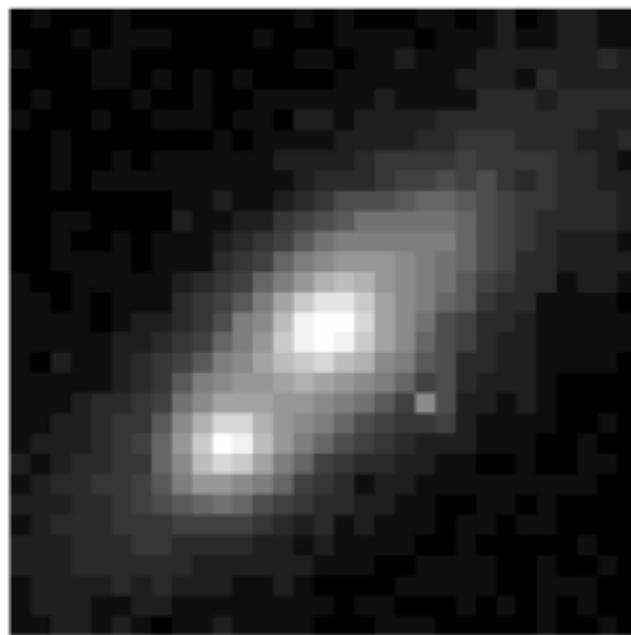
+

Curvelet component (H_2/cm^2)

(P. Didelon based on
Starck et al. 2003)



MCA based artifact removal for SNe detection



SNa and its host galaxy

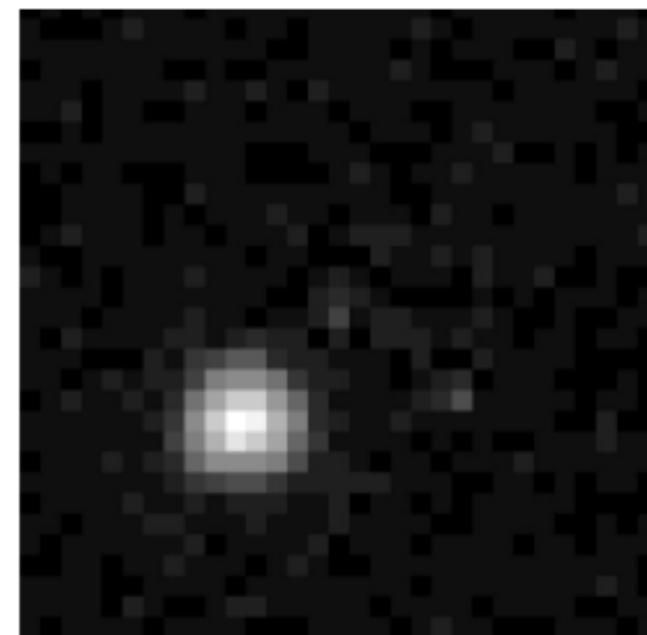
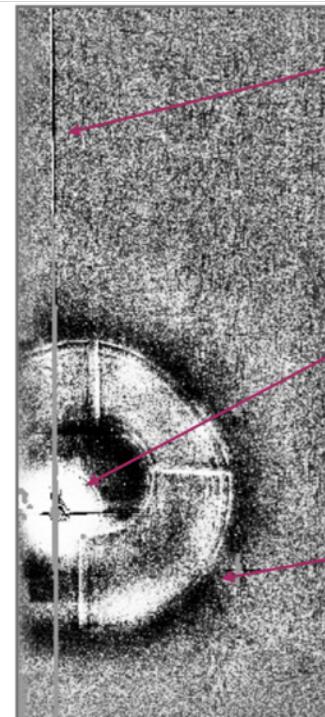
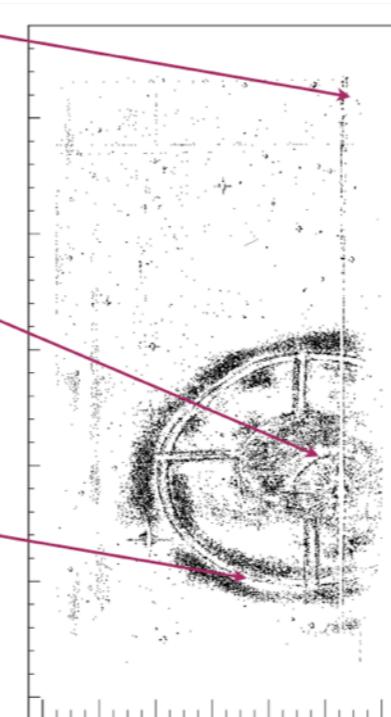


Image after subtraction of
the reference sky



Subtracted image



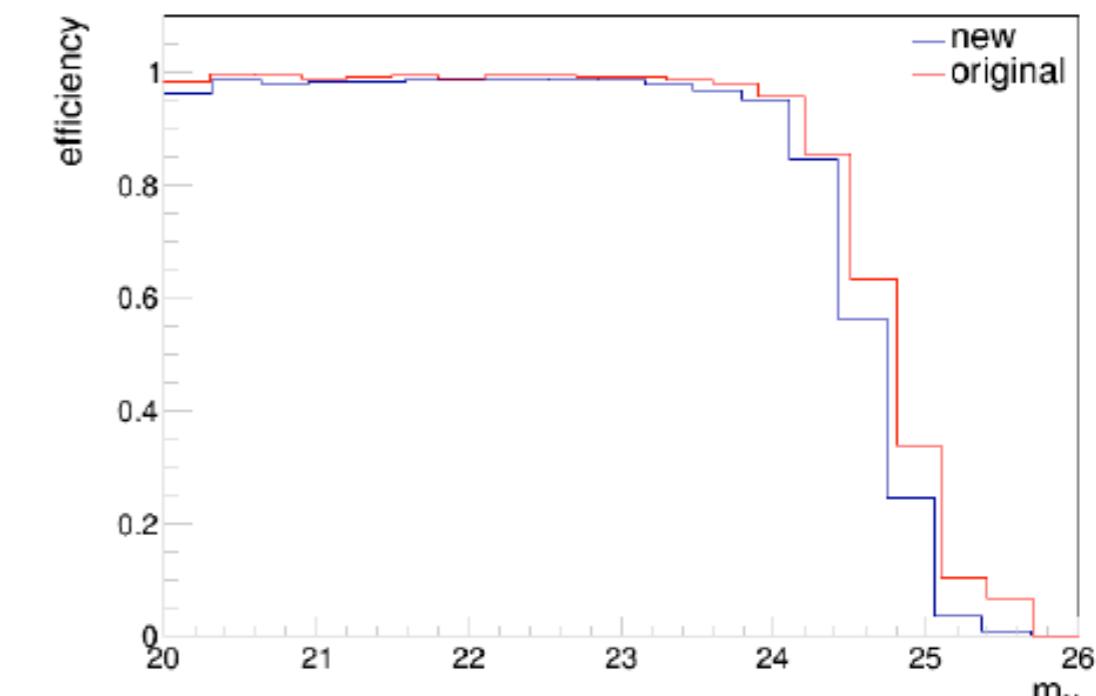
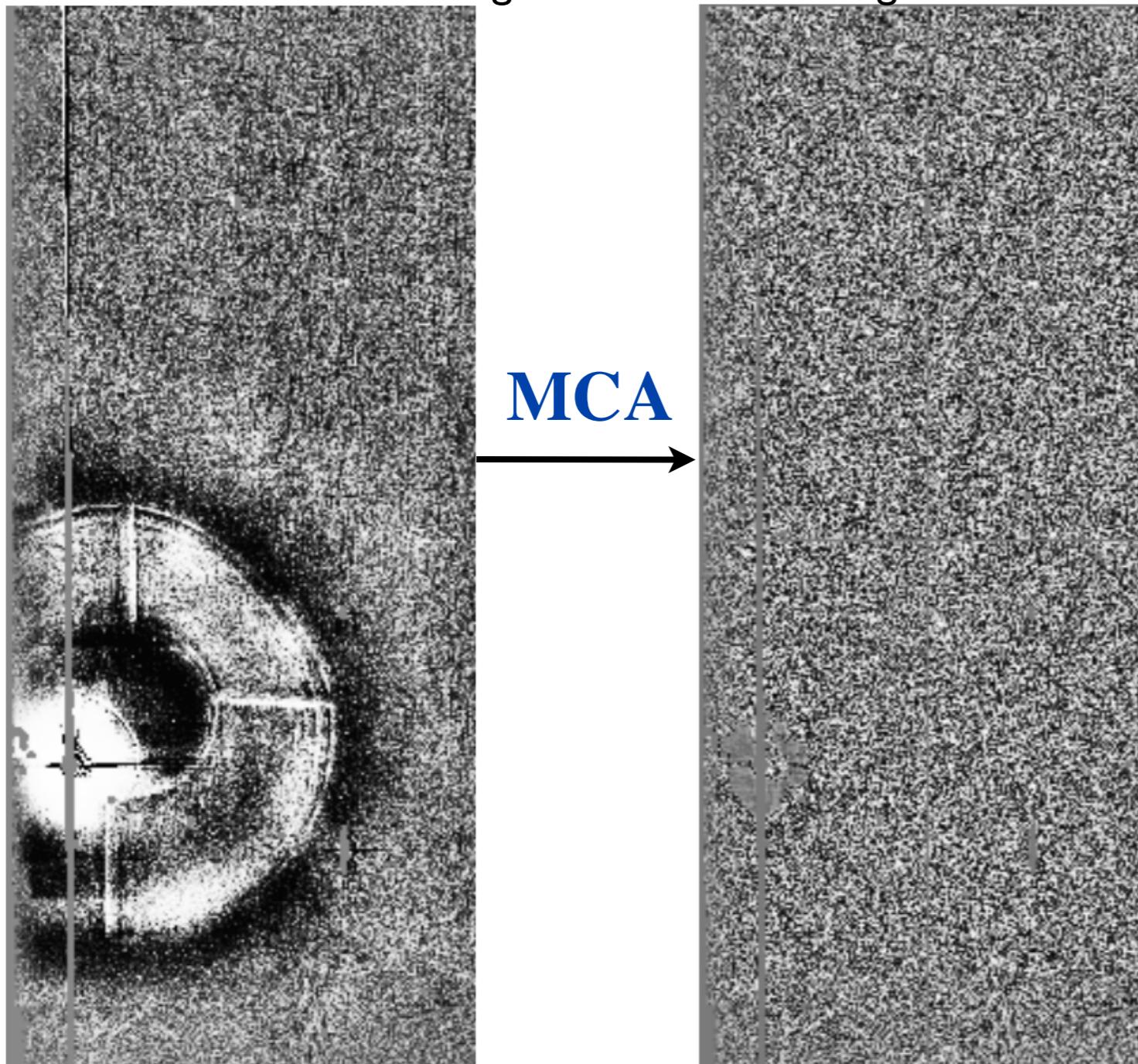
Detection catalogue

- SNe are detected by subtraction of a reference image.

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

MCA and SNe Detection

MCA cleaning of a subtracted image



Similar detection efficiency but greatly reduced number of spurious detections

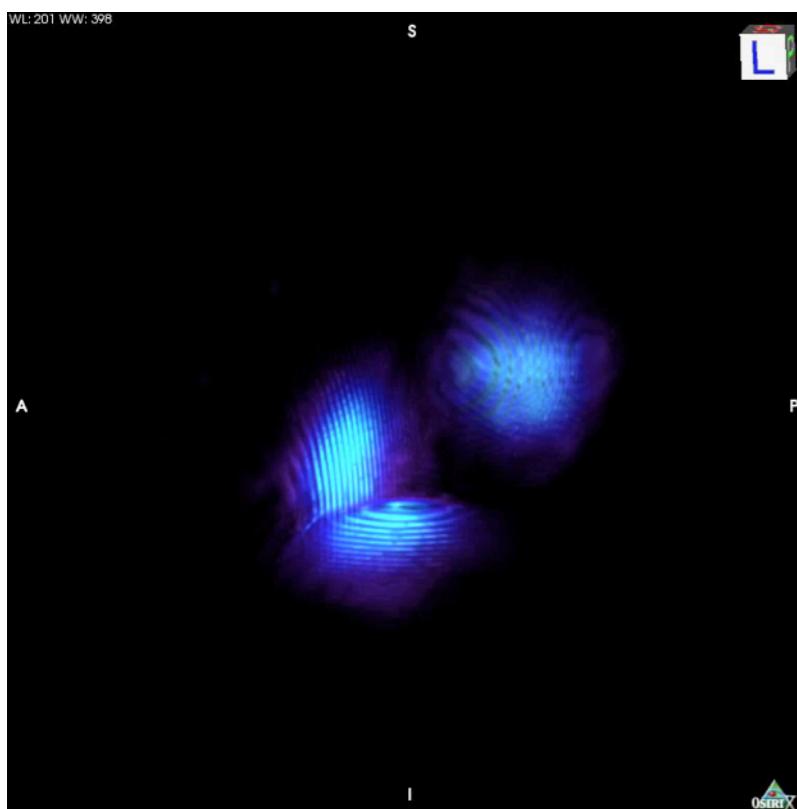
- Möller, et al, 2015, *SNIa detection in the SNLS photometric analysis using Morphological Component Analysis*, 04, Id 041, JCAP, [arxiv:1501.02110](https://arxiv.org/abs/1501.02110).

3D Morphological Component Analysis

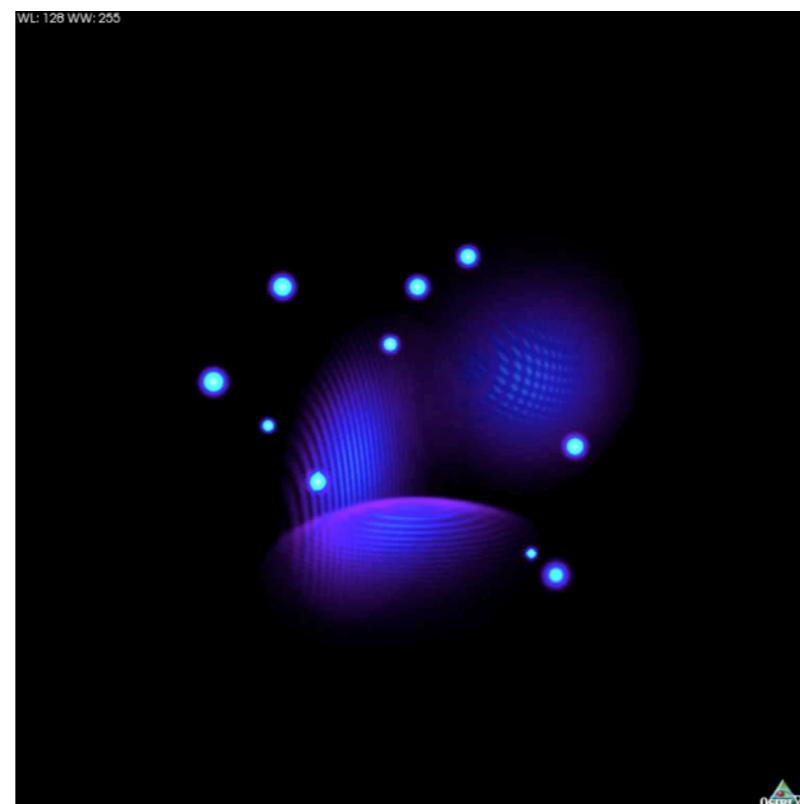


A. Woiselle

Shells

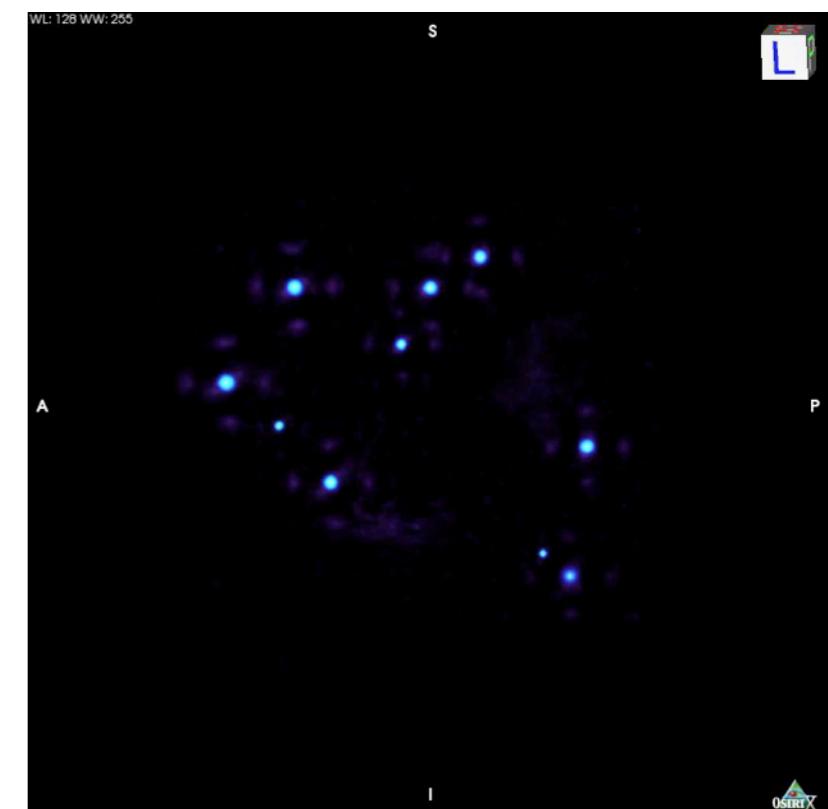


Original (3D shells + Gaussians)



Dictionary
RidCurvelets + 3D UDWT.

Gaussians



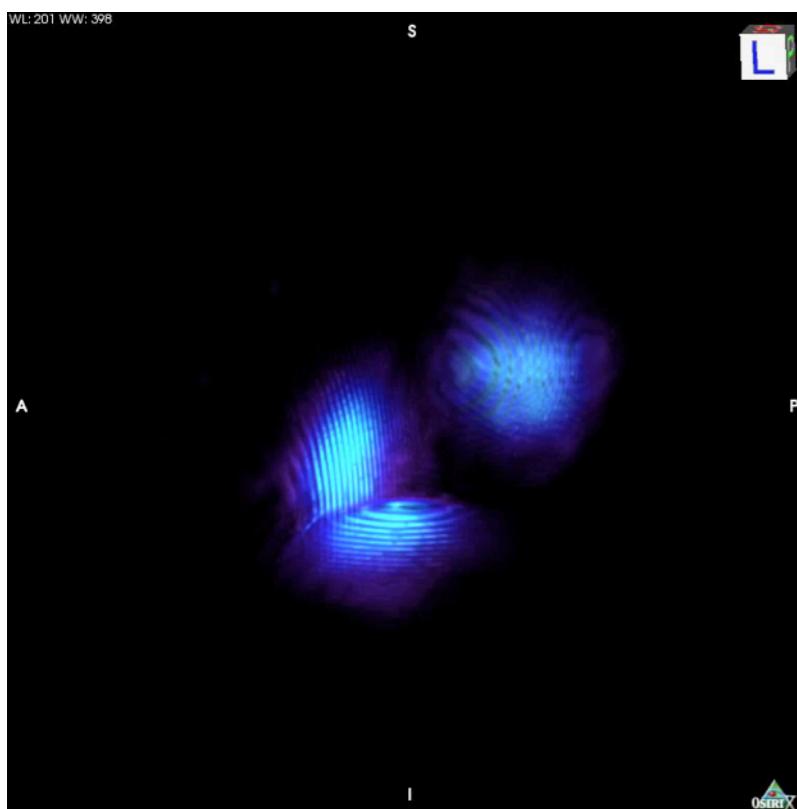
- A. Woiselle, J.L. Starck, M.J. Fadili, "[3D Data Denoising and Inpainting with the Fast Curvelet transform](#)", *JMIV*, 39, 2, pp 121-139, 2011.
- A. Woiselle, J.L. Starck, M.J. Fadili, "[3D curvelet transforms and astronomical data restoration](#)", *Applied and Computational Harmonic Analysis*, Vol. 28, No. 2, pp. 171-188, 2010.

3D Morphological Component Analysis

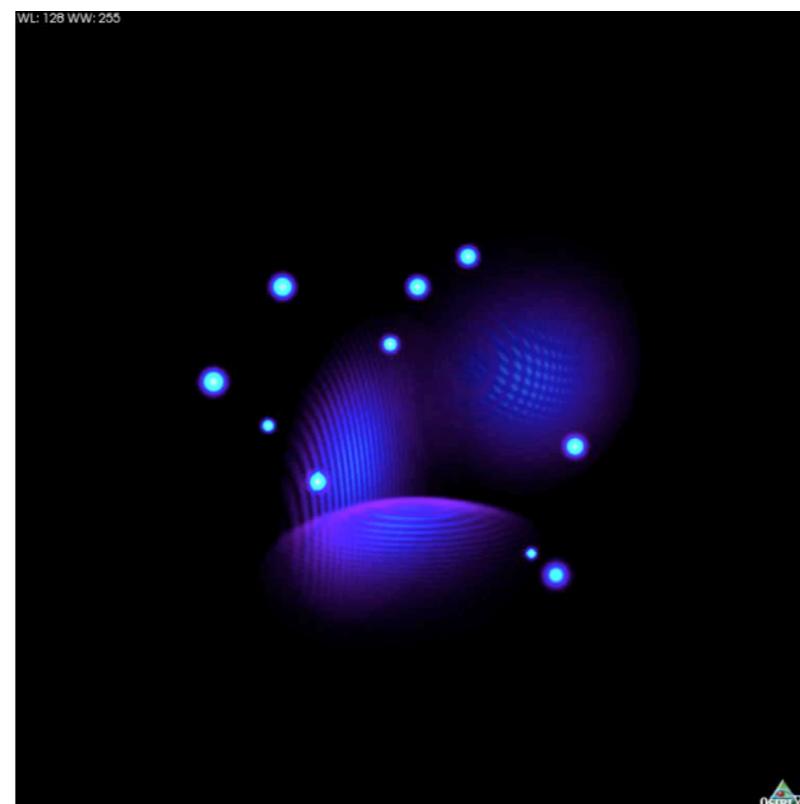


A. Woiselle

Shells

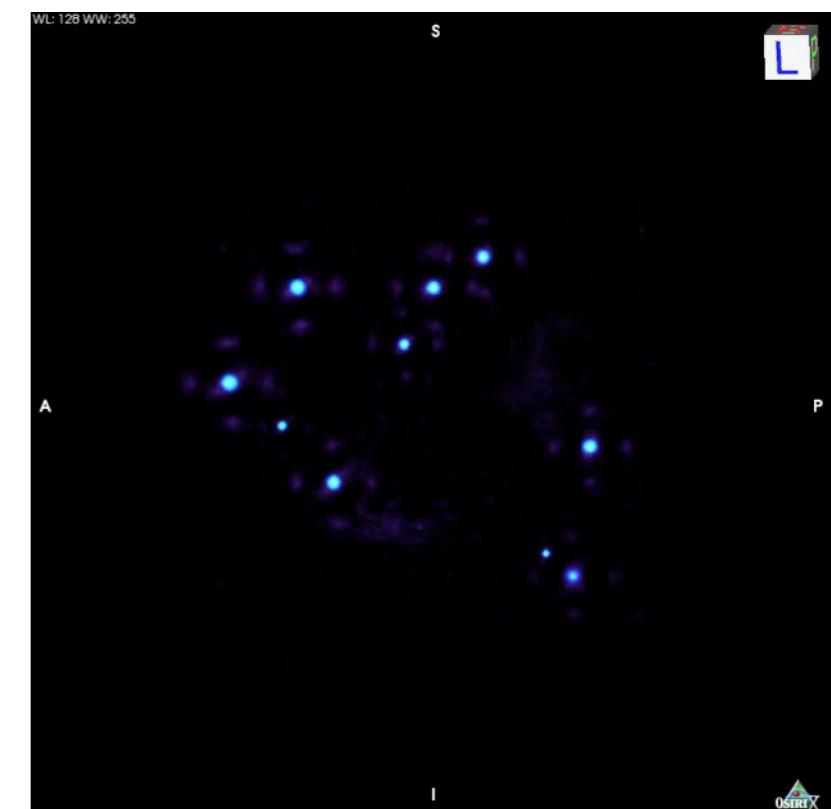


Original (3D shells + Gaussians)



Dictionary
RidCurvelets + 3D UDWT.

Gaussians



- A. Woiselle, J.L. Starck, M.J. Fadili, "[3D Data Denoising and Inpainting with the Fast Curvelet transform](#)", *JMIV*, 39, 2, pp 121-139, 2011.
- A. Woiselle, J.L. Starck, M.J. Fadili, "[3D curvelet transforms and astronomical data restoration](#)", *Applied and Computational Harmonic Analysis*, Vol. 28, No. 2, pp. 171-188, 2010.



Multichannel Unmixing

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

$H_i = \text{Id}$ The mixing matrix A is assumed to be known
$$Y = AX + N$$

$H_i = \text{Id}$ The mixing matrix A is assumed unknown
$$Y = AX + N$$

$H_i \neq \text{Id}$ The mixing matrix A is assumed unknown
$$Y = HAX + N$$

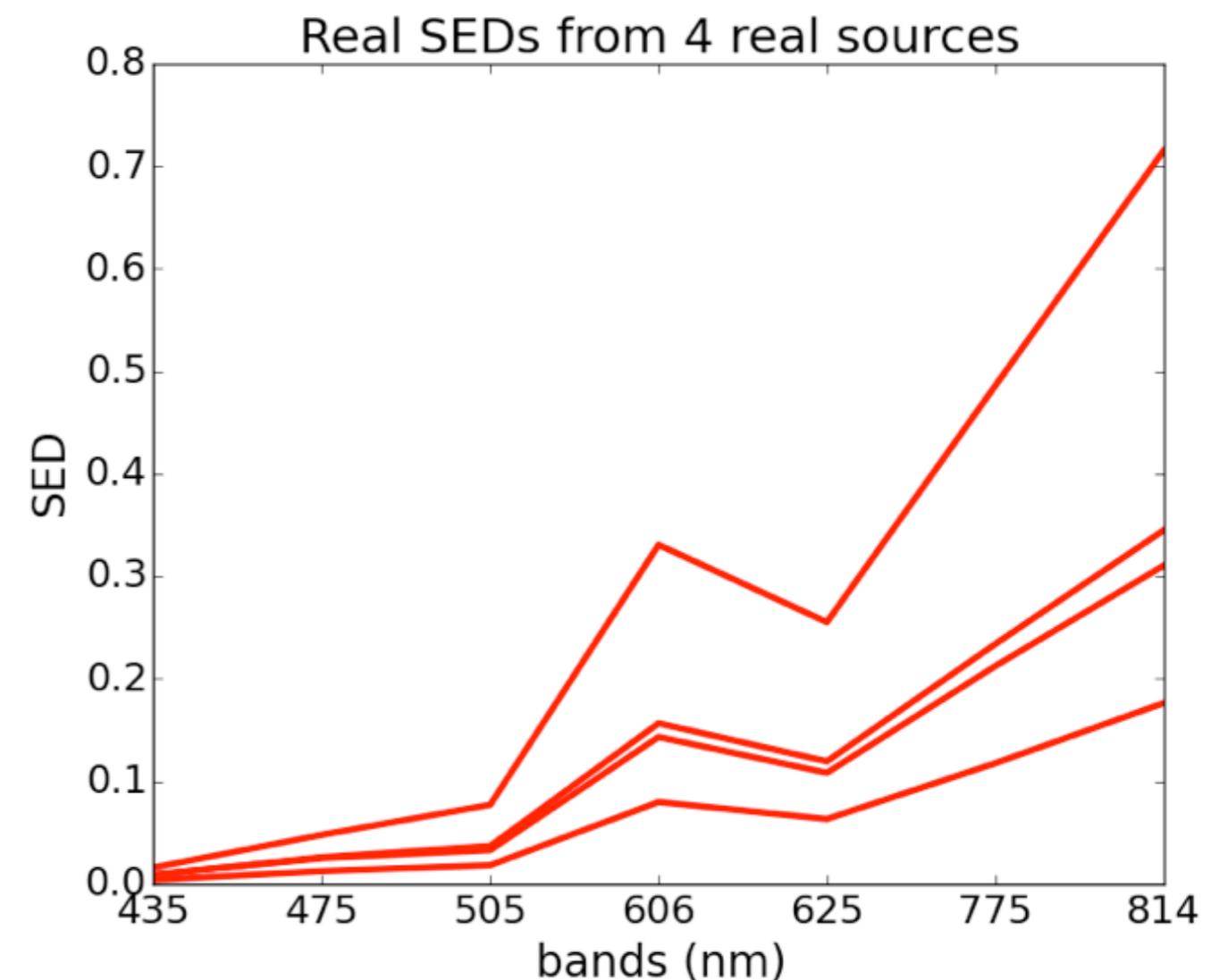
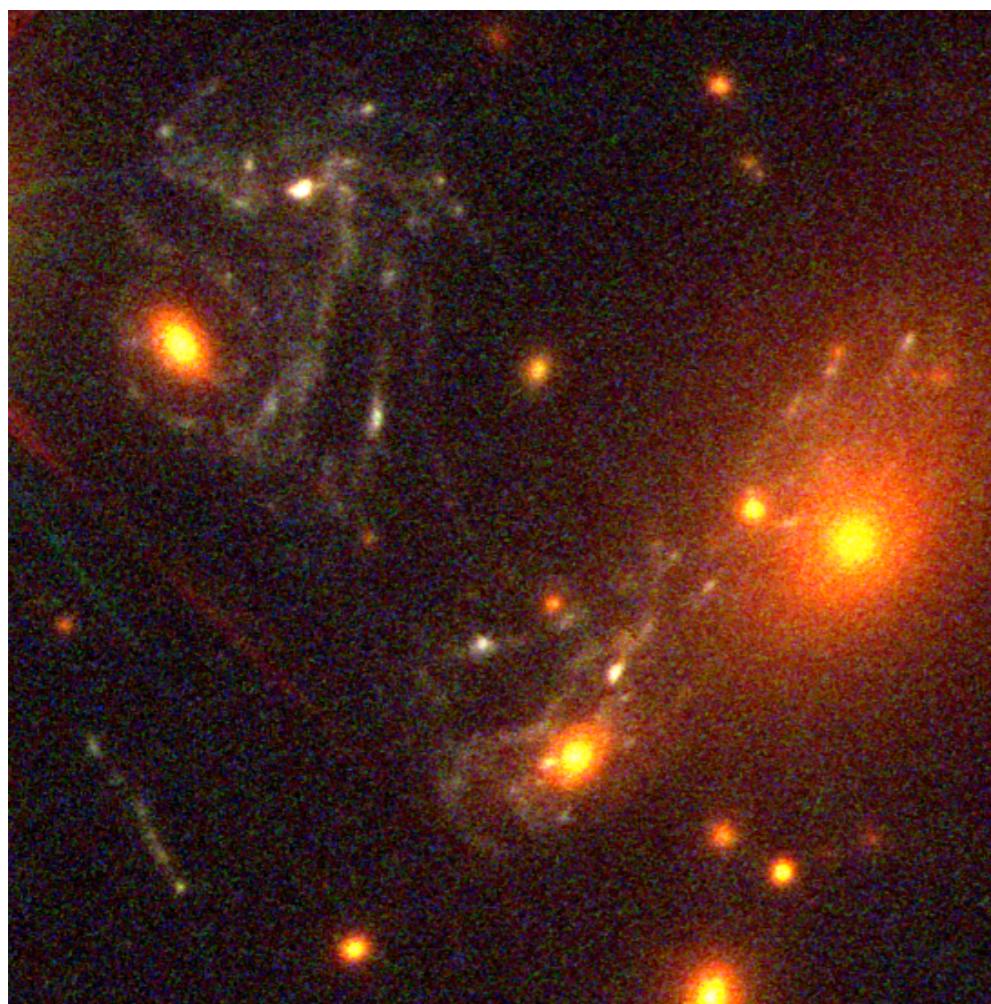
Multichannel data

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

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

$$H_i = \text{Id}$$

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.

$$\min_X \| Y - AX \|^2 + \sum_{j=1}^J \lambda_j \| \Psi^* x_j \|_0$$

R. Jospeh, F. Courbin and J.-L. Starck, "Multi-band morpho-Spectral Component Analysis Deblending Tool (MuSCADeT): deblending colourful objects", A&A, 589, id.A2, pp 10, 2016.

Morpho-Spectral Diversity

R. Jospeh, F. Courbin and J.-L. Starck, “Multi-band morpho-Spectral Component Analysis Deblending Tool (MuSCADeT): deblending colourful objects”, A&A, 589, id.A2, pp 10, 2016.

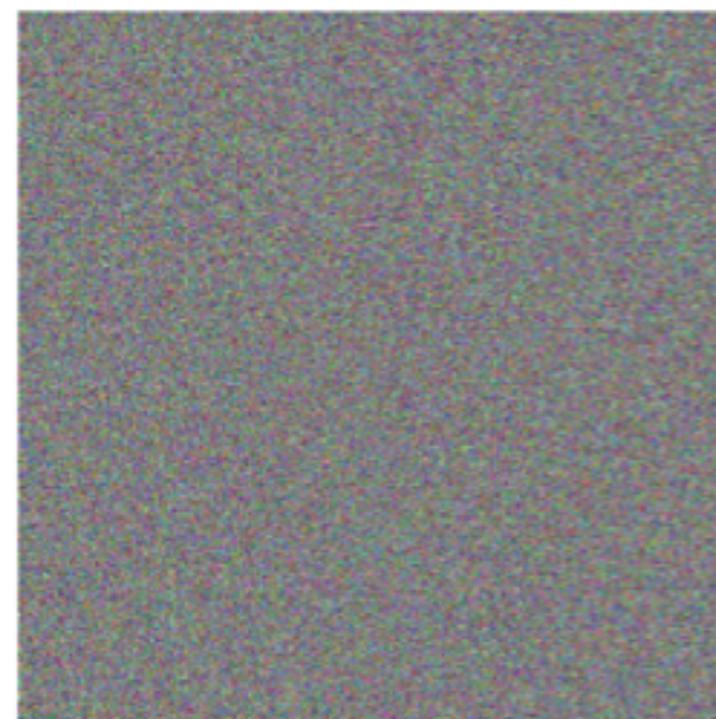
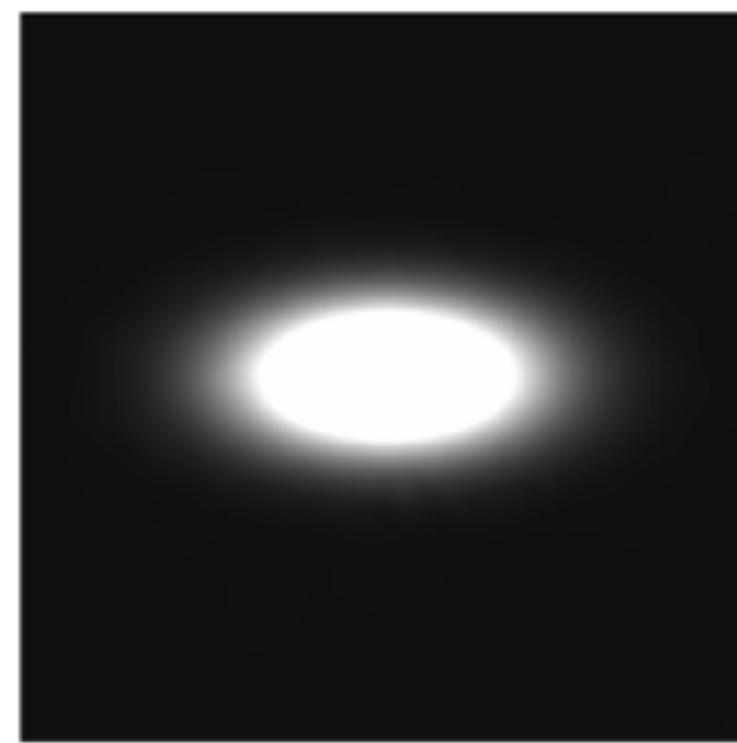
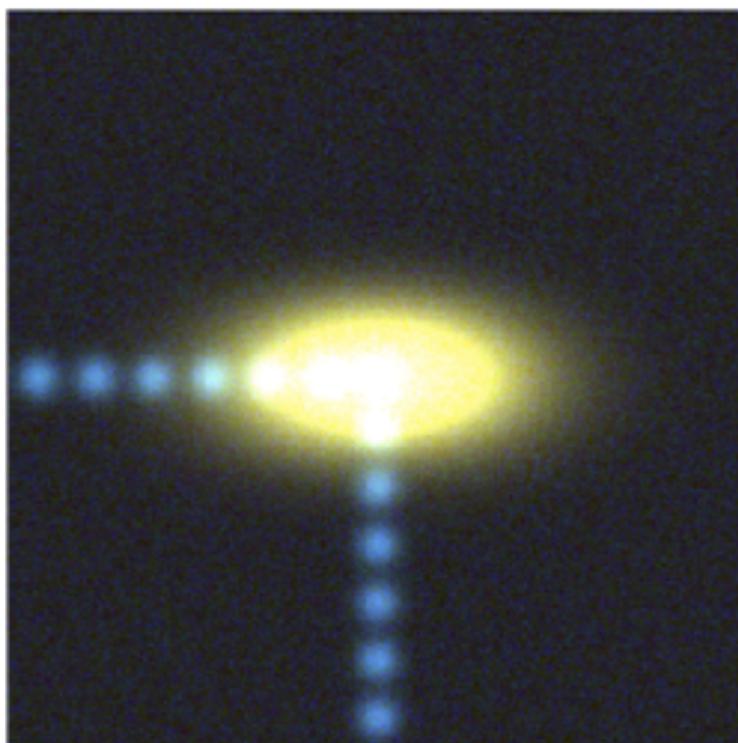
Algorithm:

Perform a gradient step: $U = X^{(n)} + \mu A^t(Y - AX^{(n)})$.

Solve for each j : $\min_{x_j^{(n+1)}} \| u_j - x_j^{(n+1)} \| + \lambda_j \| \Psi^* x_j^{(n+1)} \|_0$, and set to zero negative entries in $x_j^{(n+1)}$.

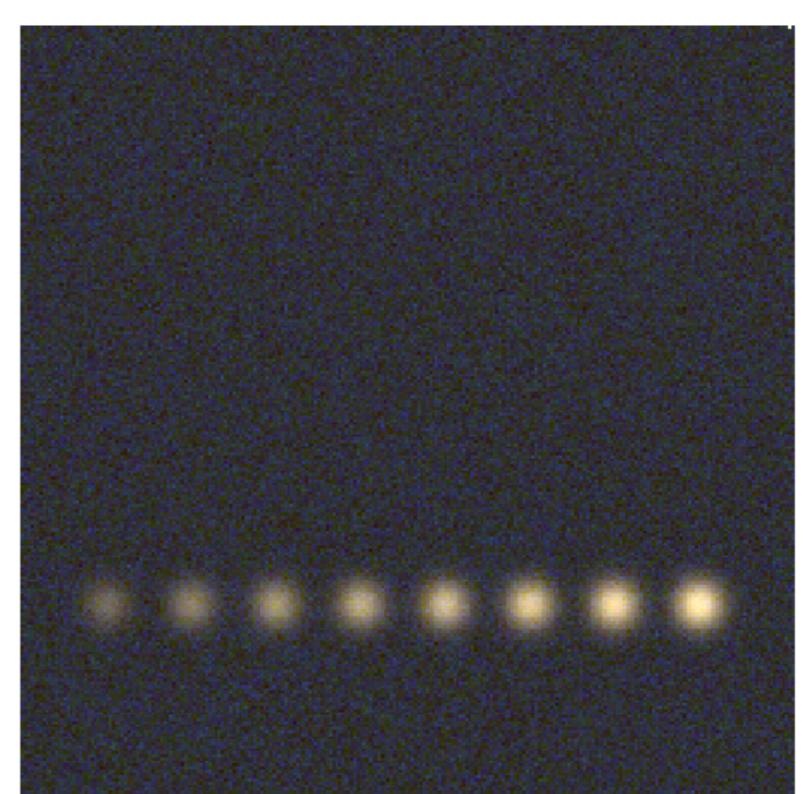
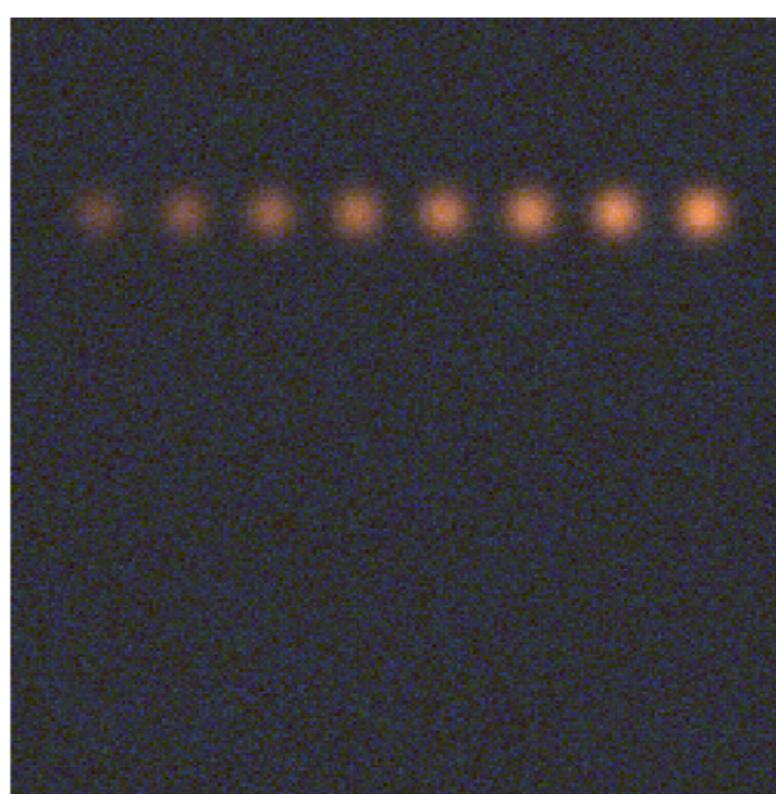
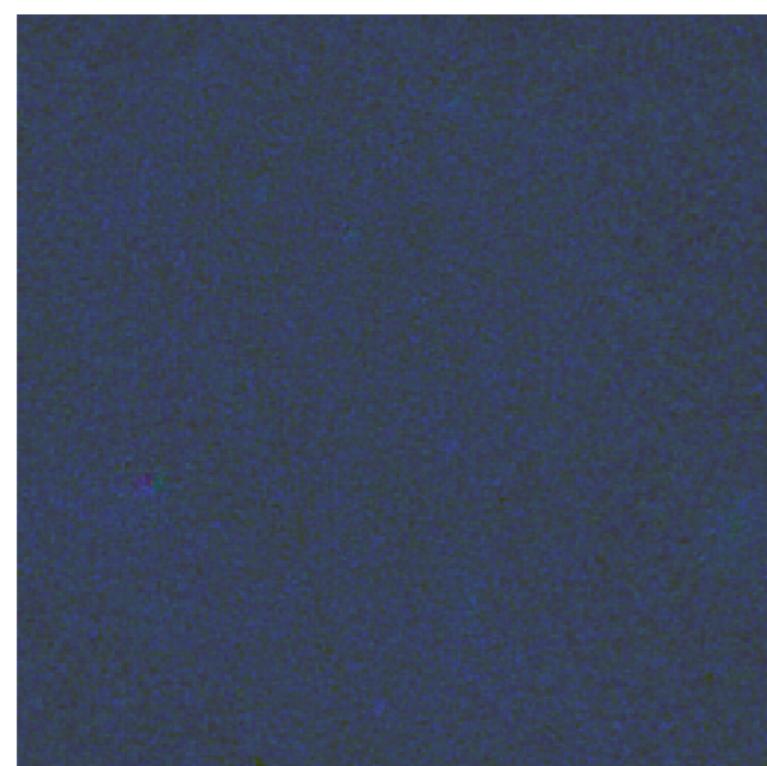
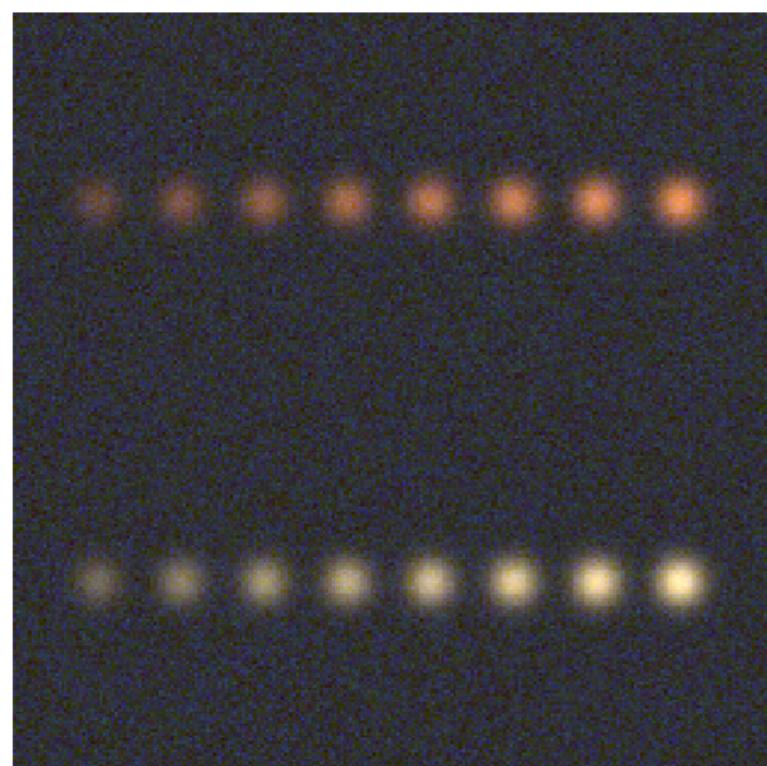
Decrease λ_j .

Multichannel data



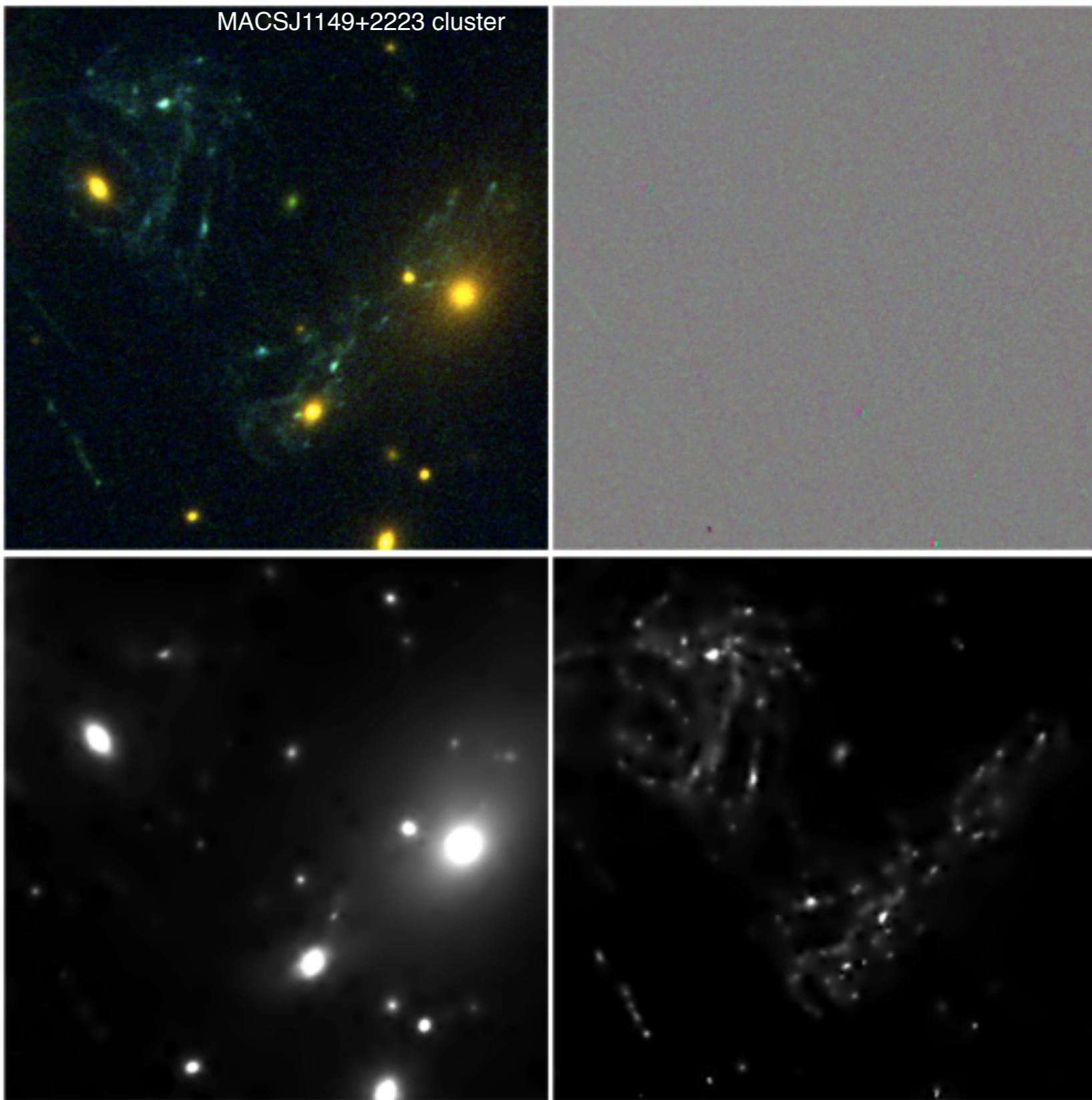
No SED variation.

Multichannel data



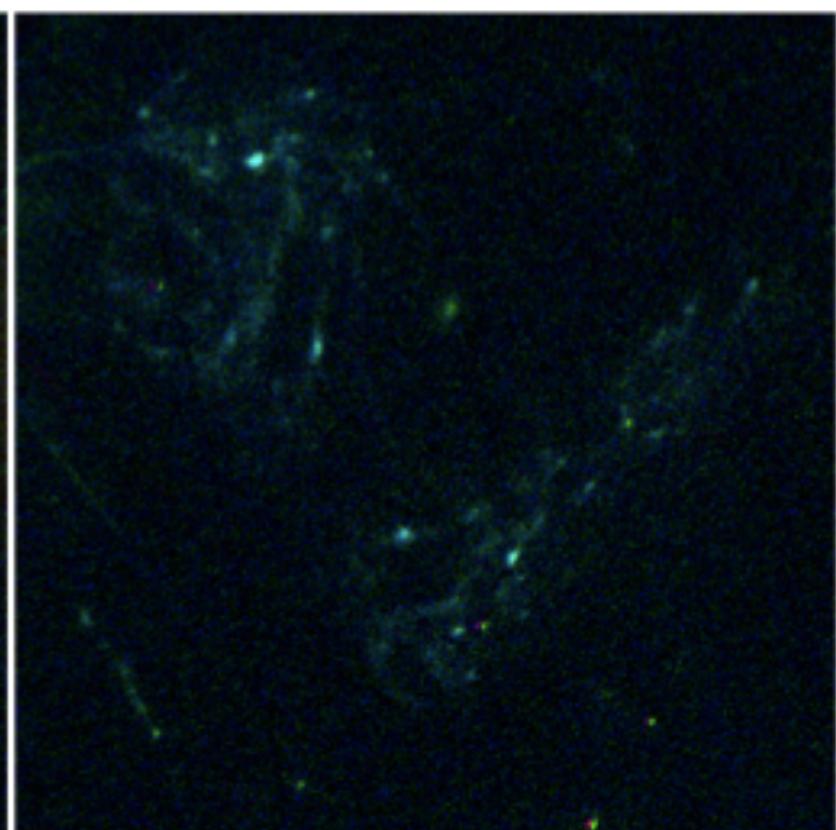
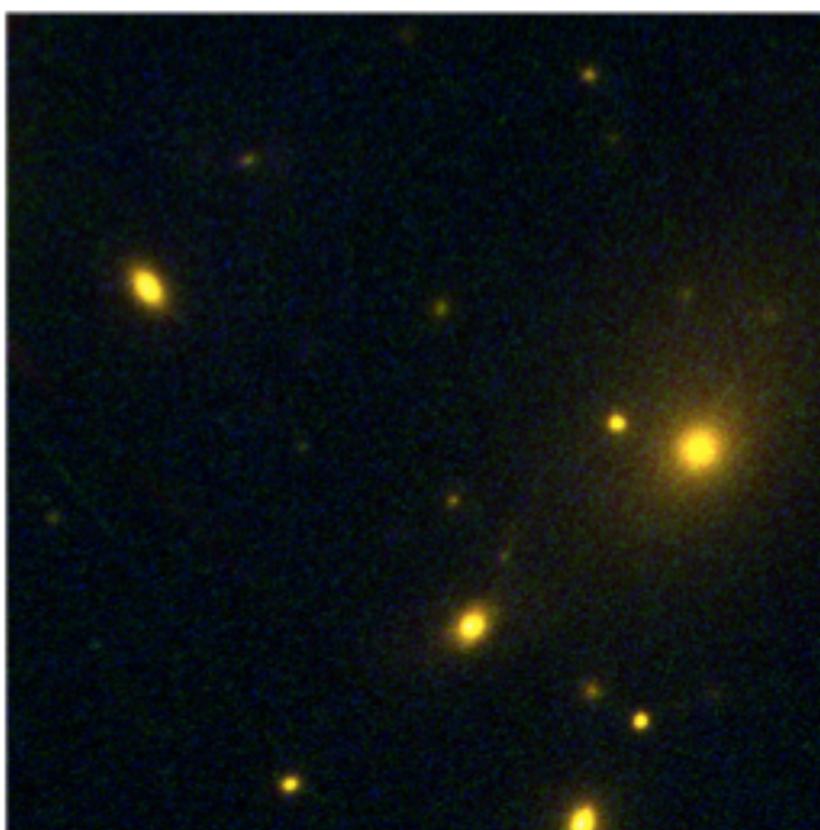
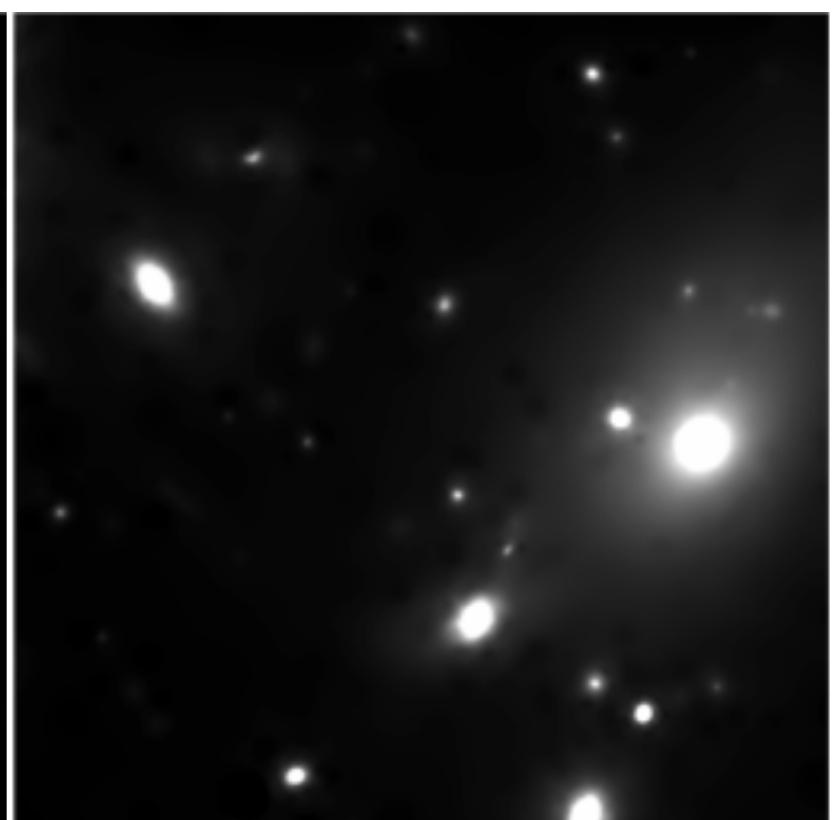
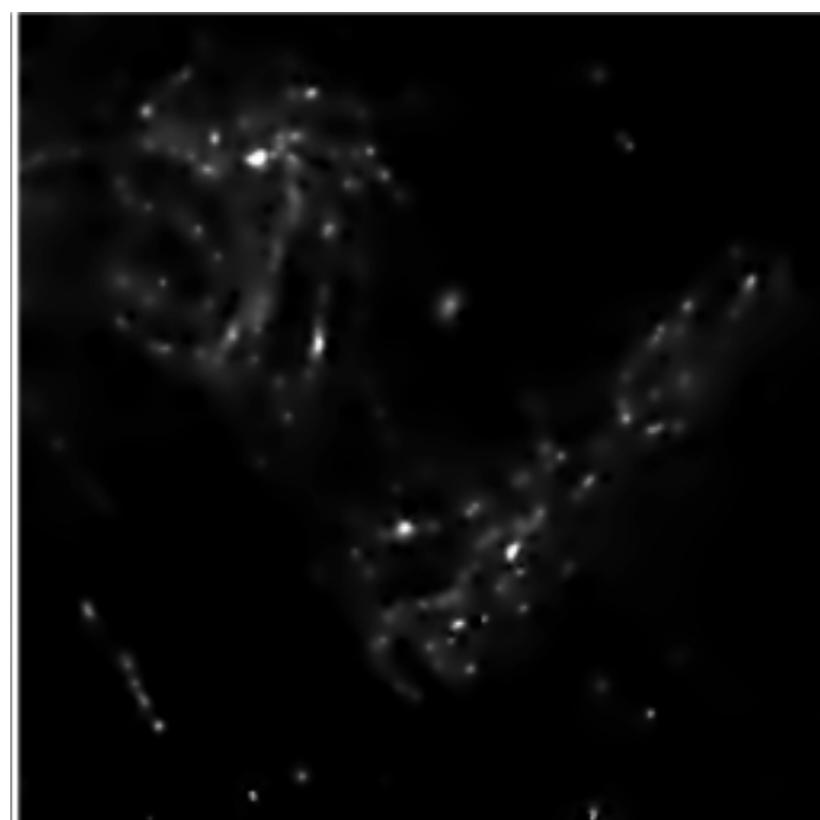
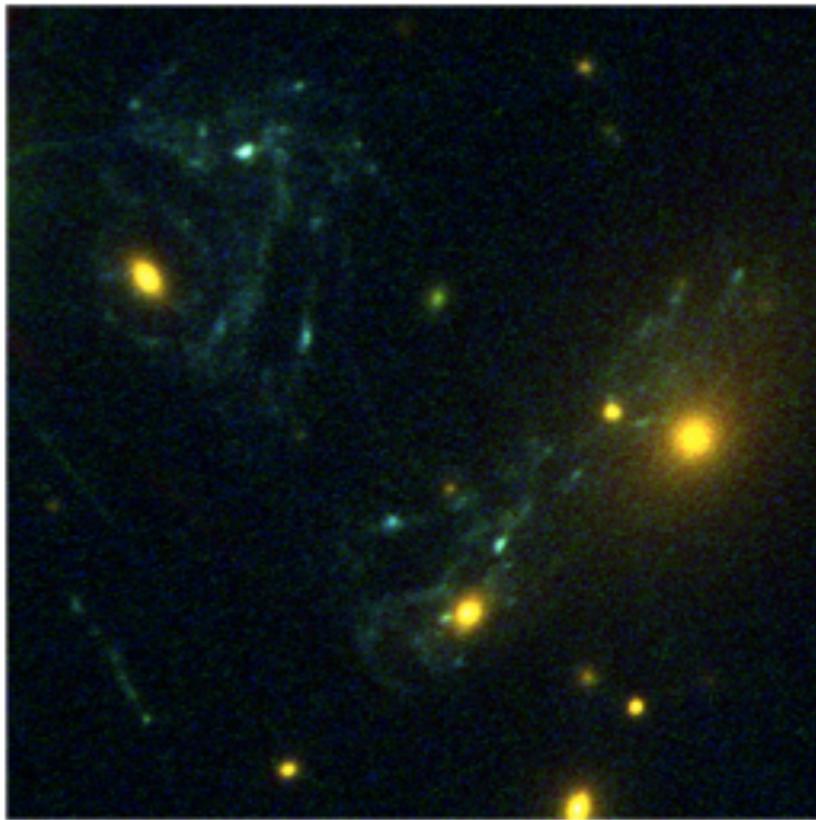
realistic SED variation.

Multichannel data



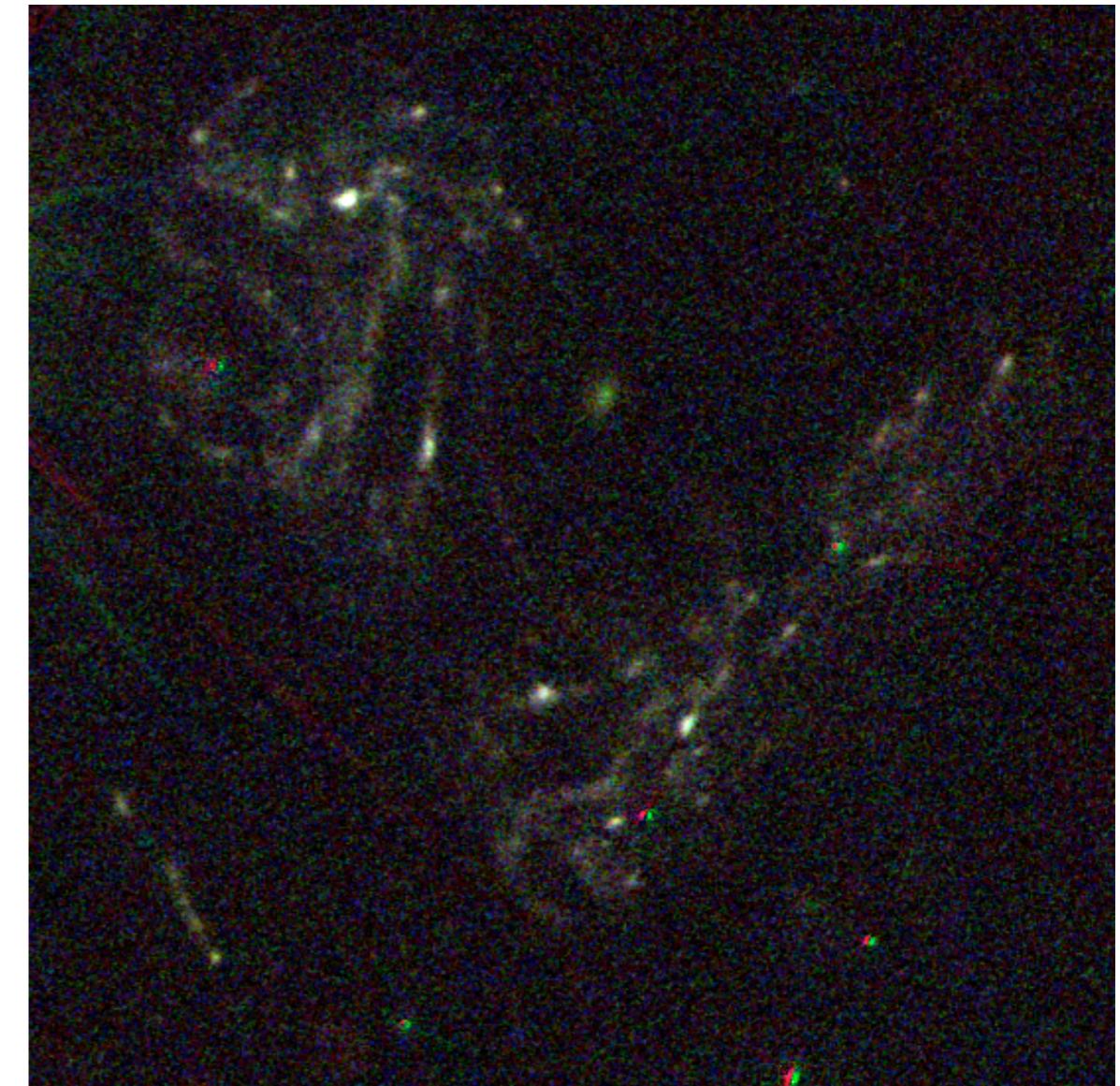
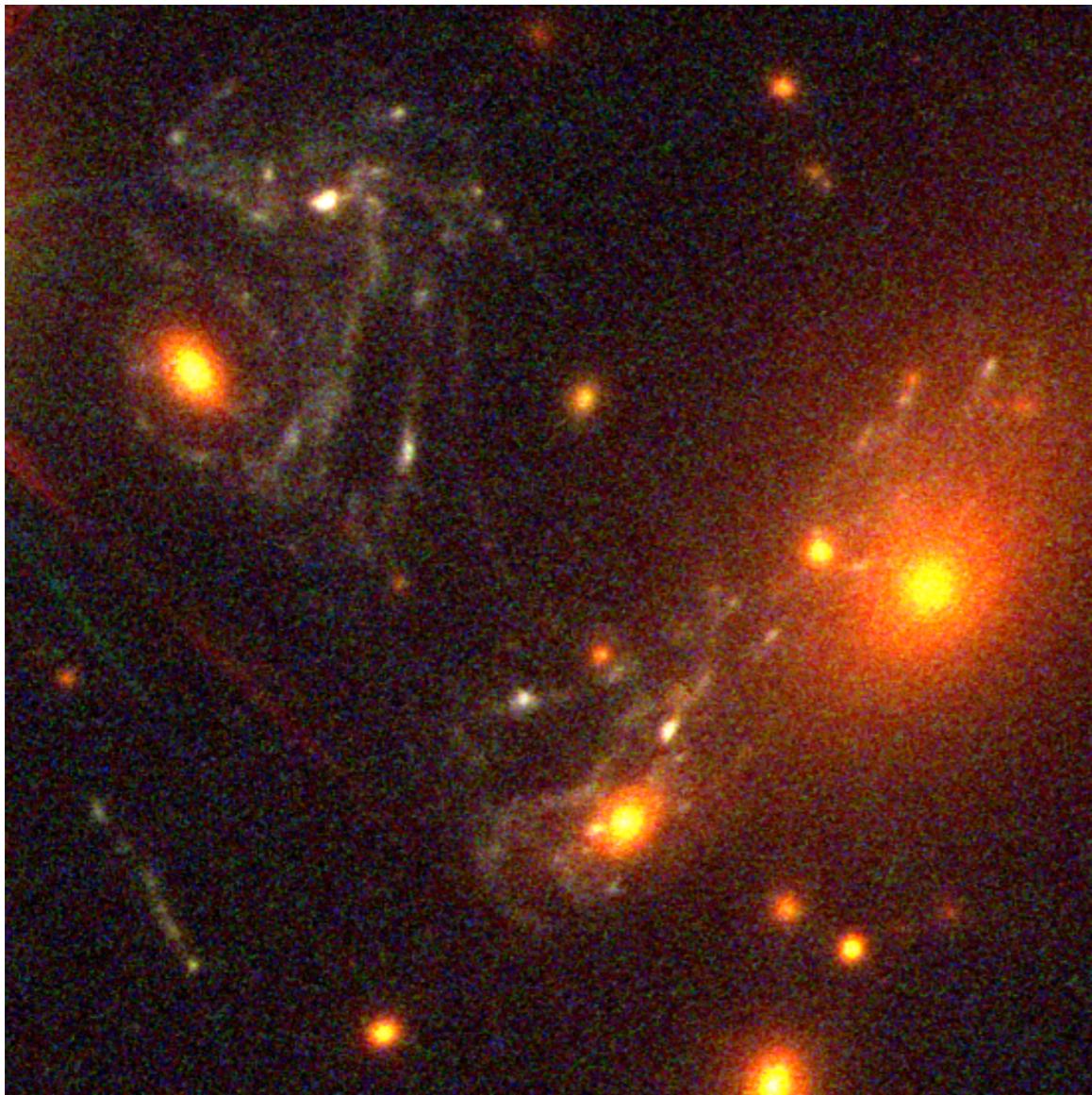
Realistic SEDs variation.

Multichannel data



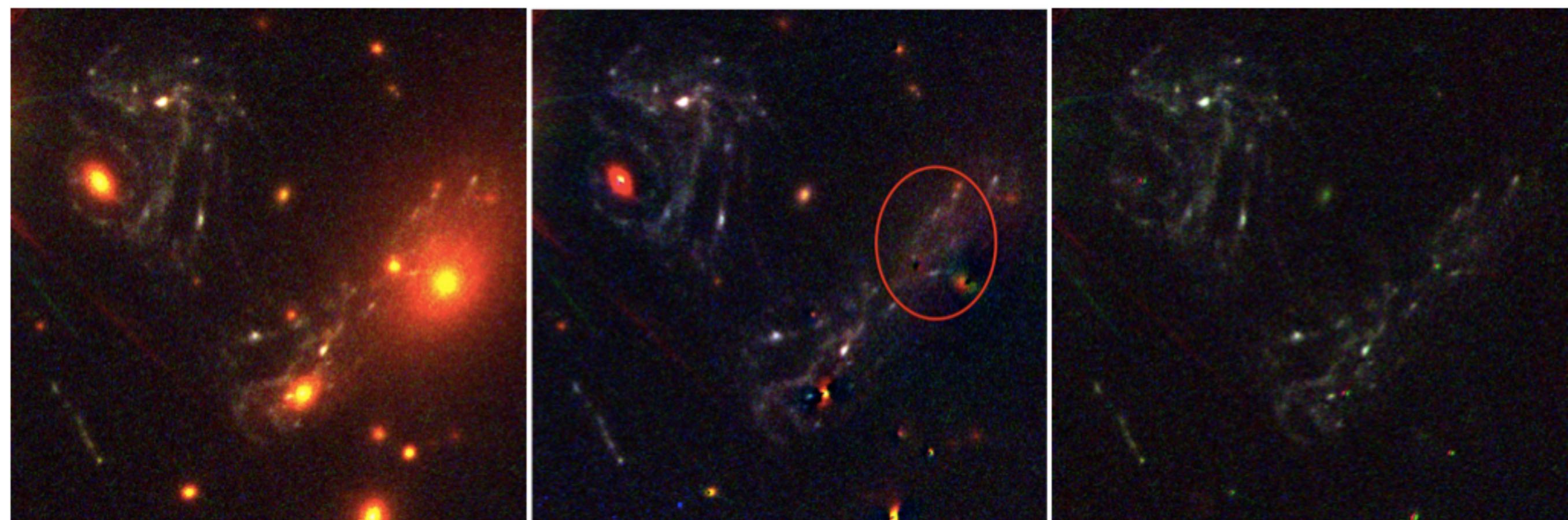
Multichannel data

galaxy cluster MACS~J1149+2223



Multichannel data

MACS~J1149+2223 cluster



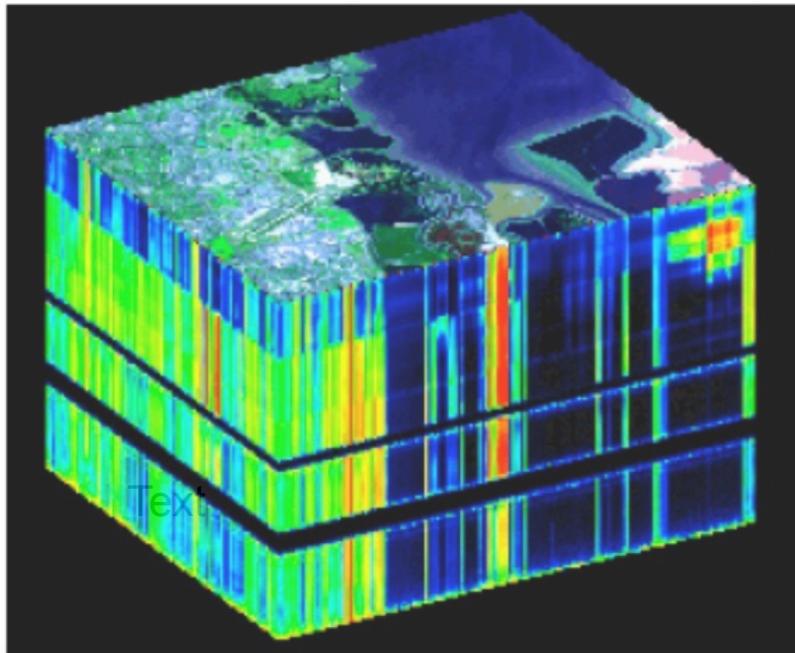
galfit subtraction of the galaxy members

Blind Source Separation

Marc LENNON

Département Image et Traitement de l'Information (ITI)

École Nationale Supérieure des Télécommunications de Bretagne



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

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

Need to add constraint

$$\min_{A,X} = \| Y - AX \|^2 \quad s.t. \quad \mathcal{C}(X, A)$$

Generalized MCA (GMCA)

- 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.



Source: $S = [s_1, \dots, s_n]$ Data: $X = [x_1, \dots, x_m] = AS$

We now assume that the sources are linear combinations of morphological components :

$$s_i = \sum_{k=1}^K c_{i,k} \quad \text{such that } \alpha_{i,k} = T_{i,k} c_{i,k} \text{ sparse}$$

$$\implies X_l = \sum_{i=1}^n A_{i,l} s_i = \sum_{i=1}^n A_{i,l} \sum_{k=1}^K c_{i,k}$$

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

$$\phi = [[\phi_{1,1}, \dots, \phi_{1,K}], \dots, [\phi_{n,1}, \dots, \phi_{n,K}]], \quad \alpha = S\phi^t = [[\alpha_{1,1}, \dots, \alpha_{1,K}], \dots, [\alpha_{n,1}, \dots, \alpha_{n,K}]]$$

GMCA aims at solving the following minimization:

$$\min_{A, c_{1,1}, \dots, c_{1,K}, \dots, c_{n,1}, \dots, c_{n,K}} = \sum_{l=1}^m \left\| X_l - \sum_{i=1}^n A_{i,l} \sum_{k=1}^K c_{i,k} \right\|_2^2 + \lambda \sum_{i=1}^n \sum_{k=1}^K \|T_{i,k} c_{i,k}\|_p$$

Sparse Component Separation: the GMCA Method

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

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

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

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

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

BSS experiment : Noiseless case

Original Sources



2 of 4 Mixtures



Noiseless experiment, 4 random mixtures, 4 sources

GMCA Experiment

•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.



GMCA Experiment

• 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.



GMCA Experiment

•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.



GMCA Experiment

• 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.



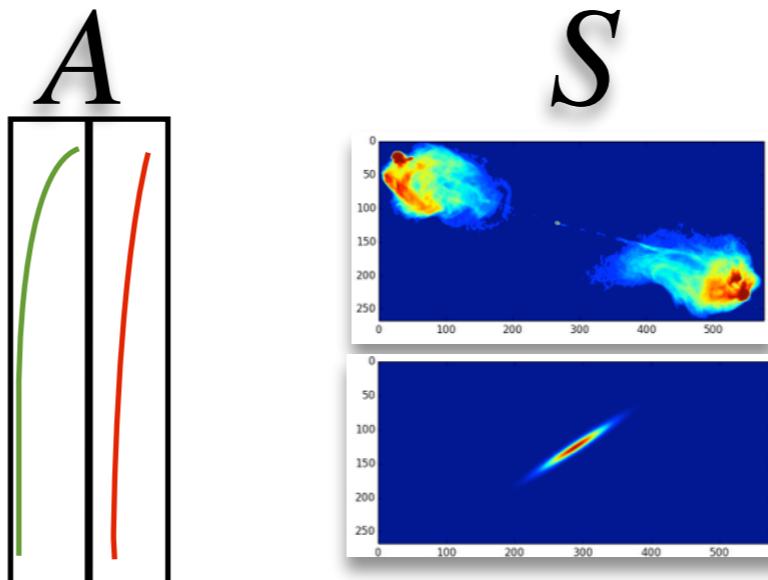
GMCA Experiment

•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.



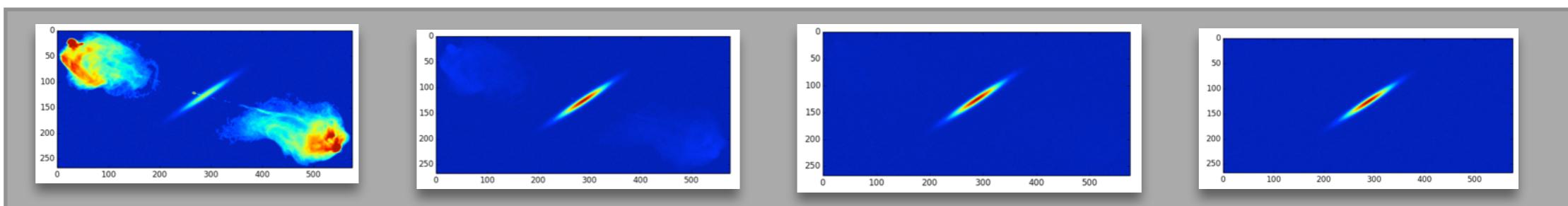
Deconvolved Blinded Source Separation

Ground Truth



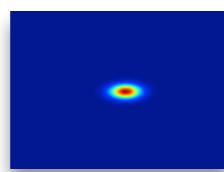
Mixtures

$$X = AS$$

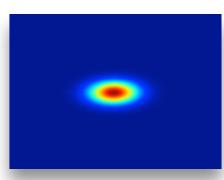


$$\text{PSF } H$$

chan 1

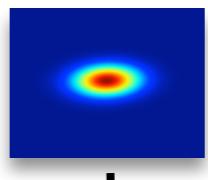


chan 4

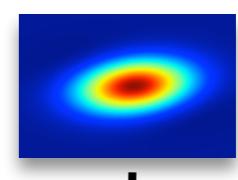


*

chan 7

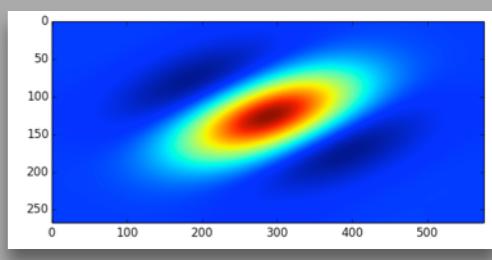
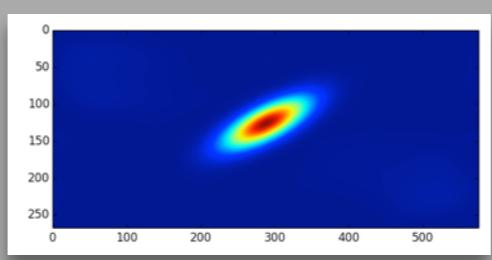
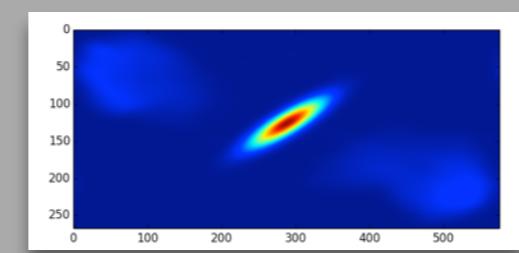
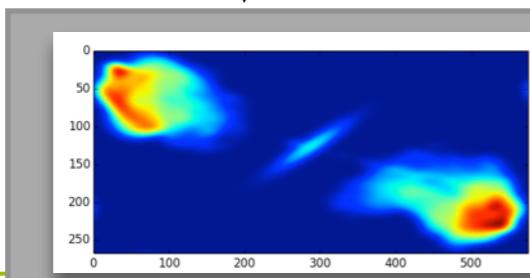


chan 10

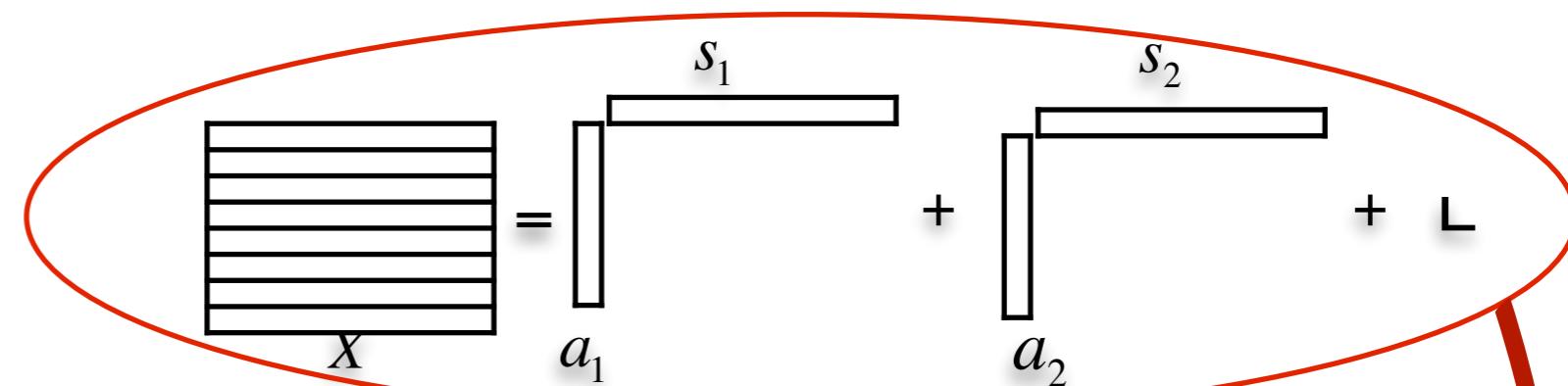


Data

$$Y = HX + N$$



Deconvolved Blinded Source Separation



- BSS(Blind Source Separation) problem

Statistical approach: ICA (FastICA(*A. HYVÄRINEN et al.*)), etc.

Methods based on morphological diversity: GMCA(*J. BOBIN et al.*) and its variations

- Deconvolution

$$Y = H(X) \Rightarrow X = H^{-1}(Y)$$

e.g. ForWaRD(*R.N. NEELAMANI et al.*)

Joint BSS and Deconvolution?

Very few literatures!

Our method: ForWaRD+GMCA = fGMCA

- Problem formulation

$$\min_{\{S_{j,g}\}, \{A_v\}} \frac{1}{2} \sum_{v,k} \|\hat{Y}_{v,k} - H_{v,k} A_v \hat{S}_k\|_2^2 + \sum_j \lambda_j \|S_{j,g} \Phi^t\|_0$$

$$\hat{Y} = H e \begin{matrix} A_v \\ \vdots \\ A \end{matrix} \hat{S} + N$$

\hat{S}

A_v has a green block in the top-right corner.
 \hat{S}_k has a red block in the middle-left column.

- Estimation of S

$$\min_{\{S_{j,g}\}} \frac{1}{2} \sum_{v,k} \|\hat{Y}_{v,k} - H_{v,k} A_v \hat{S}_k\|_2^2 + \sum_j \lambda_j \|S_{j,g} \Phi^t\|_0$$

$$\hat{S}_k = (\sum_v (H_{v,k} A_v)^t (H_{v,k} A_v) + \epsilon' \mathbf{I}_n)^{-1} \sum_v H_{v,k} \hat{Y}_{v,k} A_v^t$$

$$S_{wt} = S \Phi^t$$

$$S_{wt} = HT_\lambda(S_{wt})$$

Sparsity priori in
wavelet domain

Tikhonov parameter to
avoid non inversion

- Estimation of A

$$\min_{\{A_v\}} \frac{1}{2} \sum_{v,k} \|\hat{Y}_{v,k} - H_{v,k} A_v \hat{S}_k\|_2^2$$

$$A_v = (\sum_k H_{v,k} \hat{Y}_{v,k} \hat{S}_k^*) (\sum_k (H_{v,k} \hat{S}_k) (H_{v,k} \hat{S}_k)^*)^{-1}$$

normalize the columns of A

- fGMCA algorithm

$$\min_{\{S_{j,g}\}, \{A_v\}} \frac{1}{2} \sum_{v,k} \|\hat{Y}_{v,k} - H_{v,k} A_v \hat{\mathbf{S}}_k\|_2^2 + \sum_j \lambda_j \|S_{j,g} \Phi^t\|_0$$

- Initialize $A^{(0)}$

- Iterate $i=1, \dots, N_{\text{iter}}$

- Update S knowing A

$$\min_{\{S_{j,g}\}} \frac{1}{2} \sum_{v,k} \|\hat{Y}_{v,k} - H_{v,k} A_v \hat{\mathbf{S}}_k\|_2^2 + \sum_j \lambda_j \|S_{j,g} \Phi^t\|_1$$

- Update A knowing S

$$\min_{\{A_v\}} \frac{1}{2} \sum_{v,k} \|\hat{Y}_{v,k} - H_{v,k} A_v \hat{\mathbf{S}}_k\|_2^2$$

- Decrease the thresholding λ

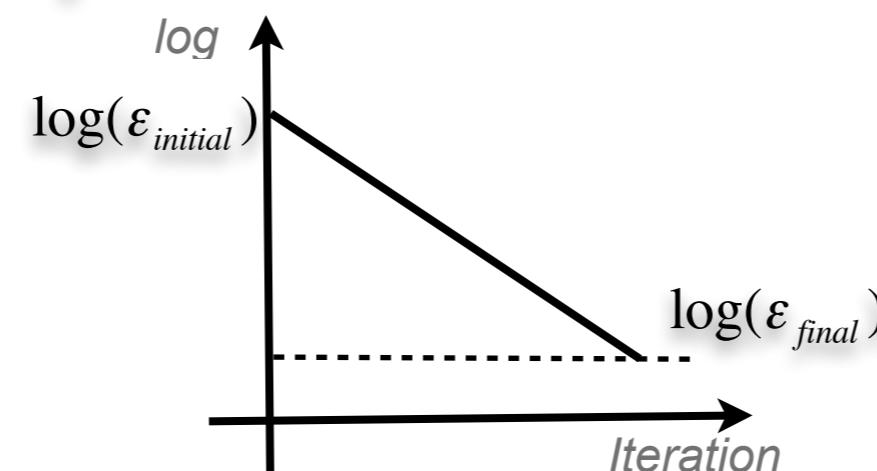
- Decrease the Tikhonov parameter ϵ

- Choice of ε'

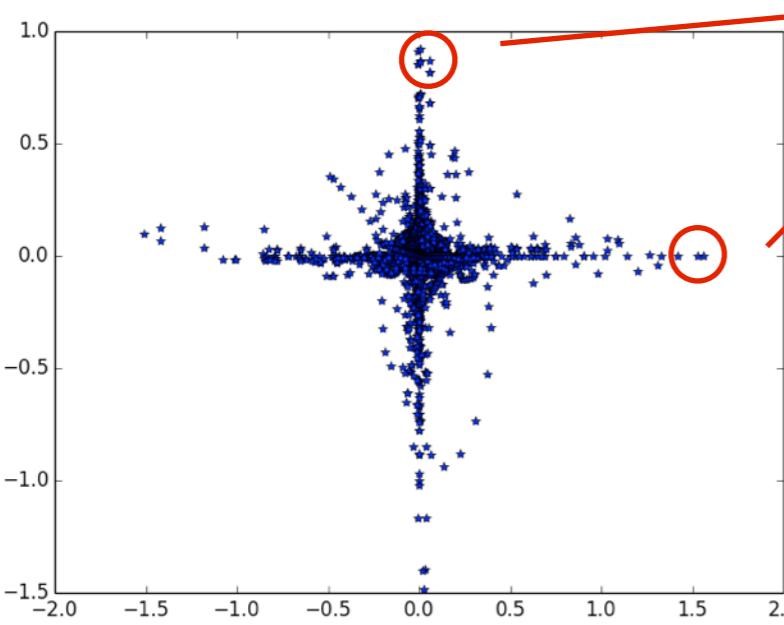
$$\hat{\mathbf{S}}_k = \left(\sum_v (H_{v,k} \mathbf{A}_v)^t (H_{v,k} \mathbf{A}_v) + \varepsilon' \mathbf{I}_n \right)^{-1} \sum_v H_{v,k} \hat{Y}_{v,k} \mathbf{A}_v^t$$

different frequency index k , the condition number is different.

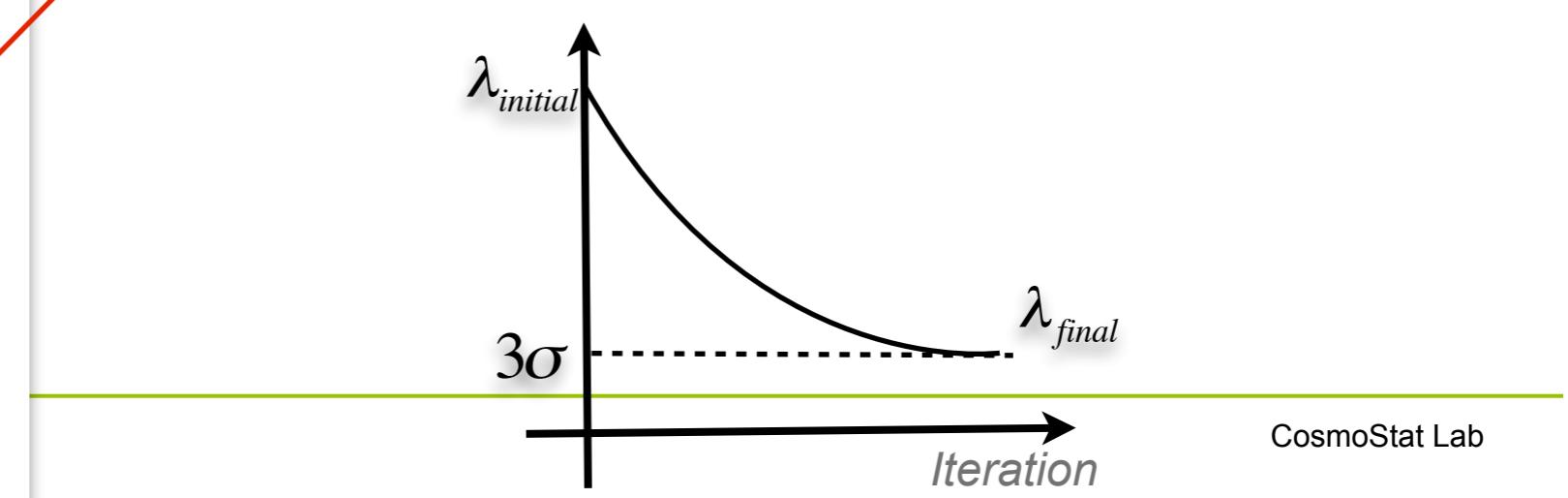
$$\varepsilon' = \varepsilon \text{SP}\left(\sum_v (H_{v,k} \mathbf{A}_v)^t (H_{v,k} \mathbf{A}_v)\right) \quad \text{SP}(\mathbf{M}) \text{ is spectral radius of matrix } \mathbf{M}$$



- Choice of λ

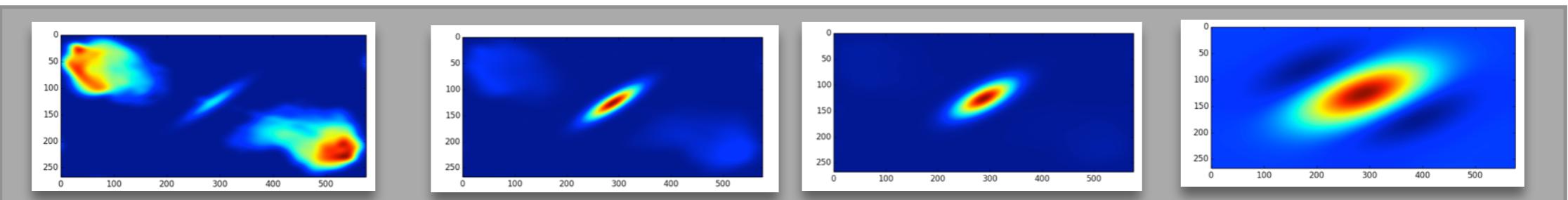


Larger initial threshold to select significant feature, easier to separate sources

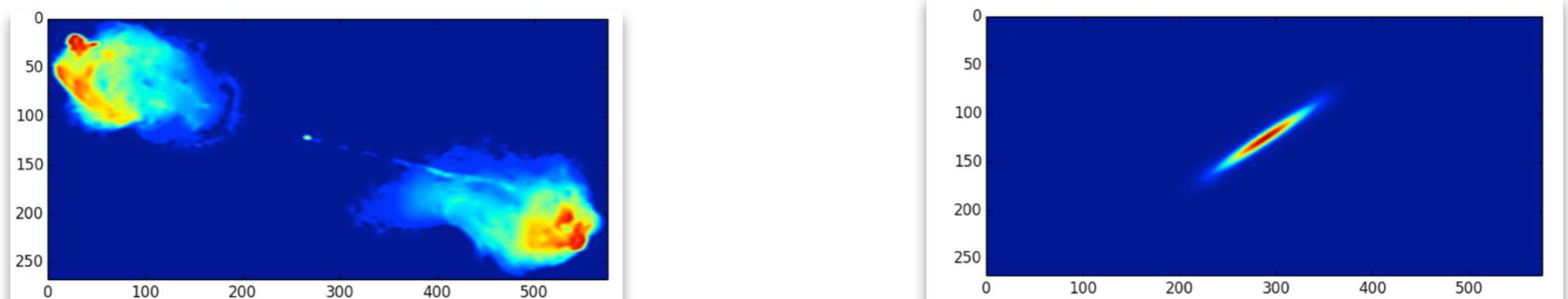


Experiments(Source reconstruction)

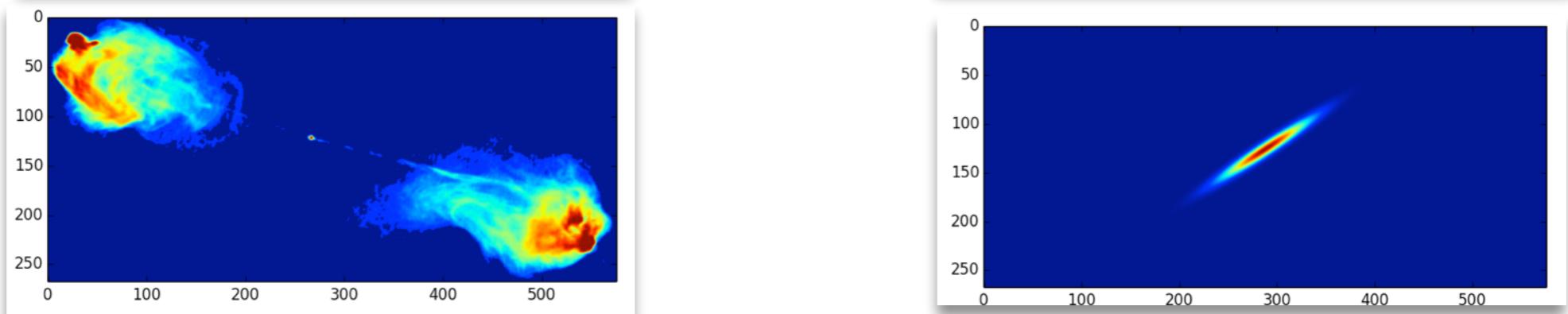
Data



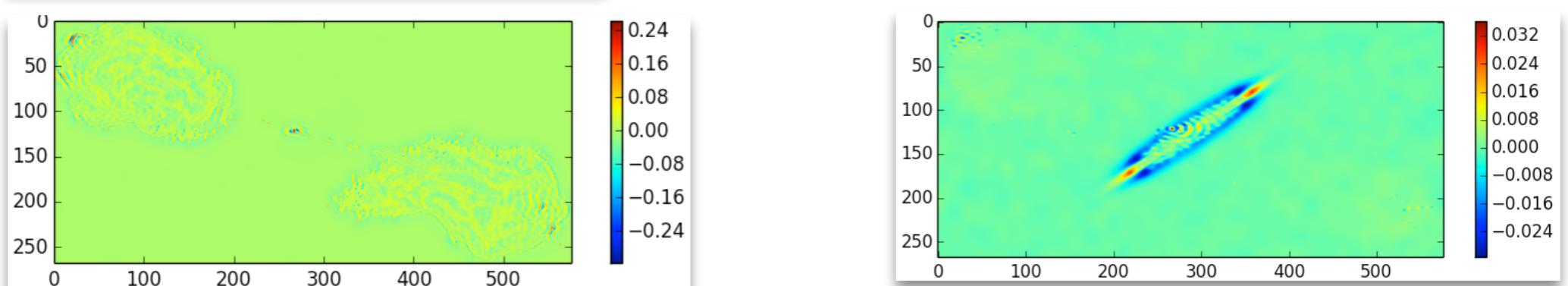
Reconstruction



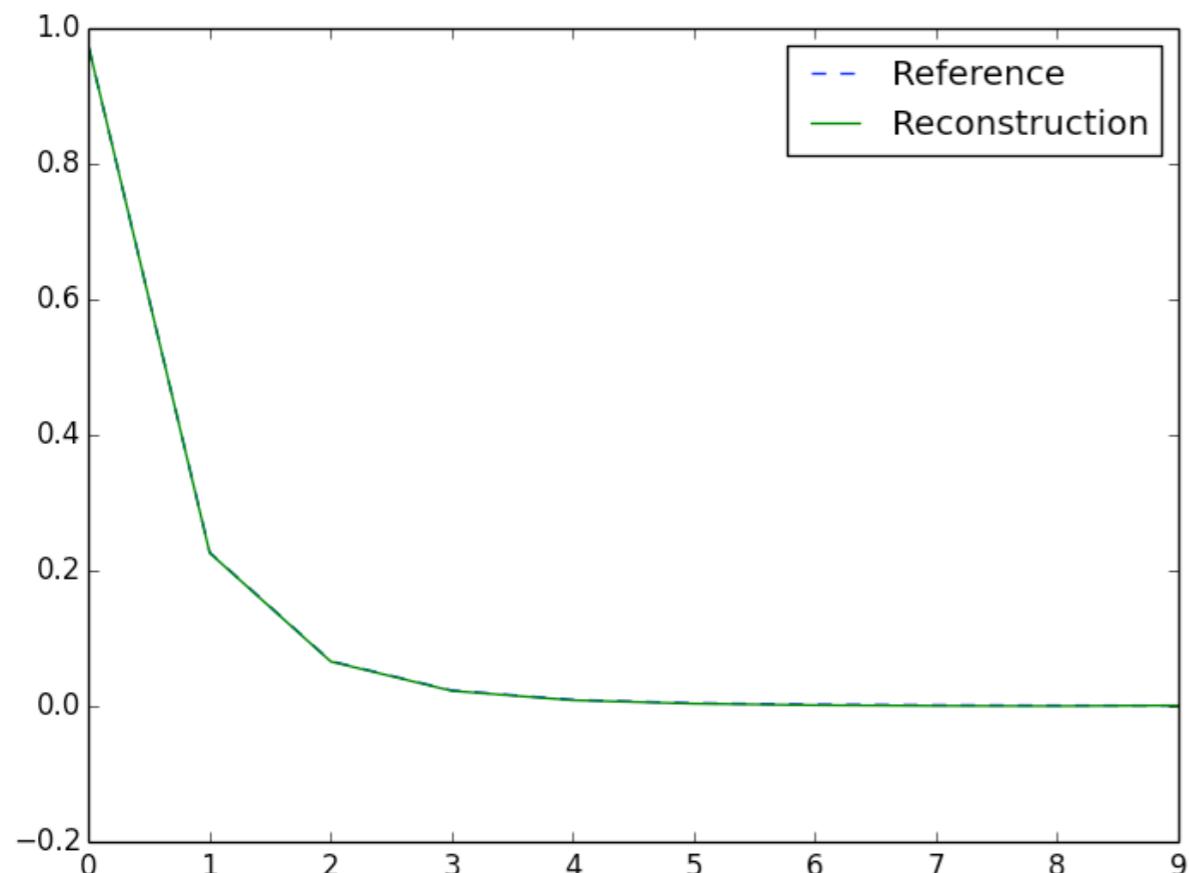
Truth



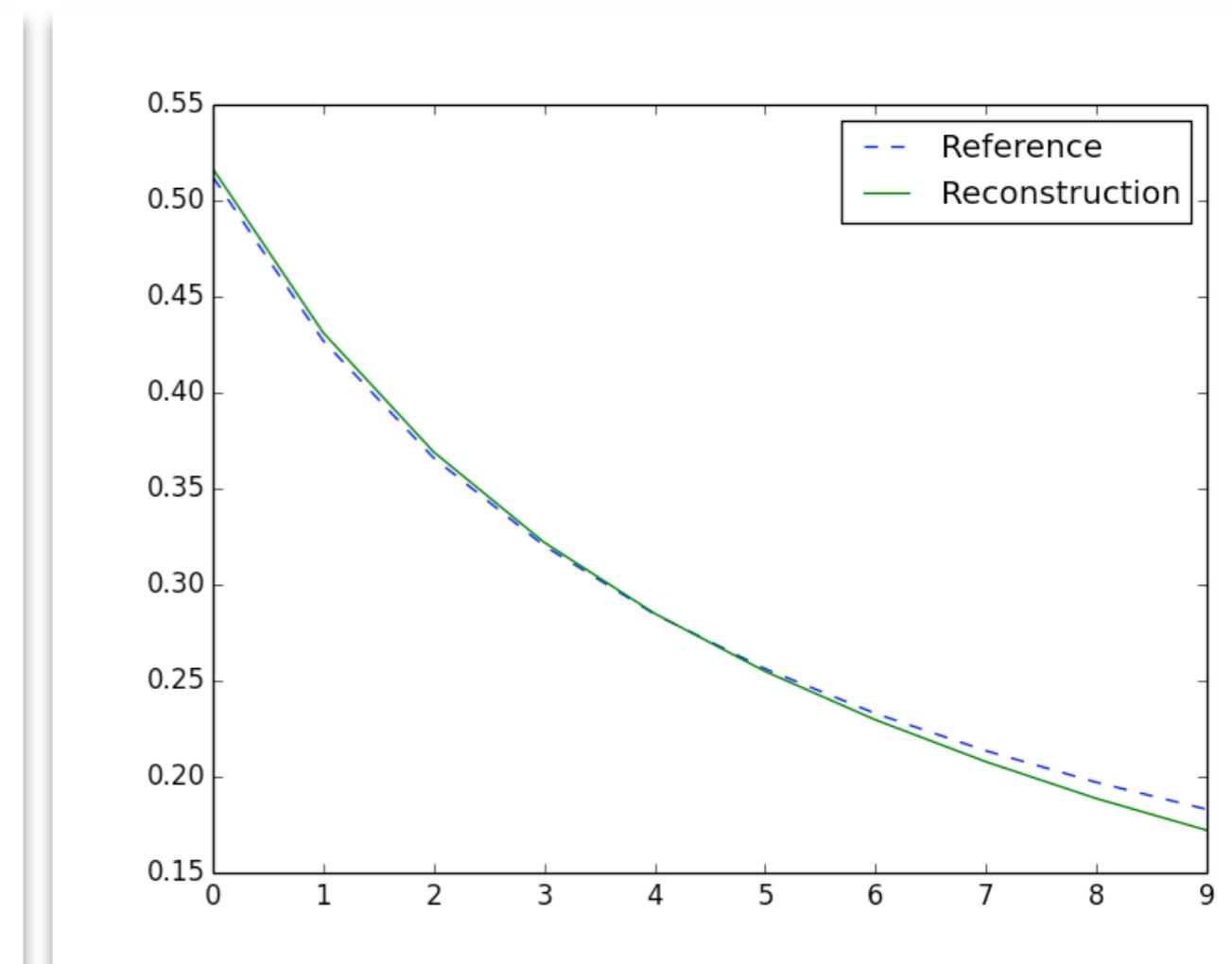
Error



Experiments(Spectra reconstruction)



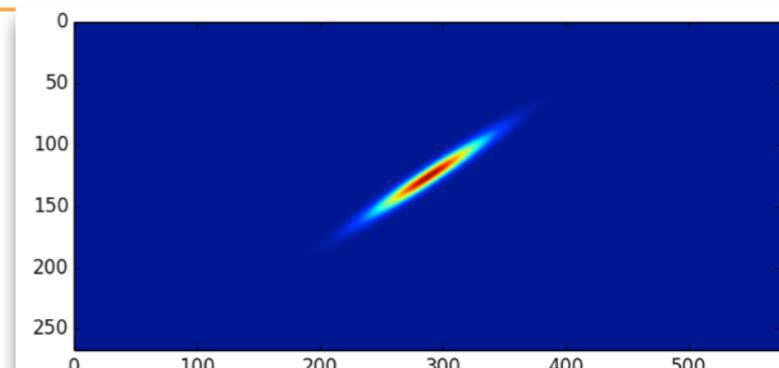
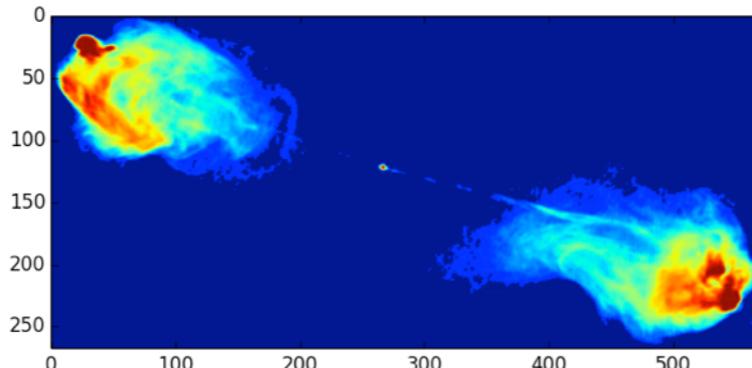
Reconstructed spectrum of S_0 v.s reference



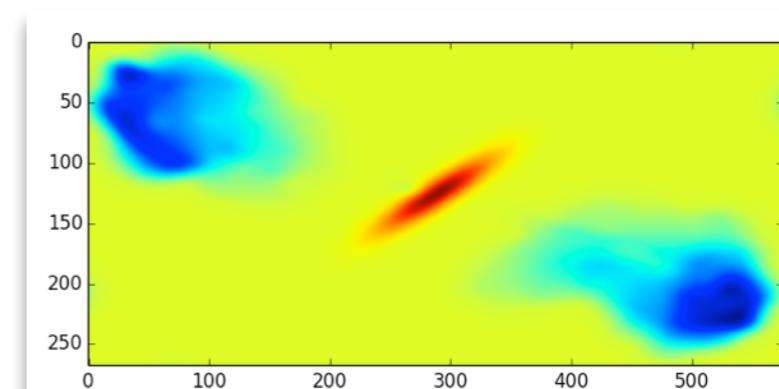
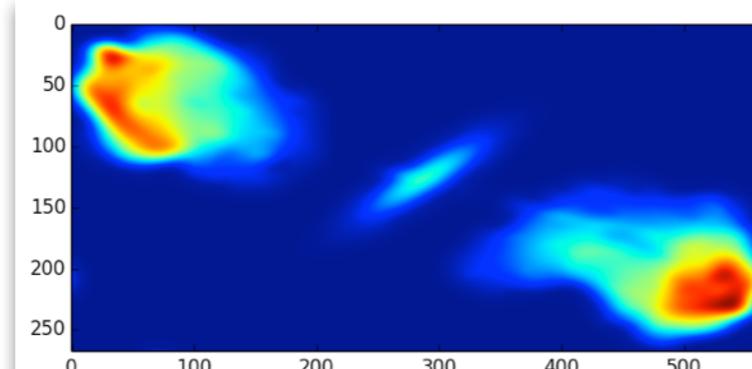
Reconstructed spectrum of S_1 v.s reference

Experiments(Source reconstruction)

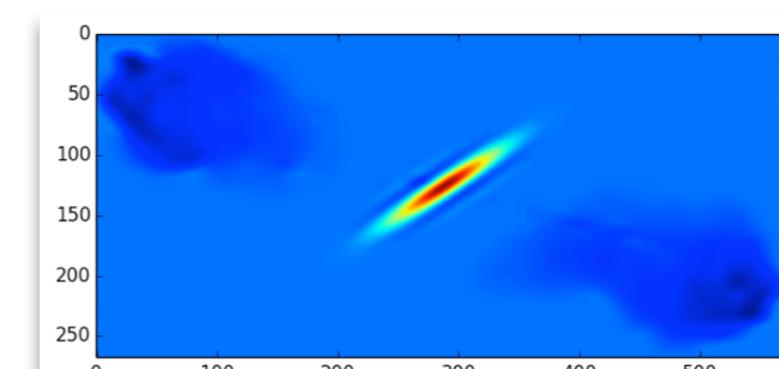
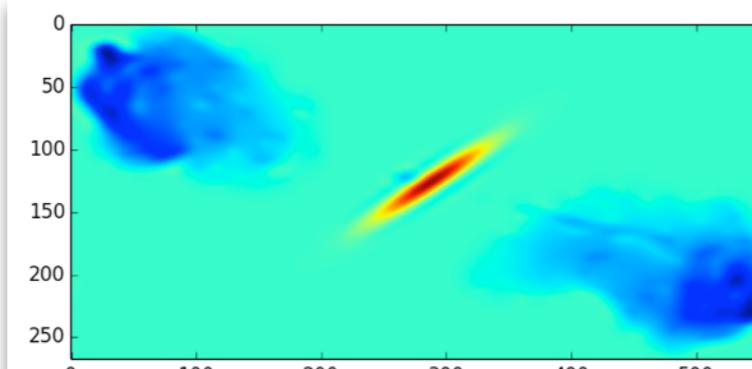
Model sources



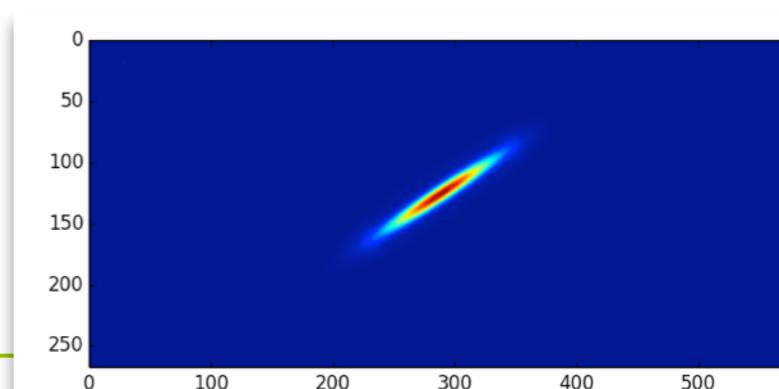
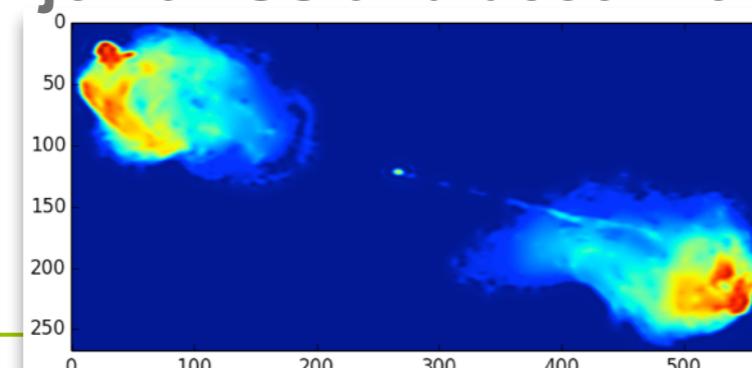
BSS only (GMCA), no deconvolution



Channel by channel deconvolution (ForWaRD) followed by a BSS (GMCA)



Our method fGMCA : joint BSS and deconvolution



- Part I: Introduction to Inverse Problems in Astrophysics
- Part II: Sparsity & Dictionaries
- Part III: Sparse Regularization
- Part IV: Unmixing
- Part V: Sparsity for Planck and Euclid Space Missions

Sparsity and the PLANCK Project

Launching of Planck

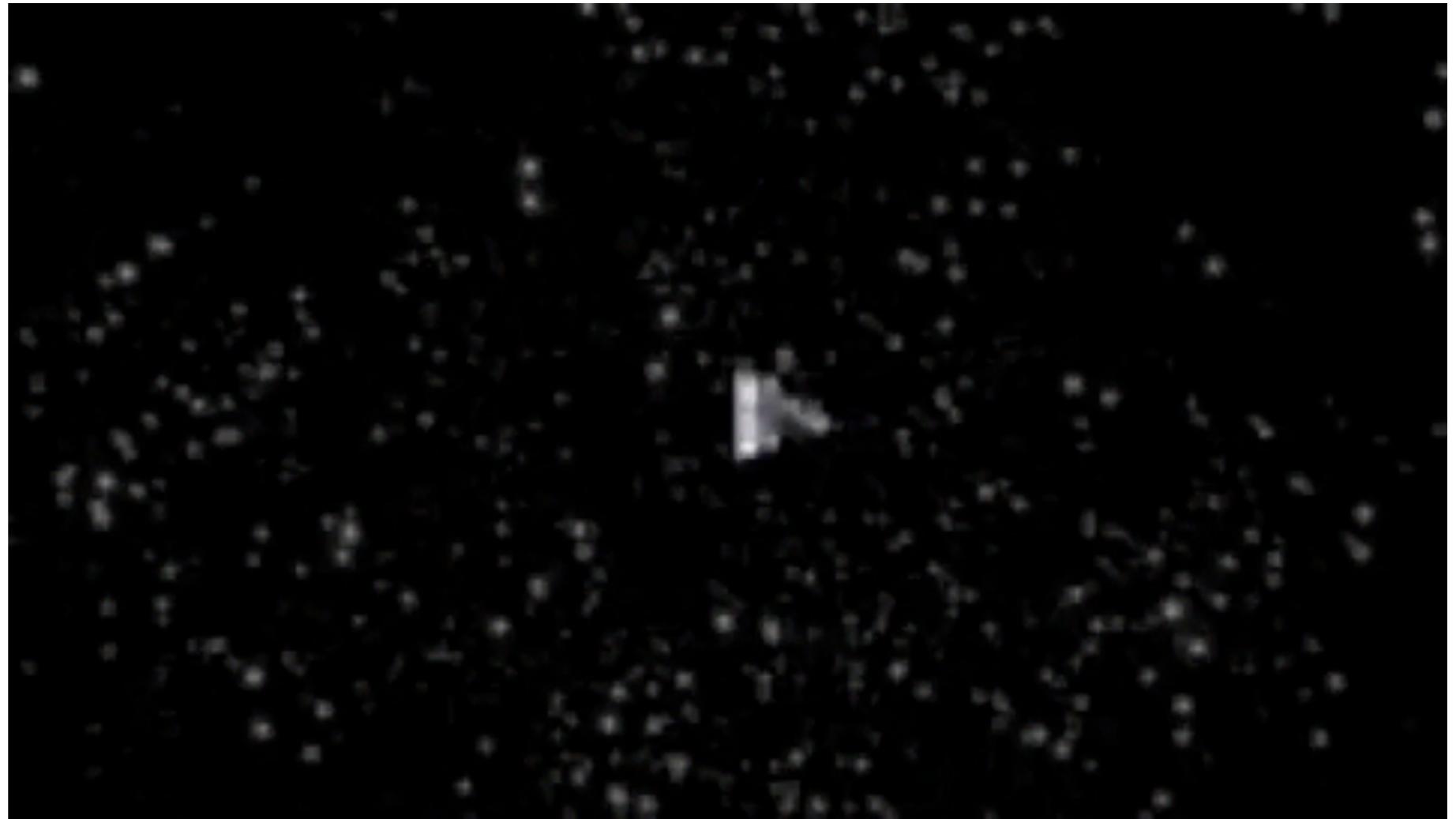
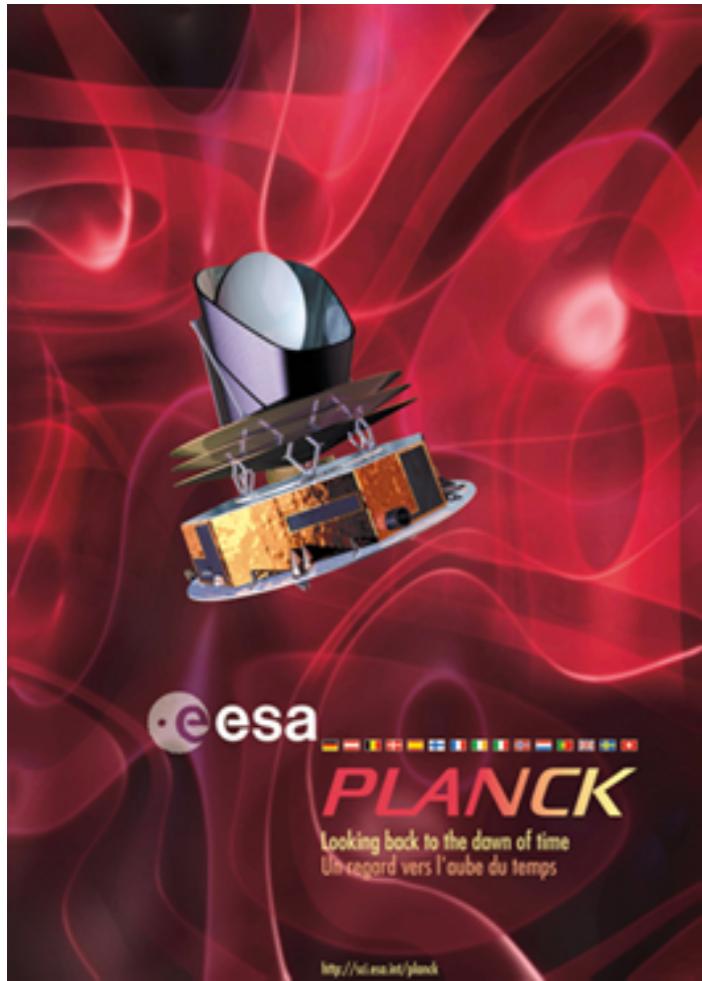
Ariane 5 : May, 14, 2009



Planck Mission Duration : 2.5 years

Until end of 2011

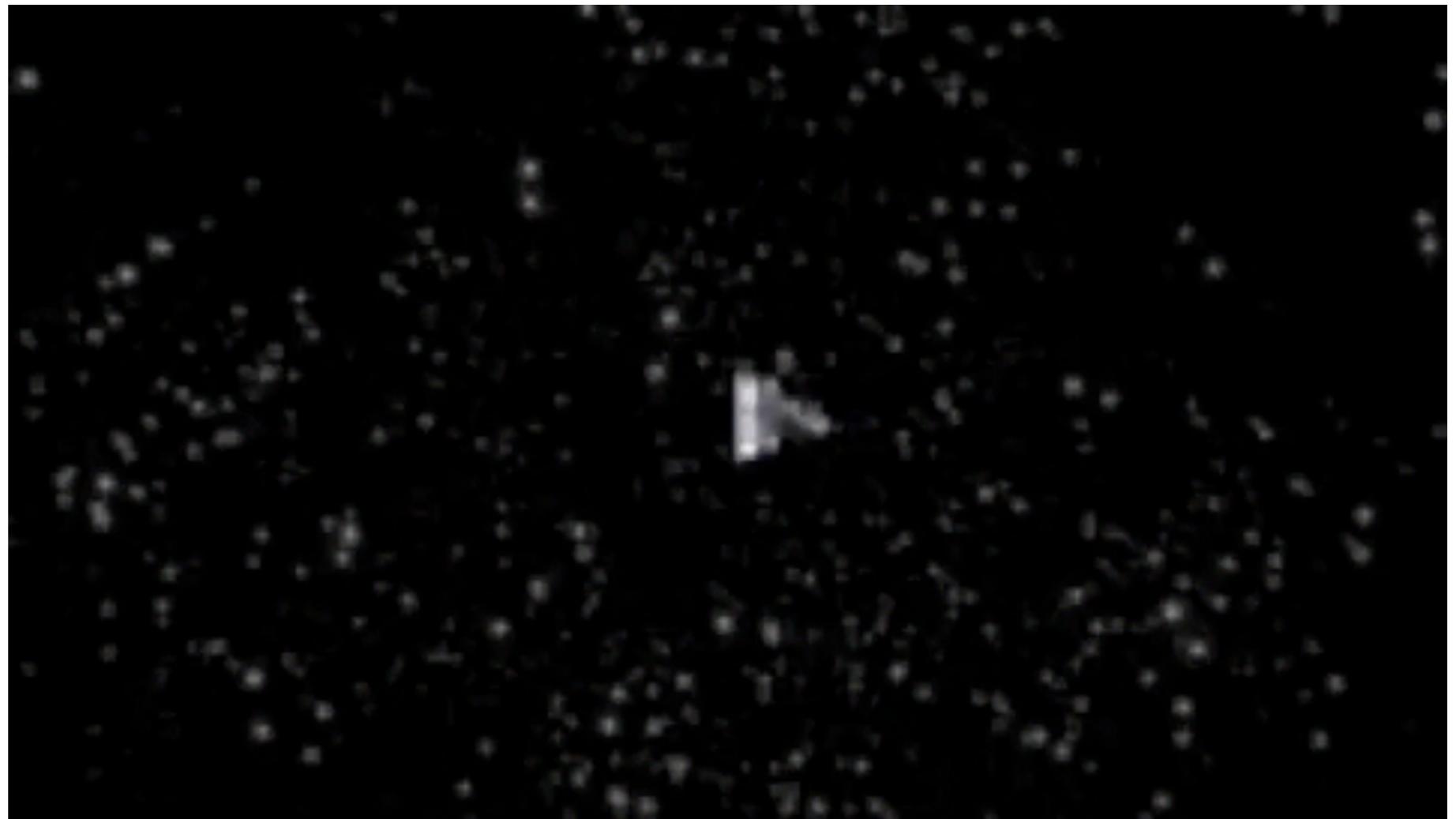
PLANCK and Sparsity



- 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'

==> PR1 DATA Released in 2013; PR2 DATA Released in 2015

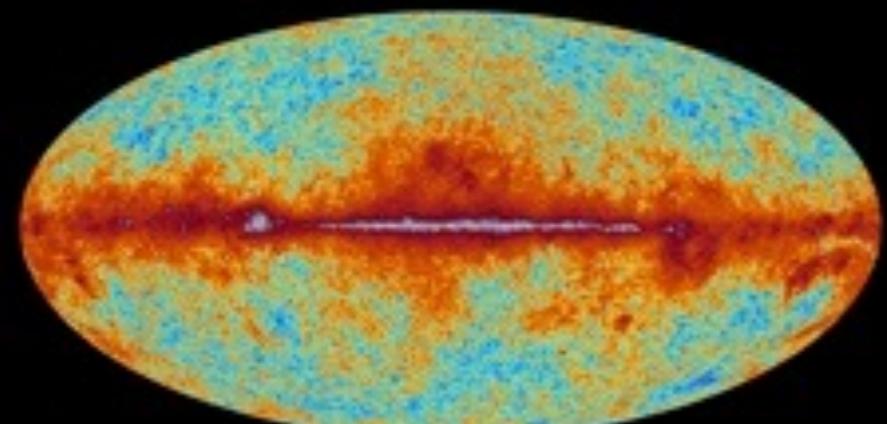
PLANCK and Sparsity



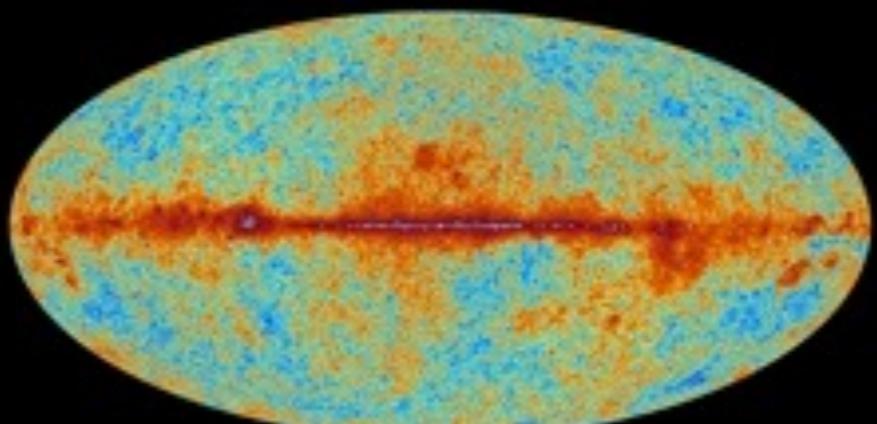
- Successor of WMAP (better resolution, better sensitivity, more channels)
- Launched on May 14, 2009
- Two instruments LFI and HFI
- Nine Temperature maps at 30,44,70,100,143,217,353,545,857 GHz + Polarization
- Angular resolutions: 33', 24', 14', 10', 7.1', 5', 5', 5', 5'

==> PR1 DATA Released in 2013; PR2 DATA Released in 2015

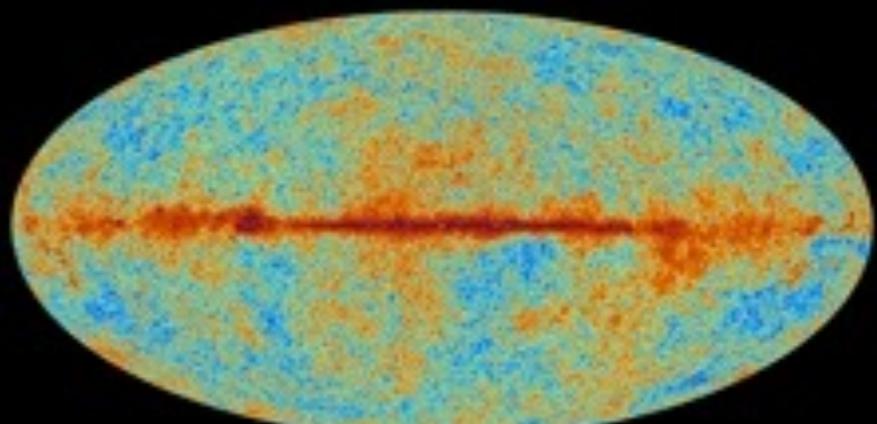
The sky as seen by Planck



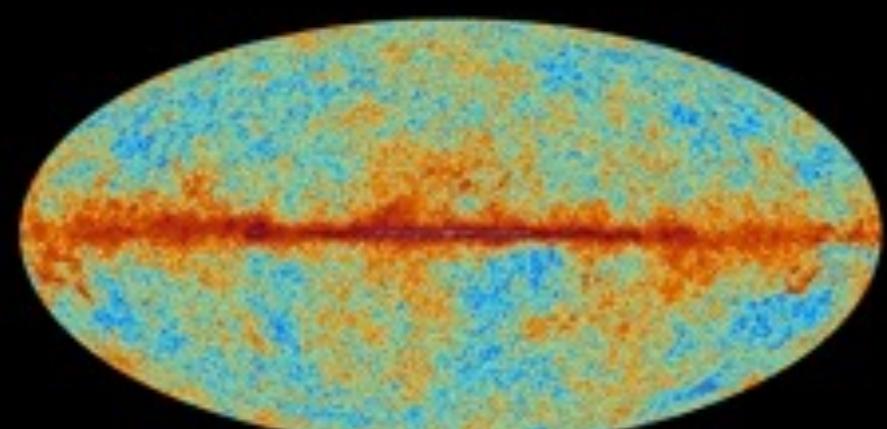
30 GHz



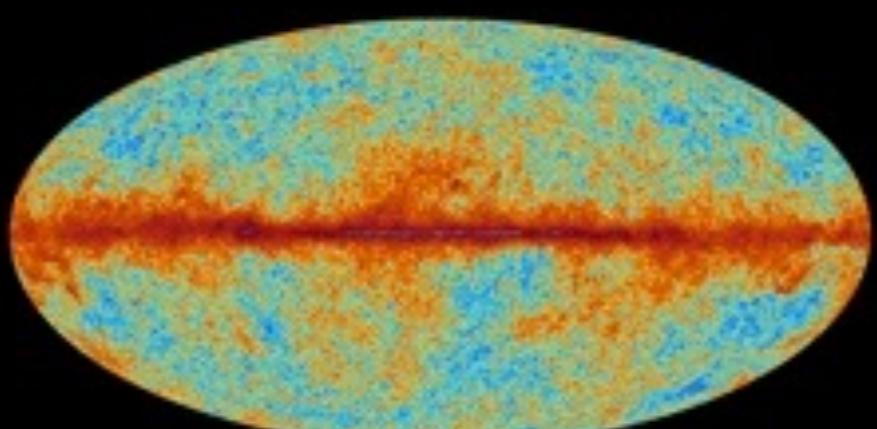
44 GHz



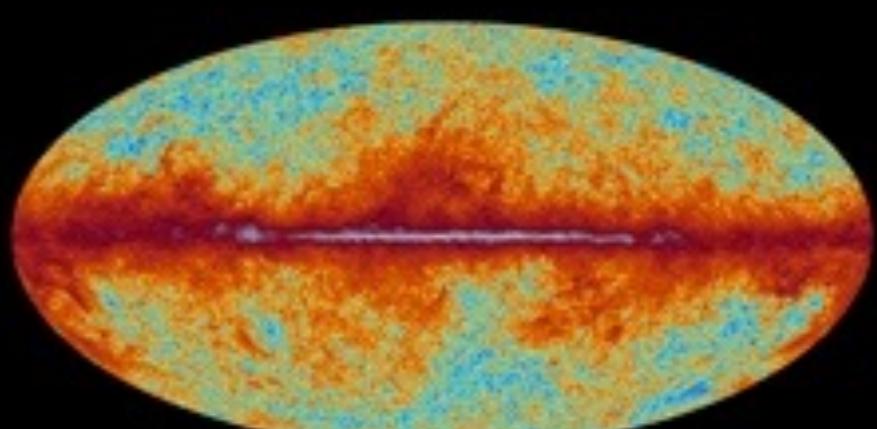
70 GHz



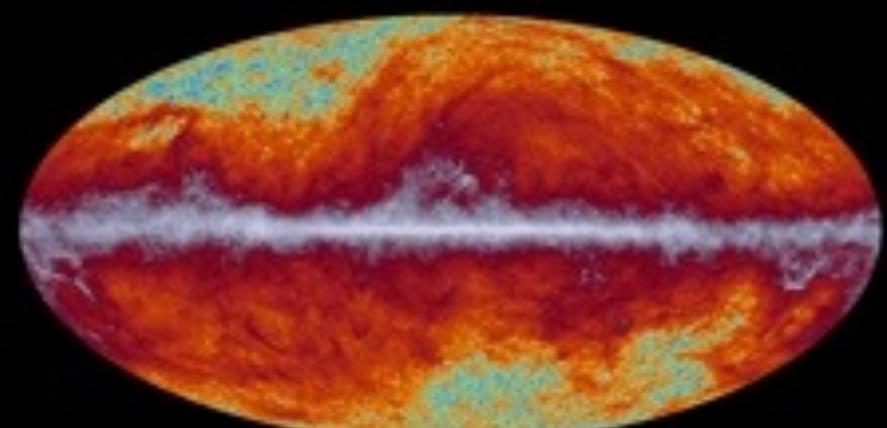
100 GHz



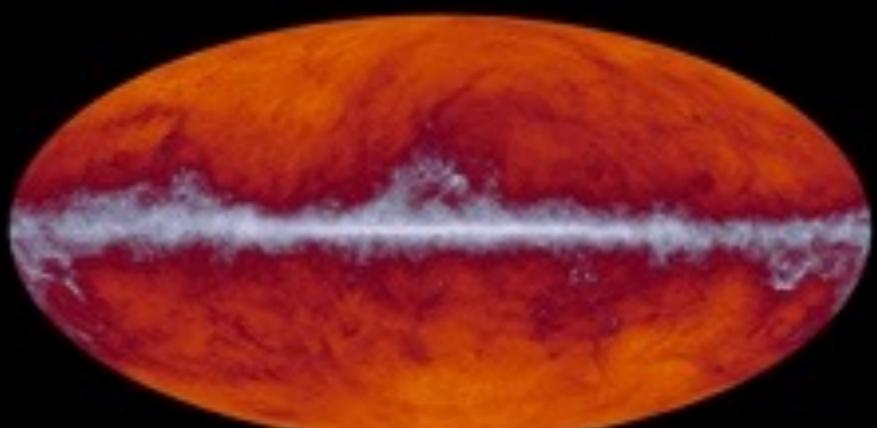
143 GHz



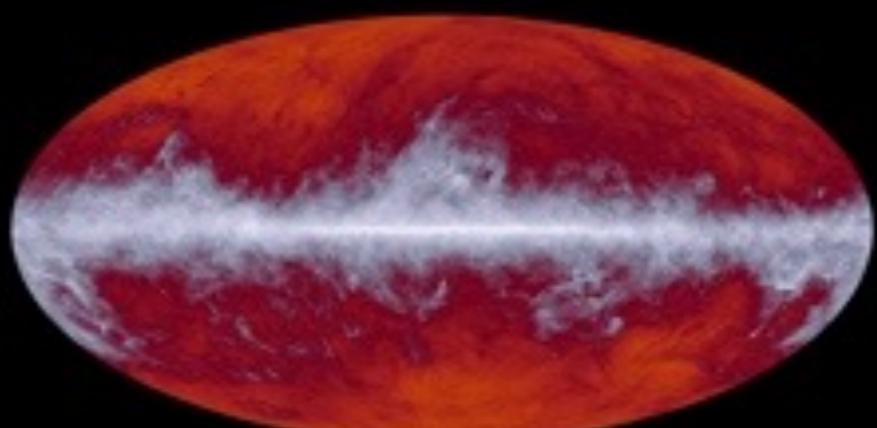
217 GHz



353 GHz



545 GHz

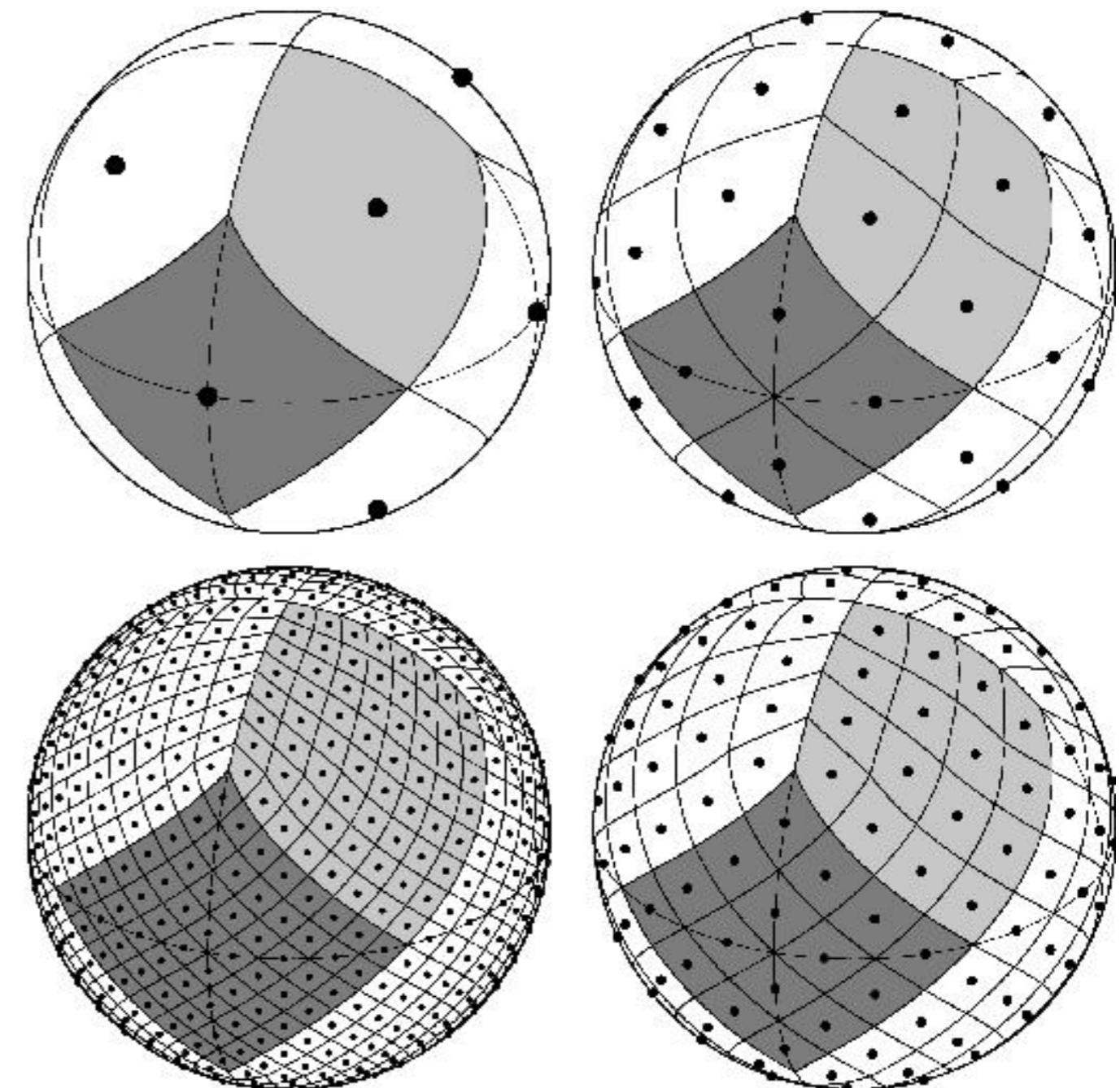


857 GHz

Healpix

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

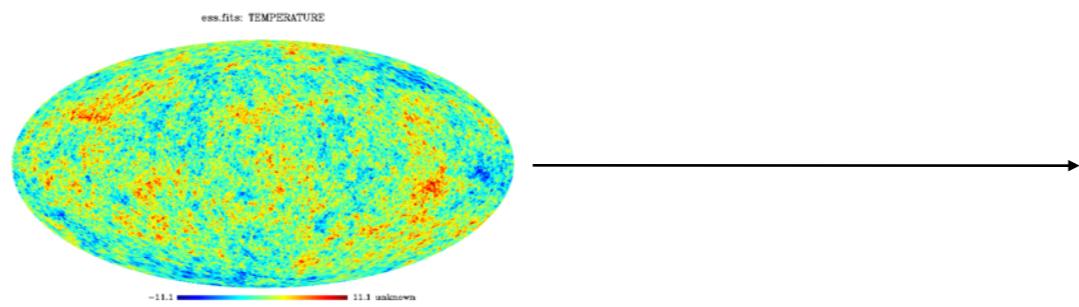
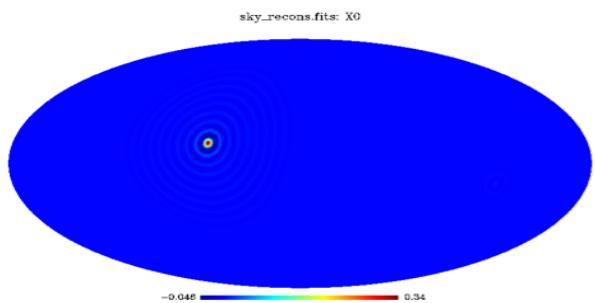
- Pixel = Rhombus
- Same Surfaces
- For a given latitude : regularly spaced
- Number of pixels:
 $12 \times (N_{\text{sides}})^2$
- Included in the software:
 - Anafast
 - Synfast



Isotropic Undecimated Wavelet on the Sphere

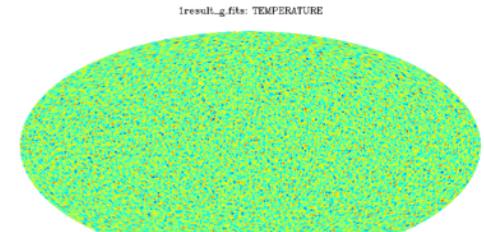
Wavelets, Ridgelets and Curvelets on the Sphere, Astronomy & Astrophysics, 446, 1191-1204, 2006.

Undecimated
Wavelet Transform

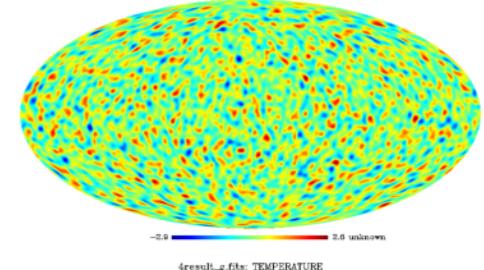


$$\hat{\psi}_{\frac{l_c}{2^j}}(l, m) = \hat{\phi}_{\frac{l_c}{2^{j+1}}} - \hat{\phi}_{\frac{l_c}{2^j}}(l, m)$$

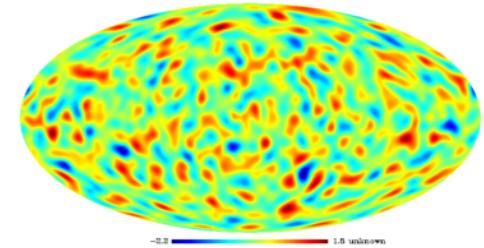
j=1



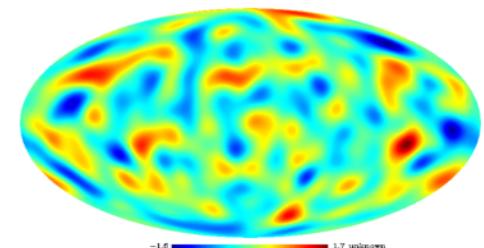
j=2



j=3



j=4



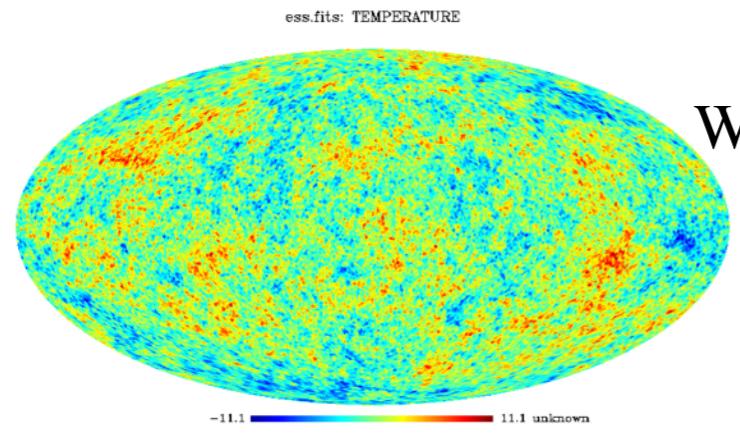
$$\hat{H}_j(l, m) = \begin{cases} \frac{\hat{\phi}_{\frac{l_c}{2^{j+1}}}(l, m)}{\hat{\phi}_{\frac{l_c}{2^j}}(l, m)} & \text{if } l < \frac{l_c}{2^{j+1}} \text{ and } m = 0 \\ 0 & \text{otherwise} \end{cases}$$

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

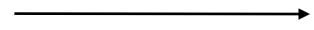
$$\hat{c}_{j+1}(l, m) = \hat{H}_j(l, m)\hat{c}_j(l, m)$$

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

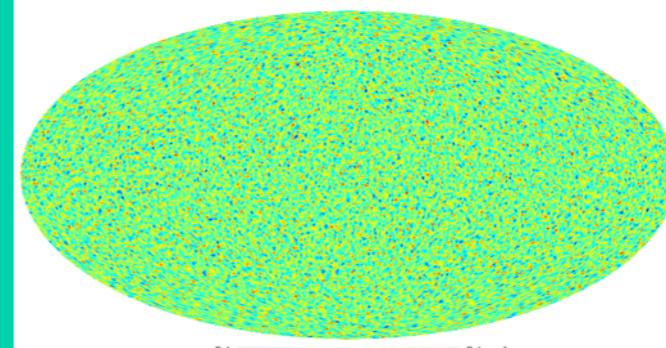
$$c_0(\vartheta, \varphi) = c_J(\vartheta, \varphi) + \sum_{j=1}^J w_j(\vartheta, \varphi)$$



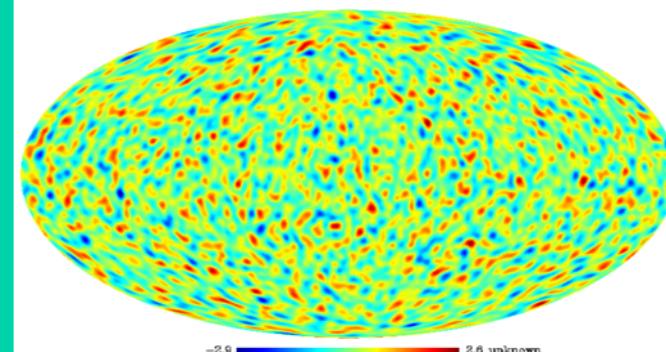
Undecimated
Wavelet Transform



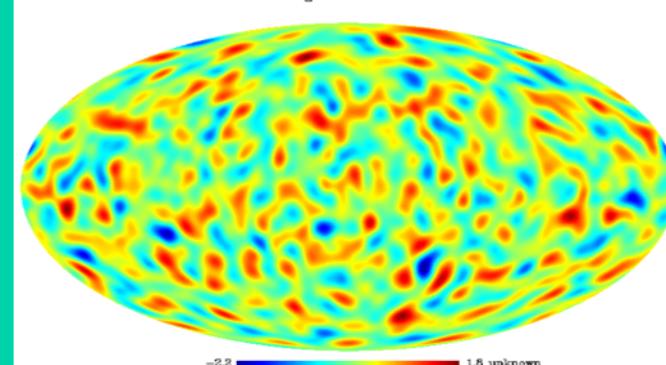
j=1



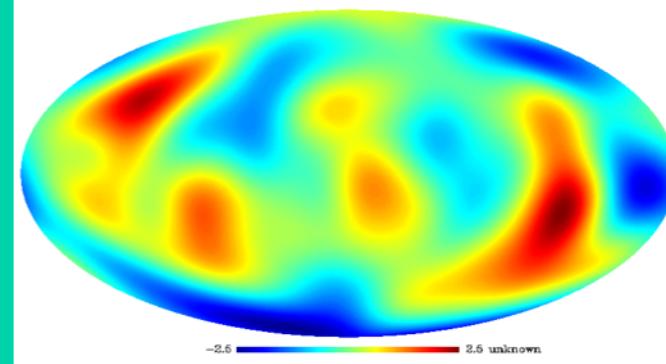
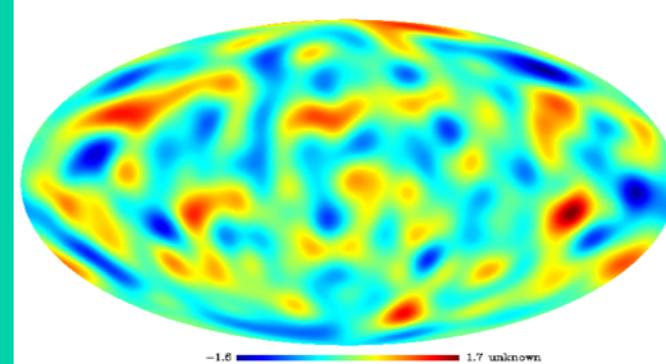
j=2



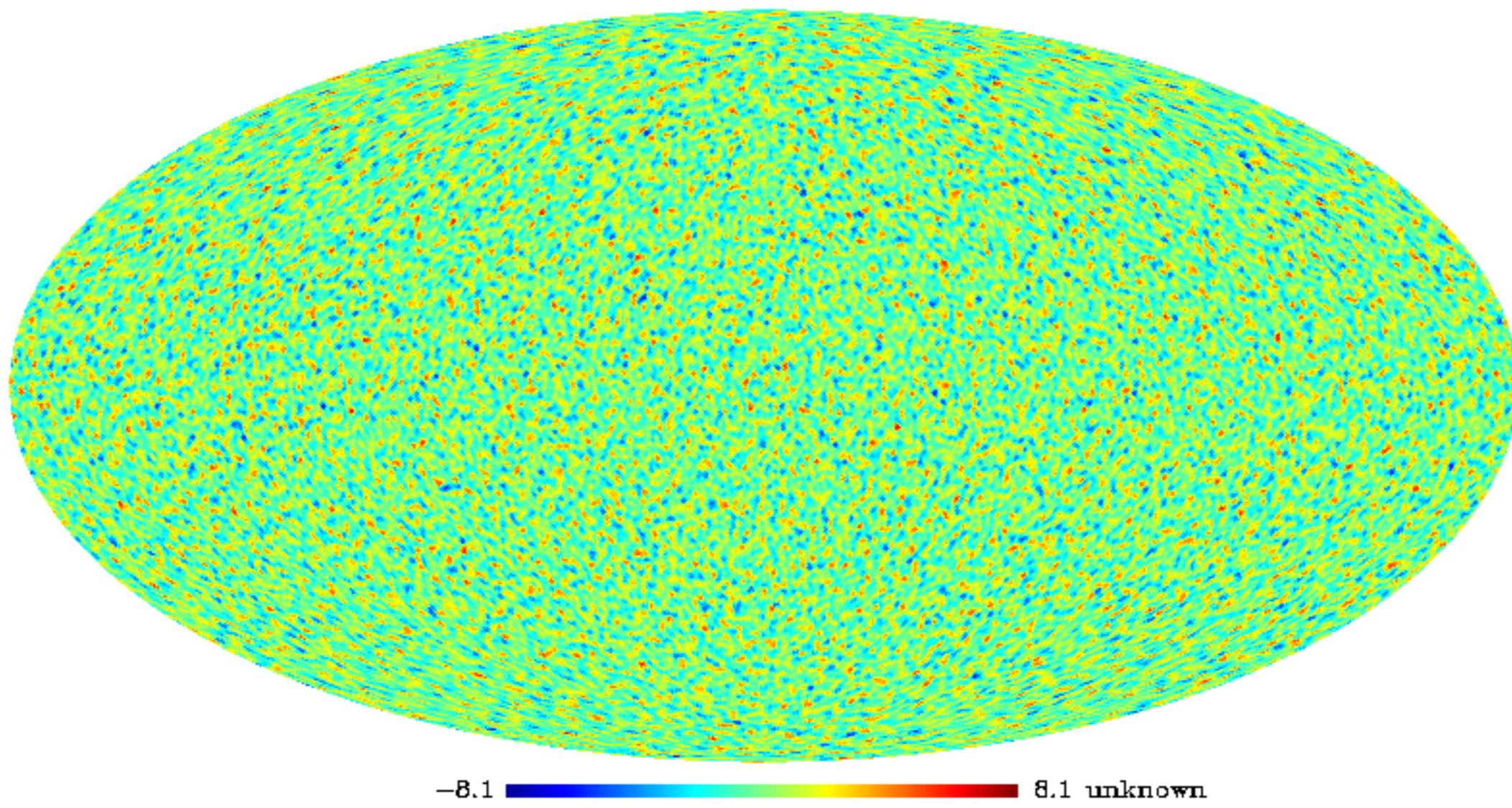
j=3



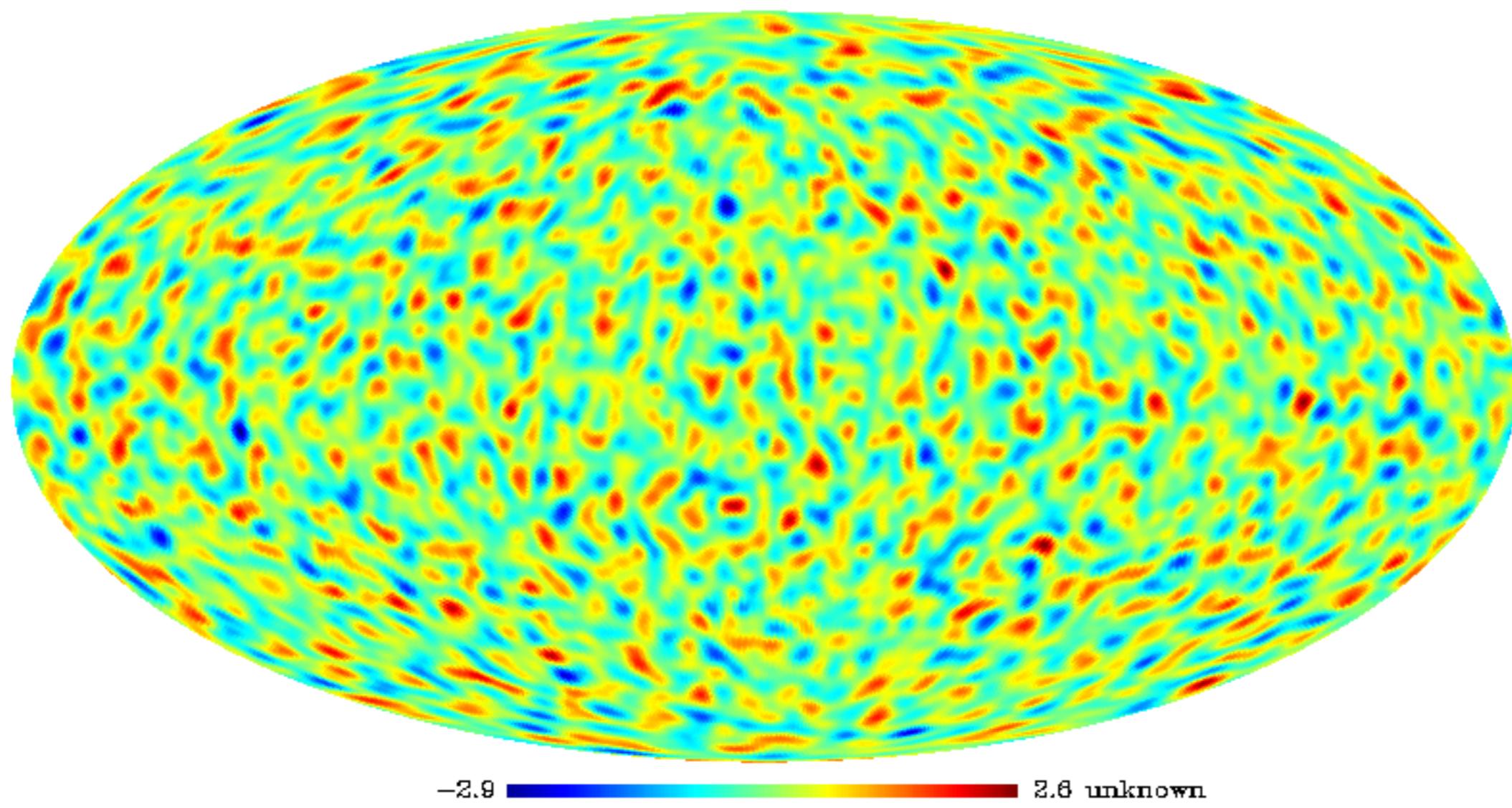
j=4



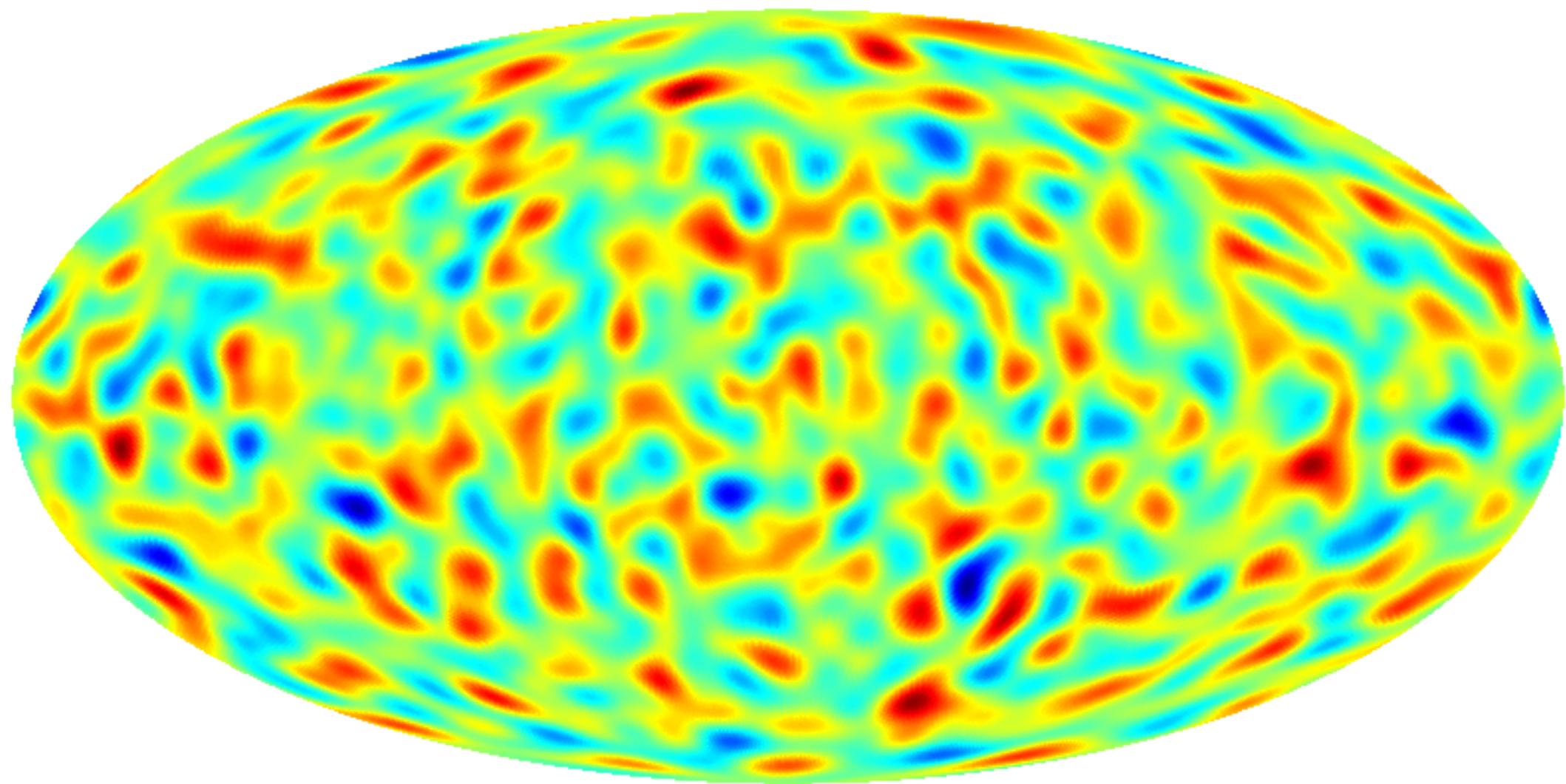
1result_g.fits: TEMPERATURE



2result_g.fits: TEMPERATURE

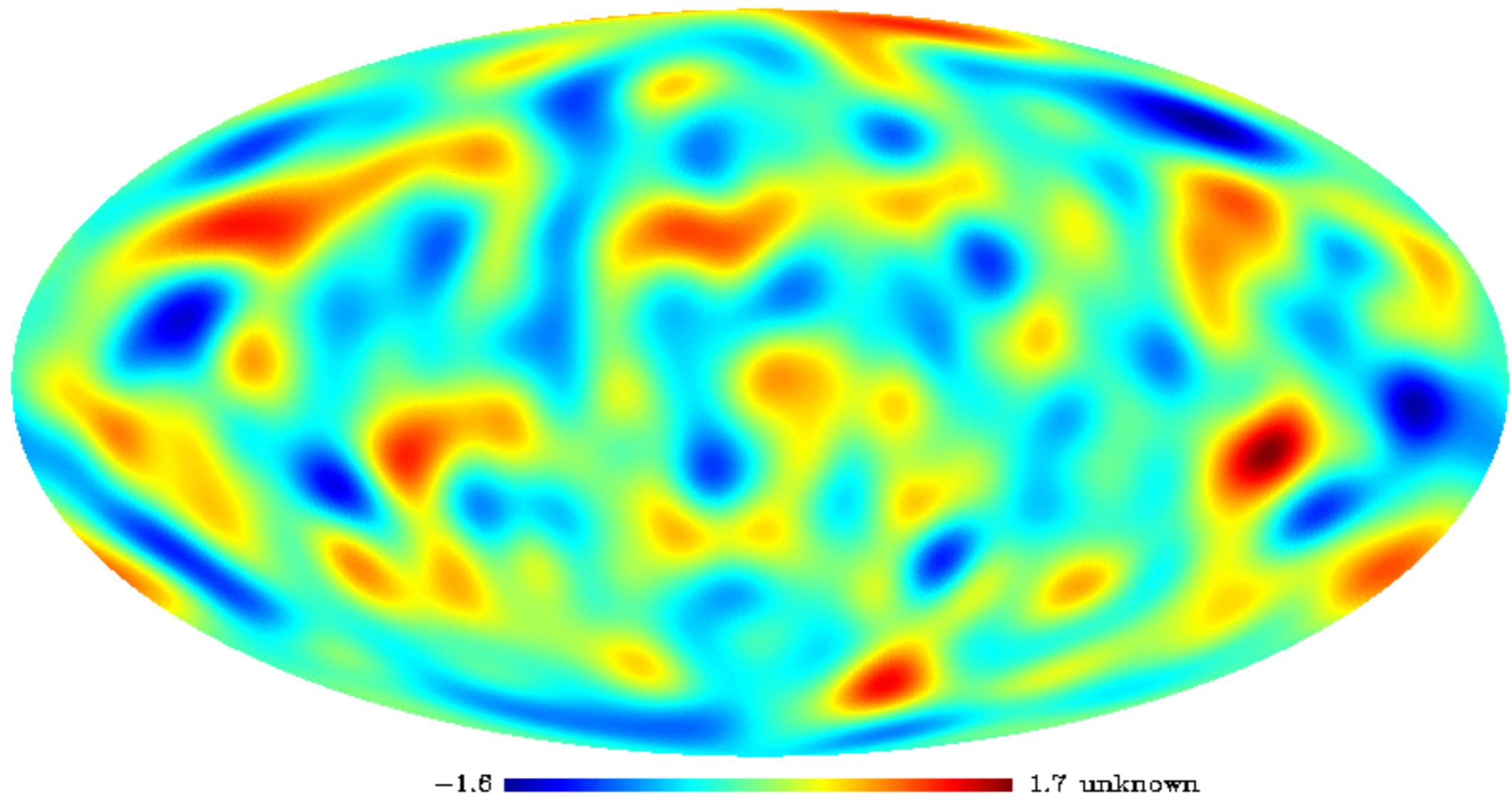


4result_g.fits: TEMPERATURE

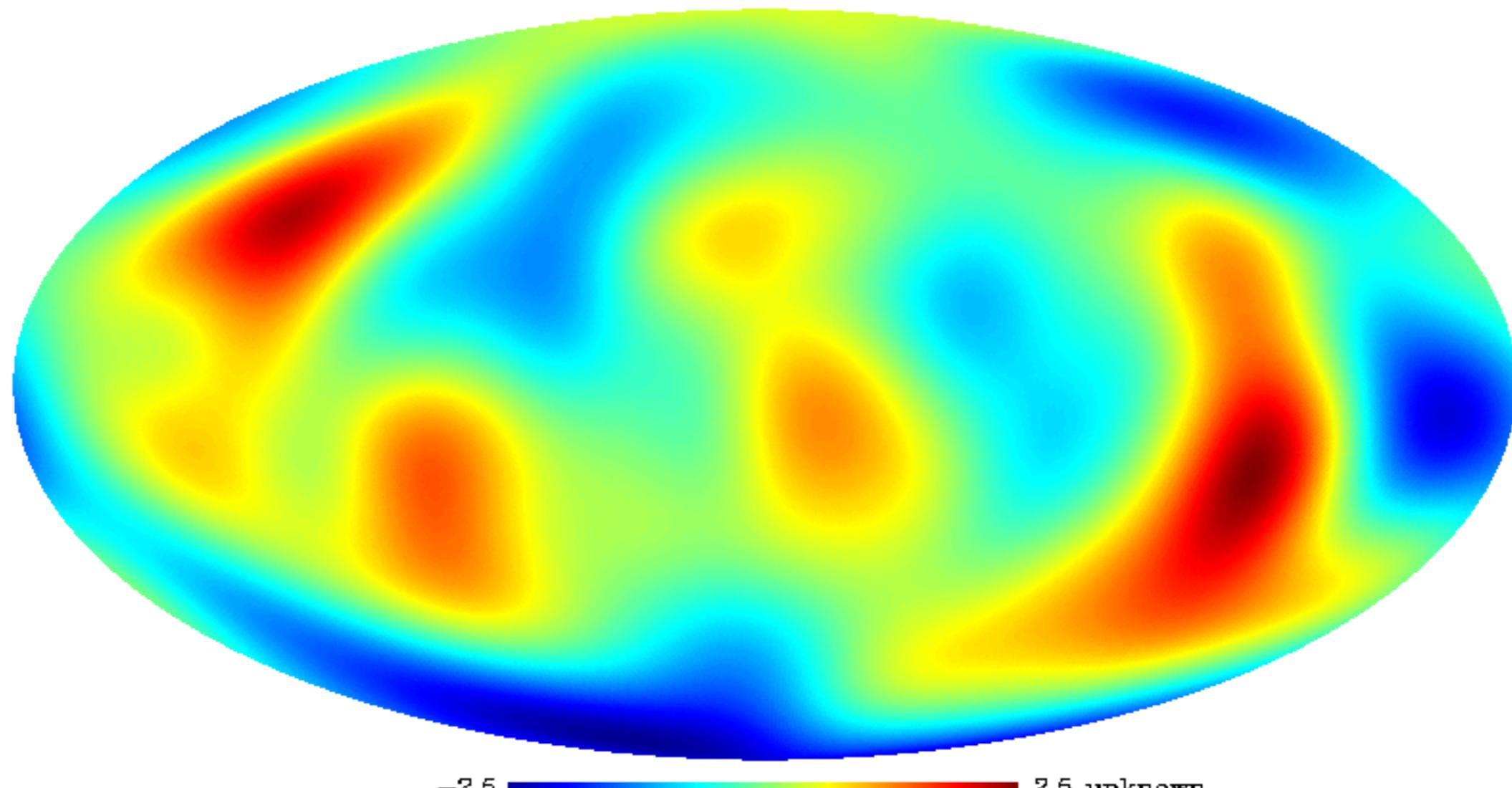


-2.2 ————— 1.8 unknown

8result_g.fits: TEMPERATURE

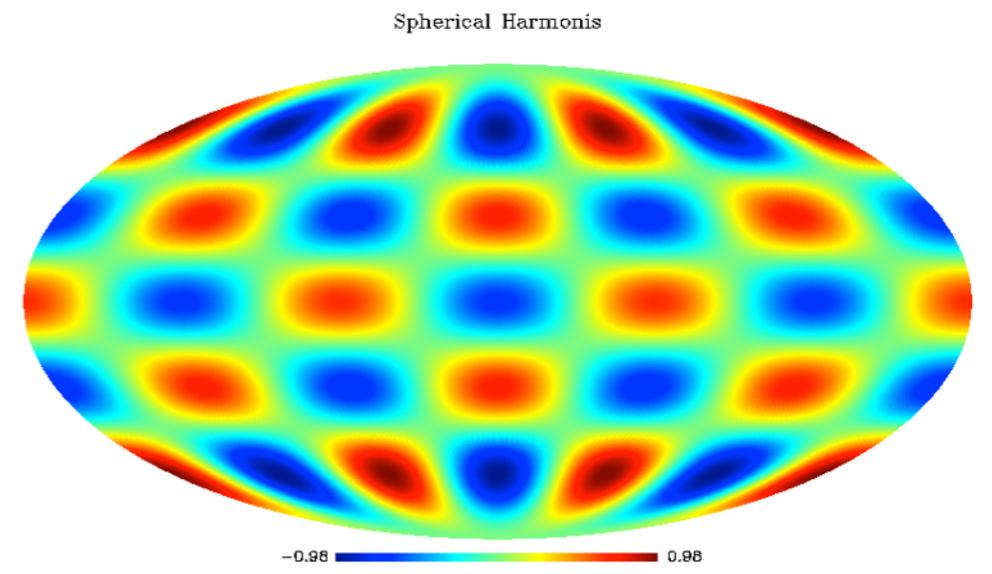


8result_h.fits: TEMPERATURE

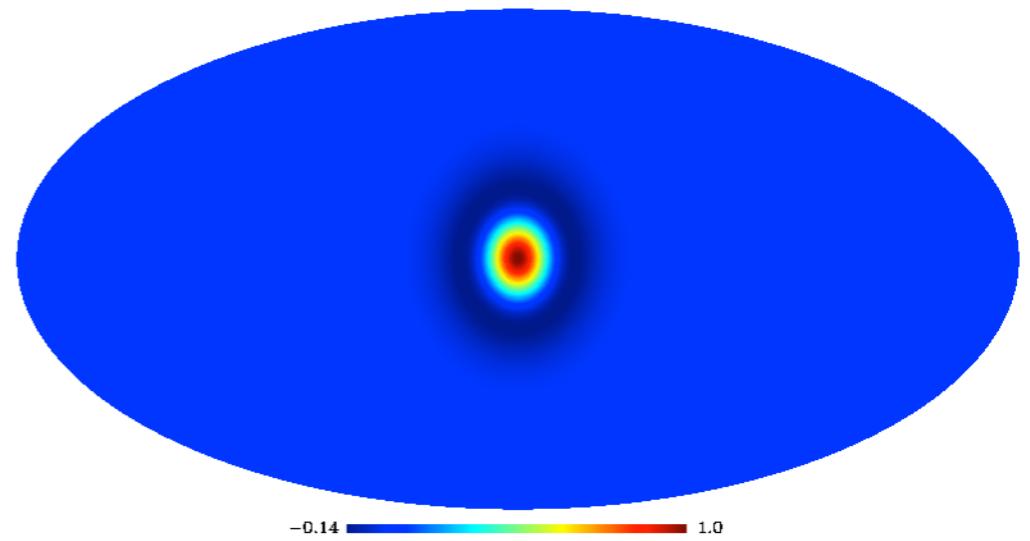


Dictionaries

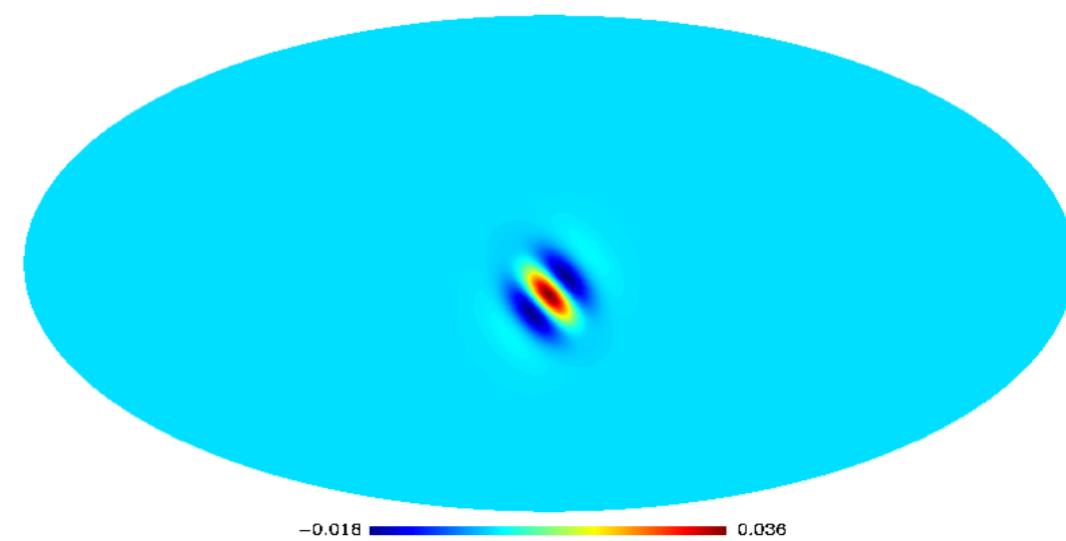
Spherical Harmonics



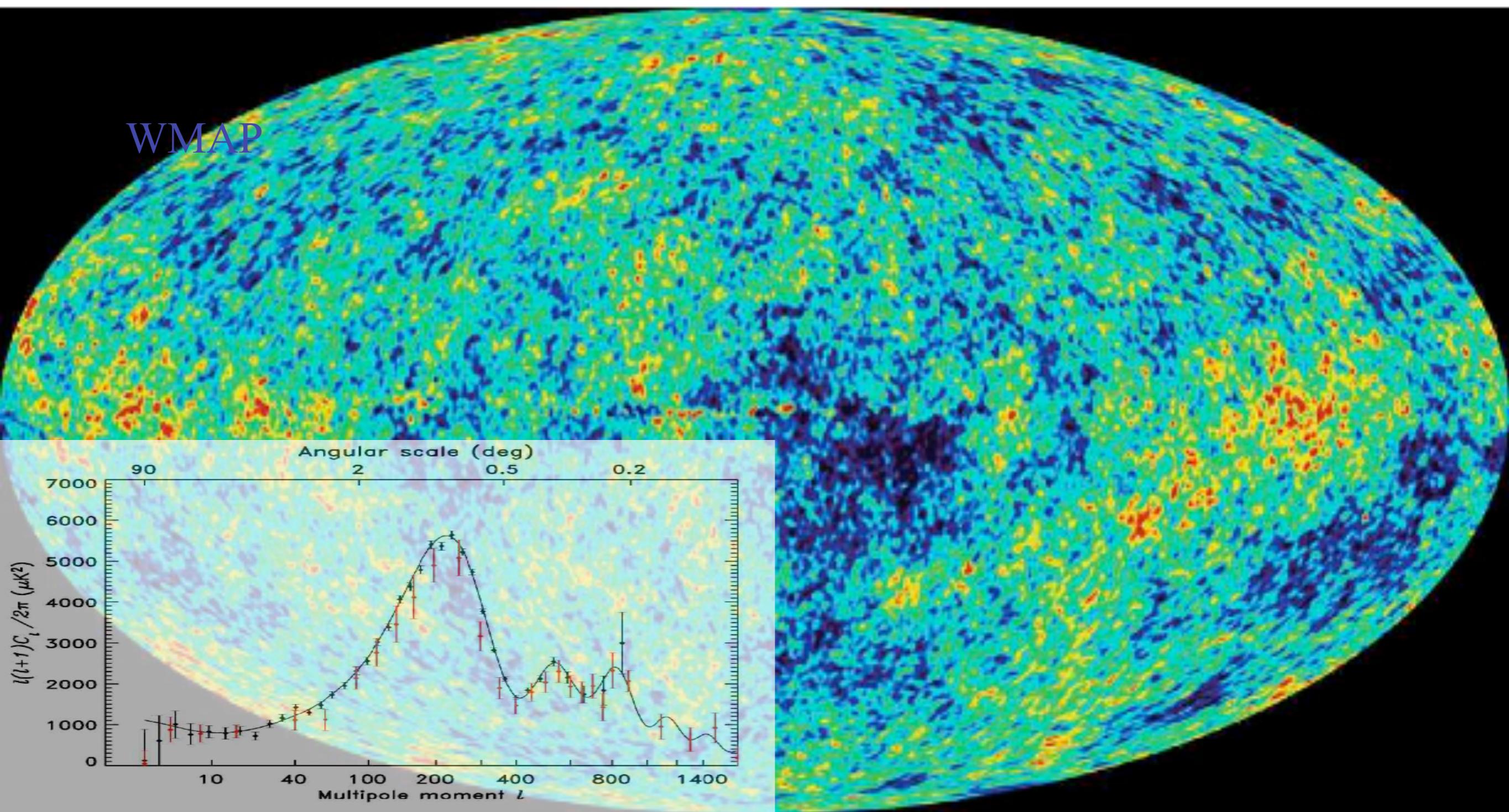
Wavelets



Curvelets

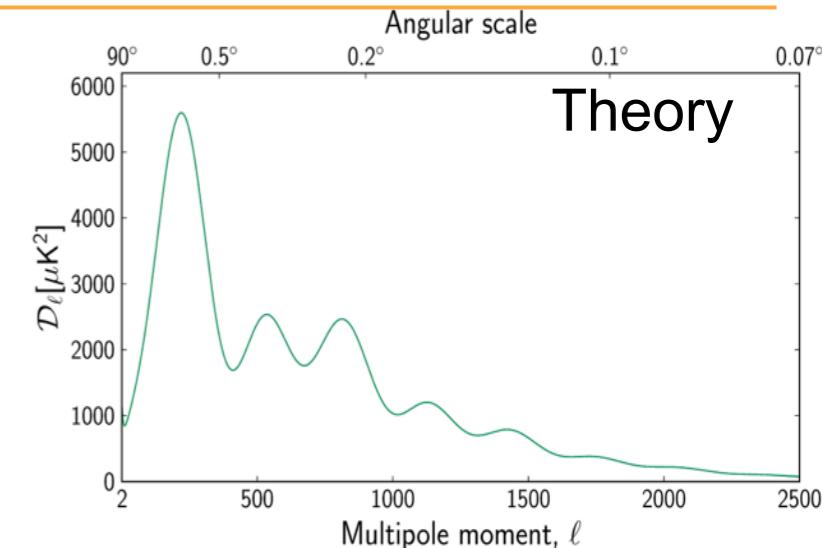
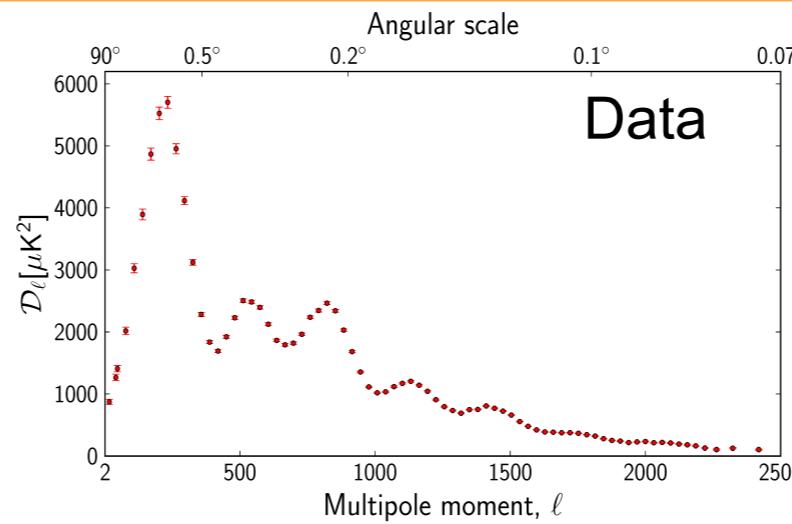
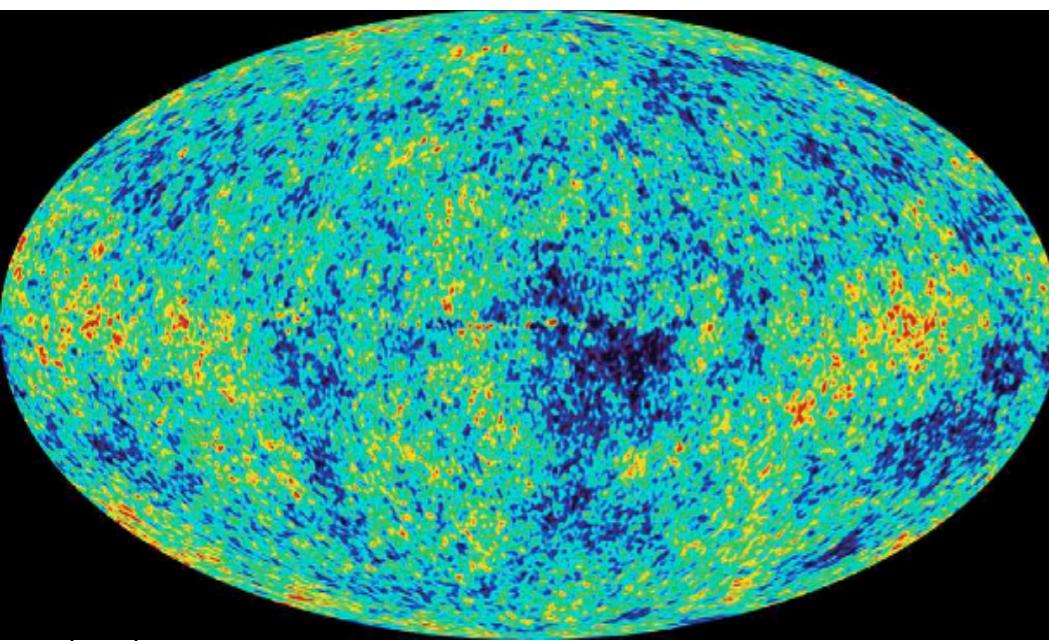


The CMB exhibits Fluctuations

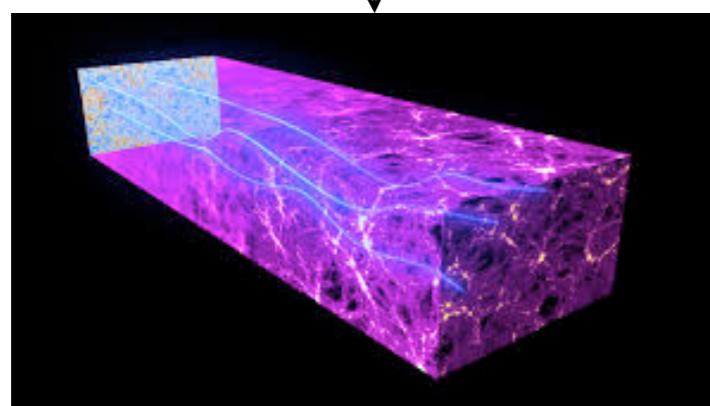


The Cosmic Microwave Background (CMB) is a relic radiation (with a temperature equals to 2.726 Kelvin) emitted 13 billion years ago when the Universe was about 370000 years old.

Statistical Properties of the CMB fluctuation



Cosmological Parameters



Search for specific signatures predicted by inflation models

Statistical analysis of the weak lensing effect

Large scale analysis

Integrated Sachs-Wolfe Effect (ISW)

constraints on inflation mod
(fnl)

gravitational potential
mapping

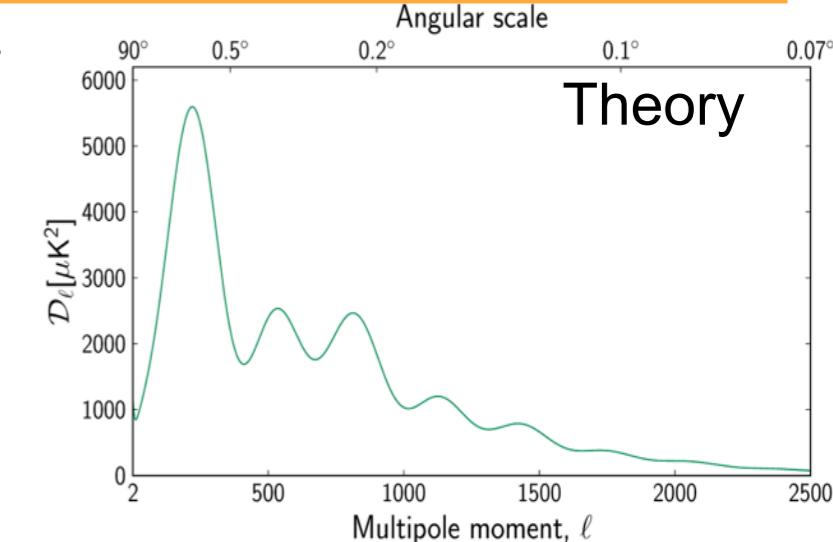
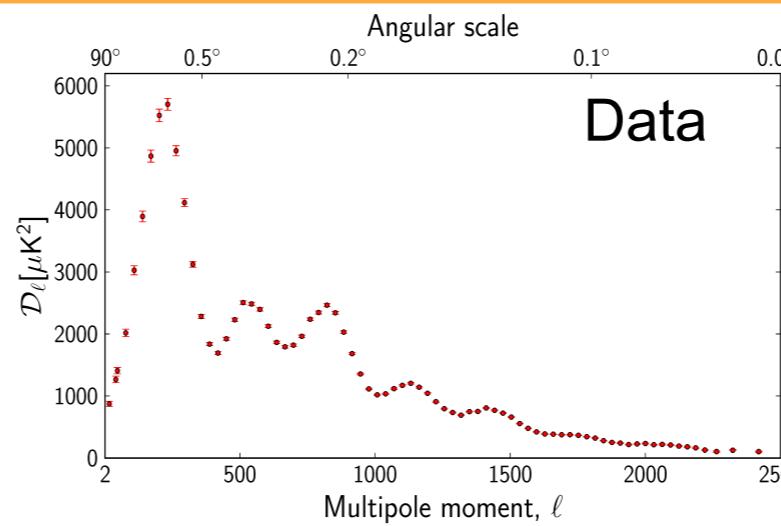
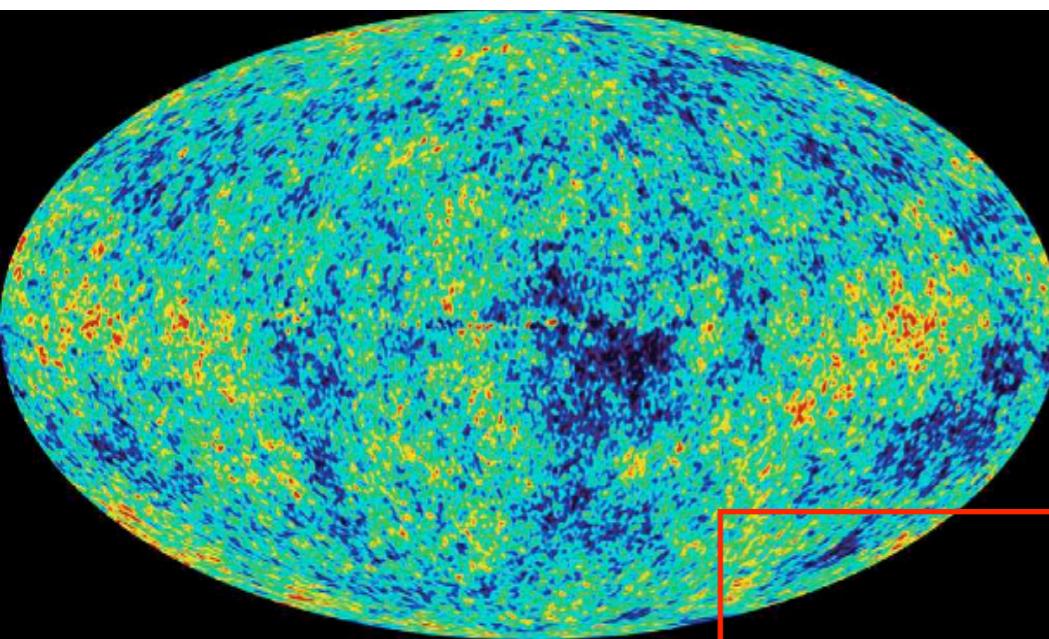
+

power spectrum

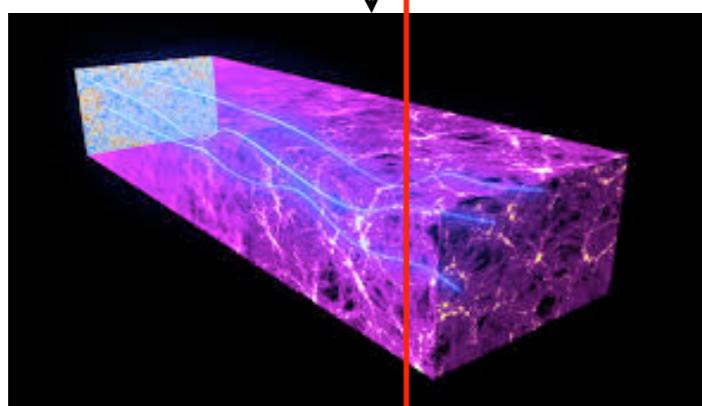
Topology of the univers,
inflation, ISW, etc

Constraint on dark energy

Statistical Properties of the CMB fluctuation



Cosmological Parameters



Search for specific signatures predicted by inflation models

Statistical analysis of the weak lensing effect

Large scale analysis

Integrated Sachs-Wolfe Effect (ISW)

constraints on inflation mod
(fnl)

gravitational potential
mapping

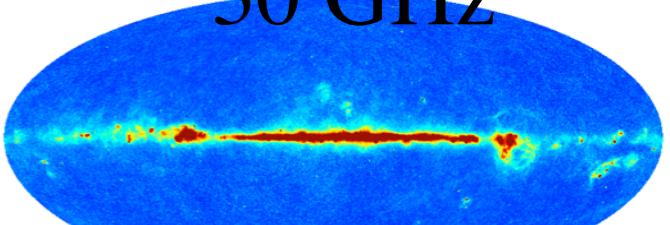
+

power spectrum

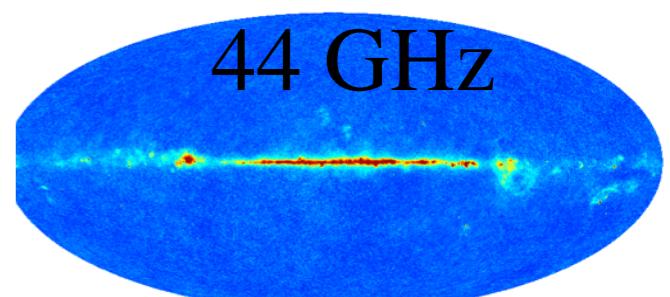
Topology of the univers,
inflation, ISW, etc

Constraint on dark energy

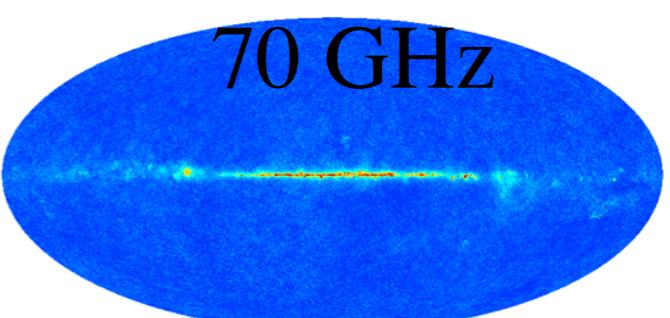
30 GHz



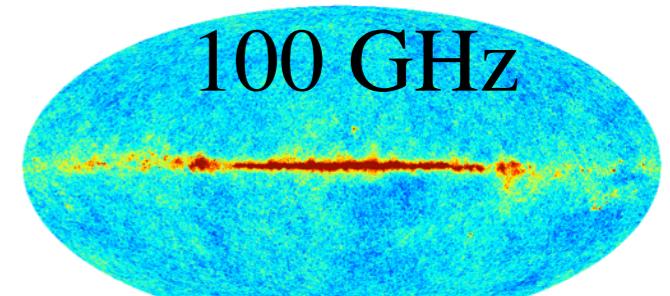
44 GHz



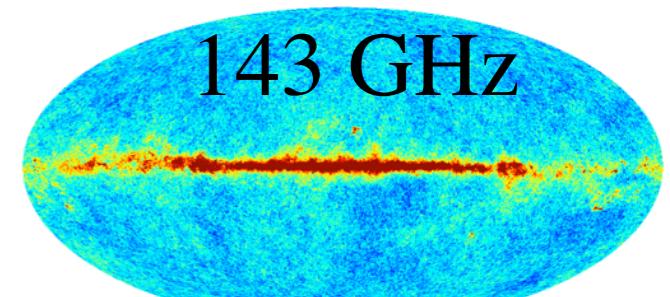
70 GHz



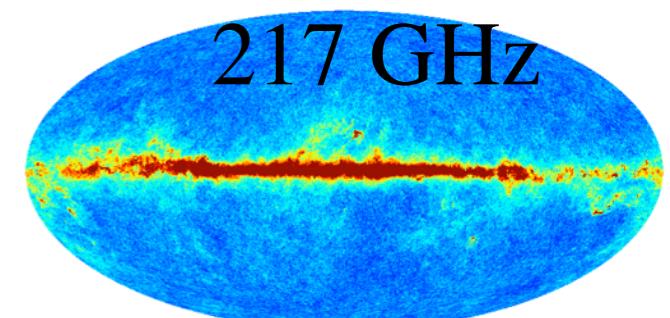
100 GHz



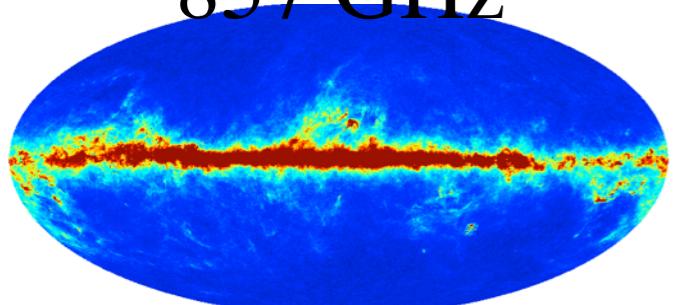
143 GHz



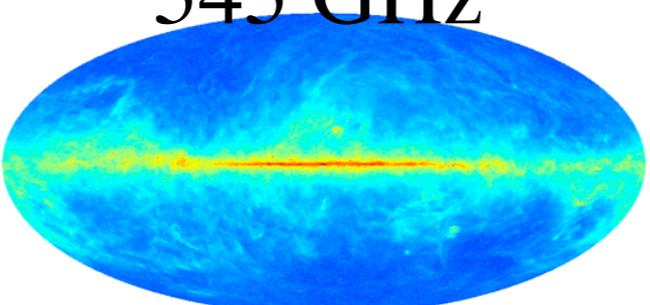
217 GHz



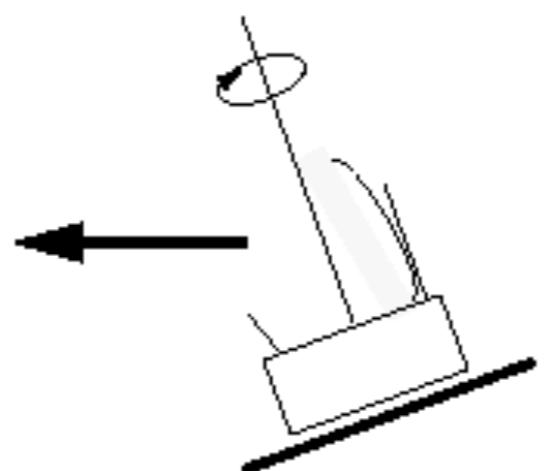
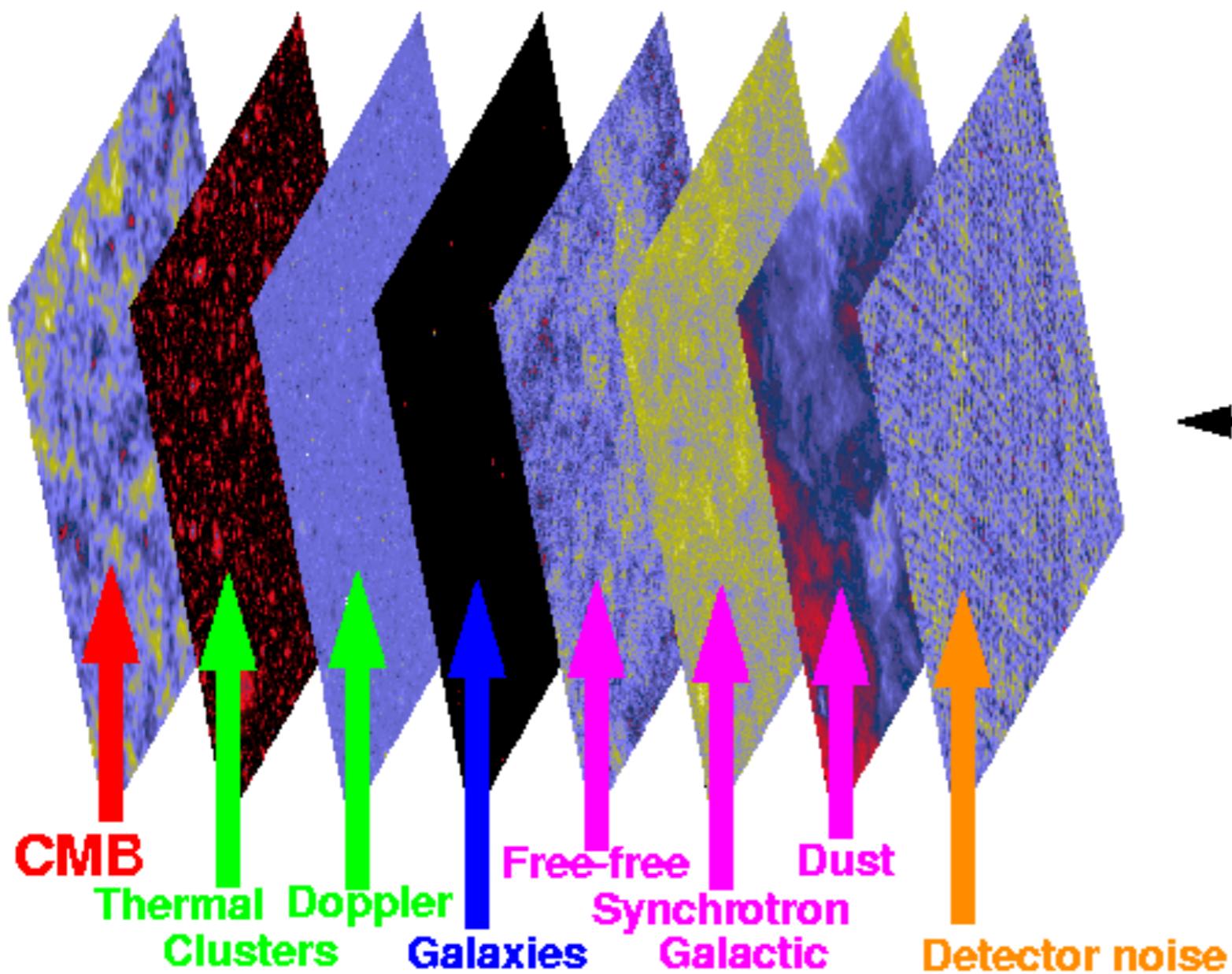
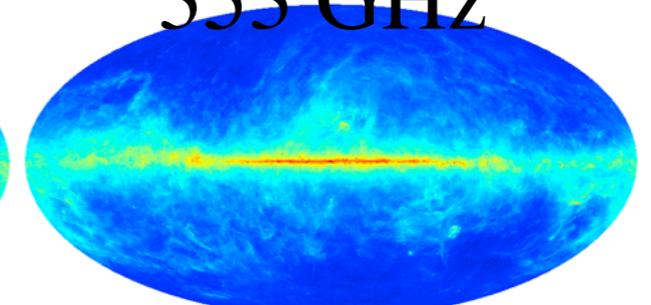
857 GHz

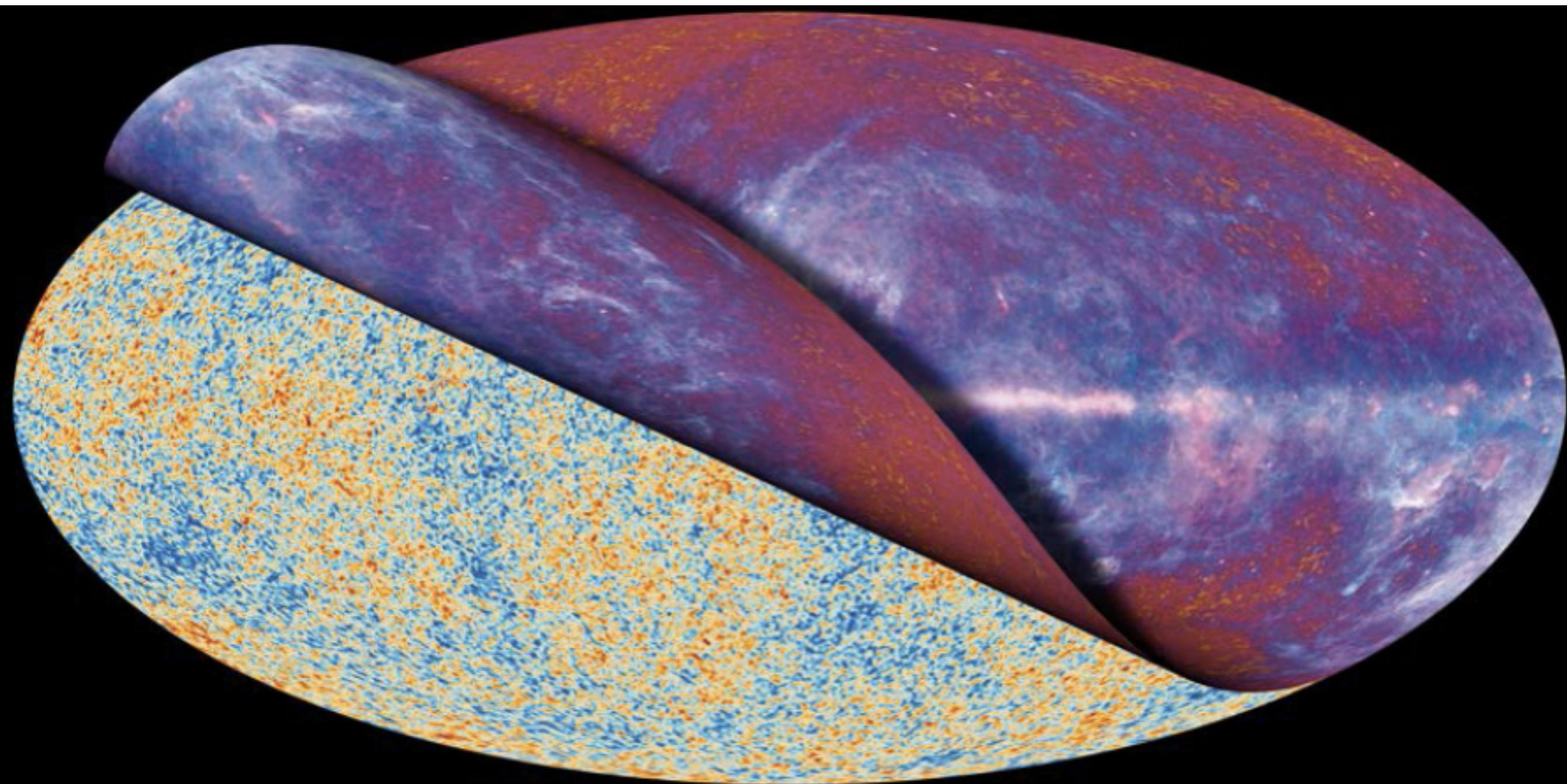


545 GHz

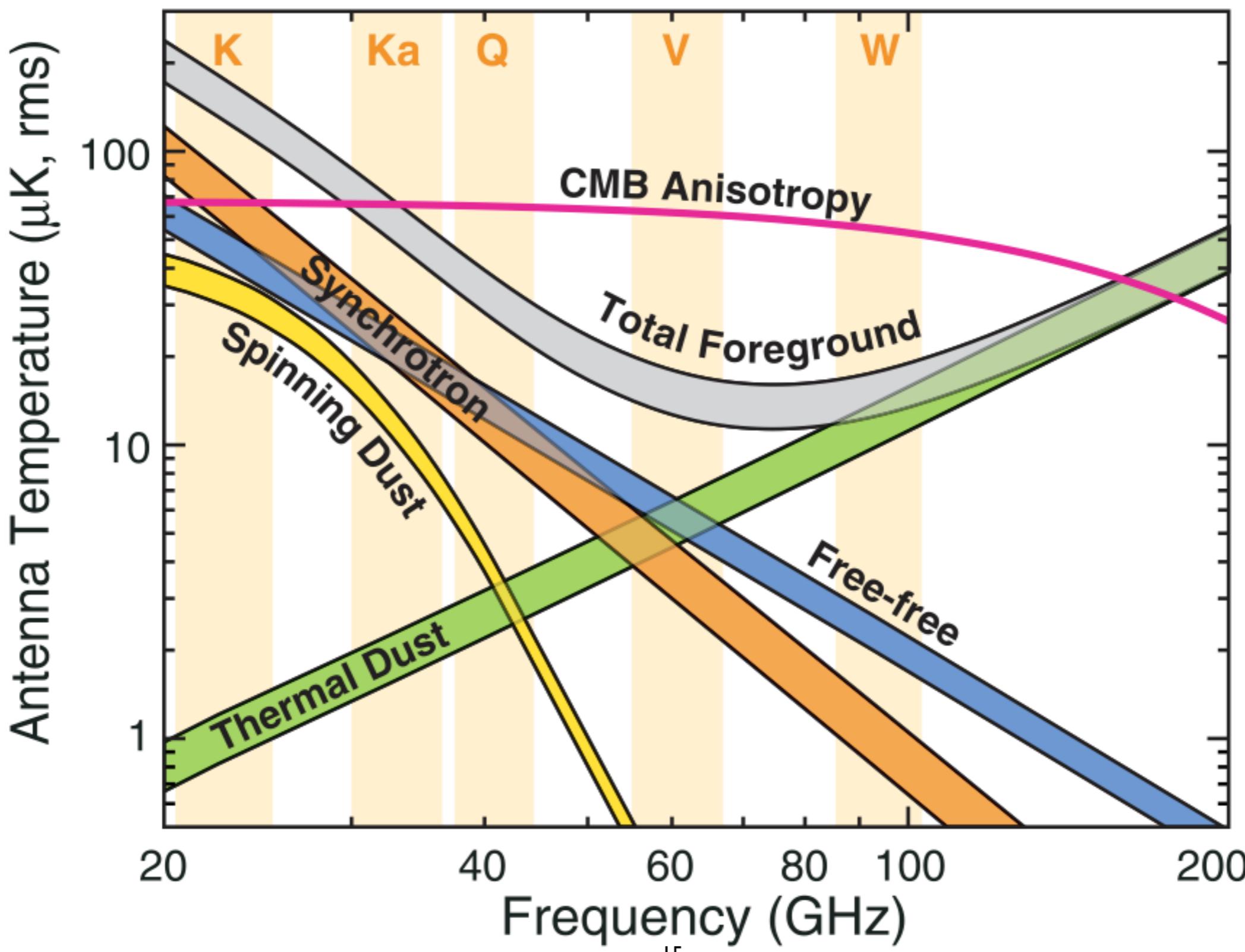


353 GHz





Component Separation: more problems



At a first order, we can model each channel map y_i is a linear combination of several sky components x_j .

For each channel map:

$$y_i = \sum_j a_{ij}x_j + n_i$$

Estimating the CMB is a Blind Source Separation (BSS) problem

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

Need to add constraint

$$\min_{A,X} = \| Y - AX \|^2 \quad s.t. \quad \mathcal{C}(X, A)$$

1) The beam:

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

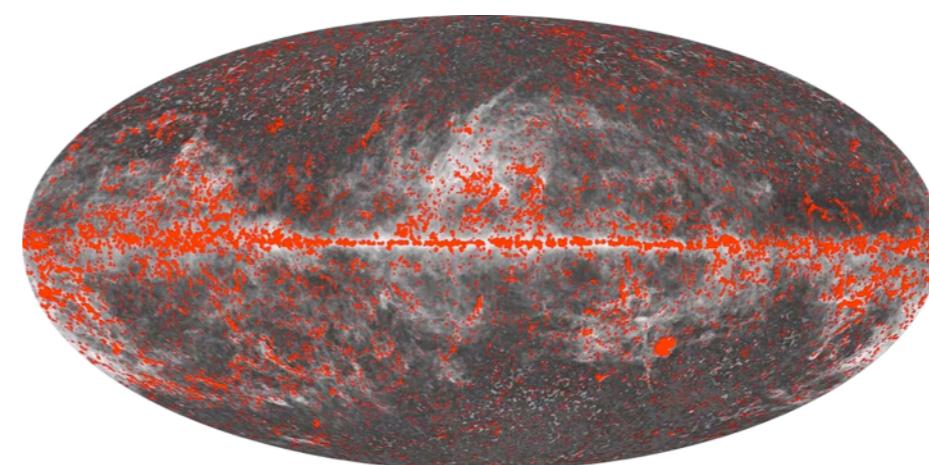
Globally: $\mathbf{Y} = \mathcal{H}(\mathbf{AX}) + \mathbf{N}$ **\mathcal{H} is singular !**

where \mathcal{H} is the multichannel convolution operator

2) Spectral behavior **varies spatially** for some components (dust, synchroton).

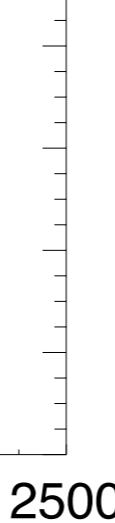
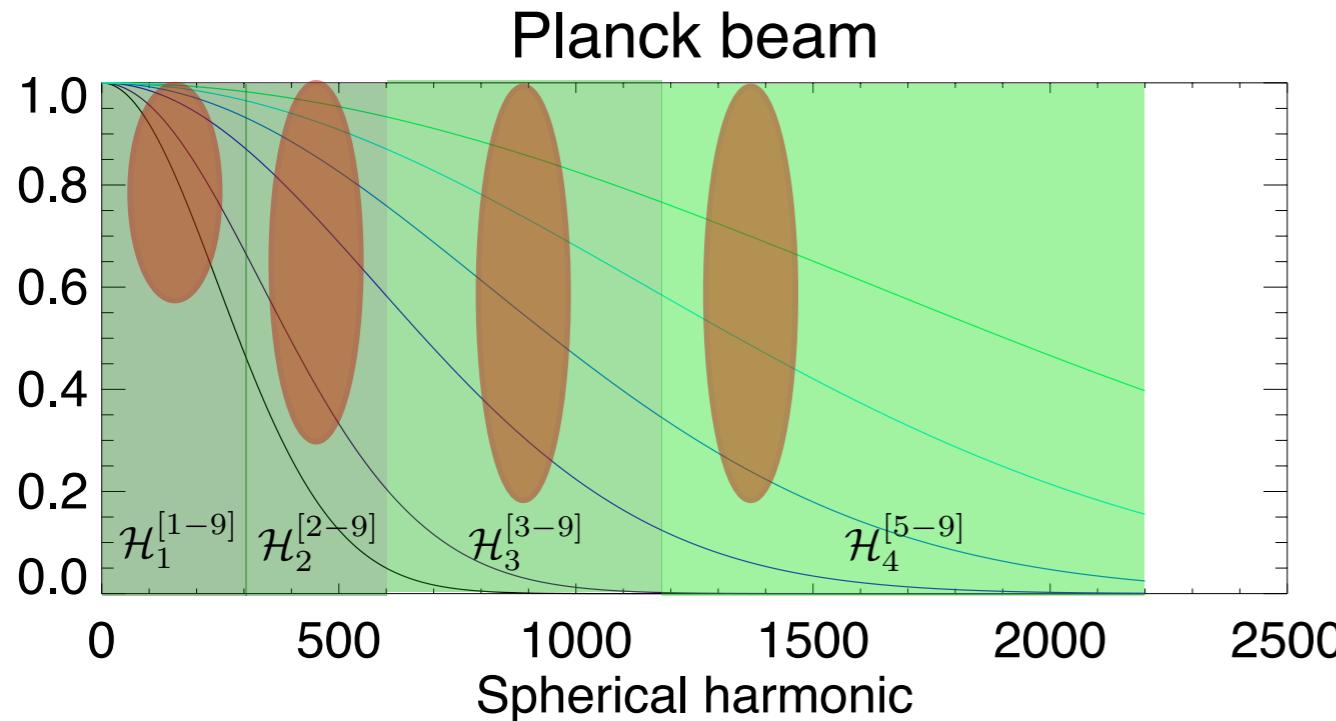
$$\mathbf{Y}[k] = \mathcal{H}(\mathbf{A}_k \mathbf{X})[k] + \mathbf{N}[k]$$

3) Point sources:

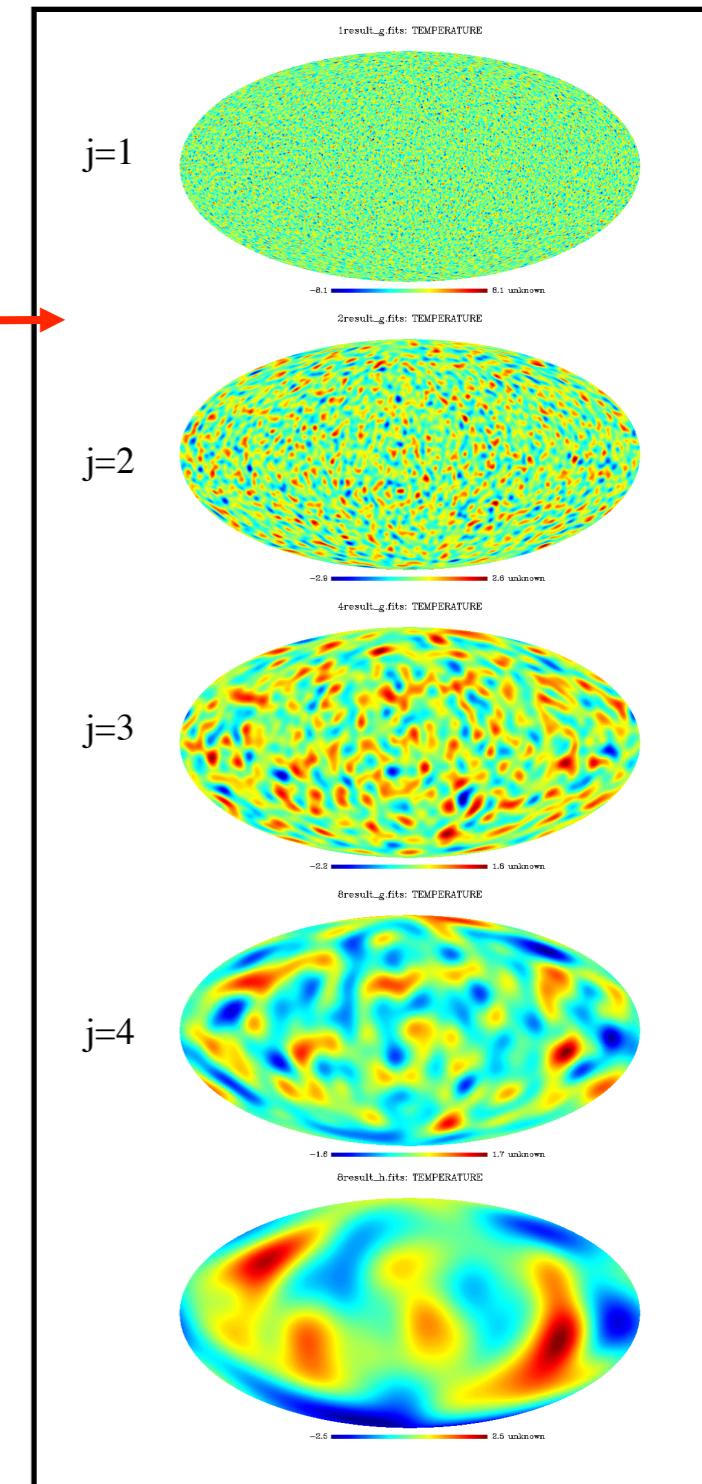


Component Separation

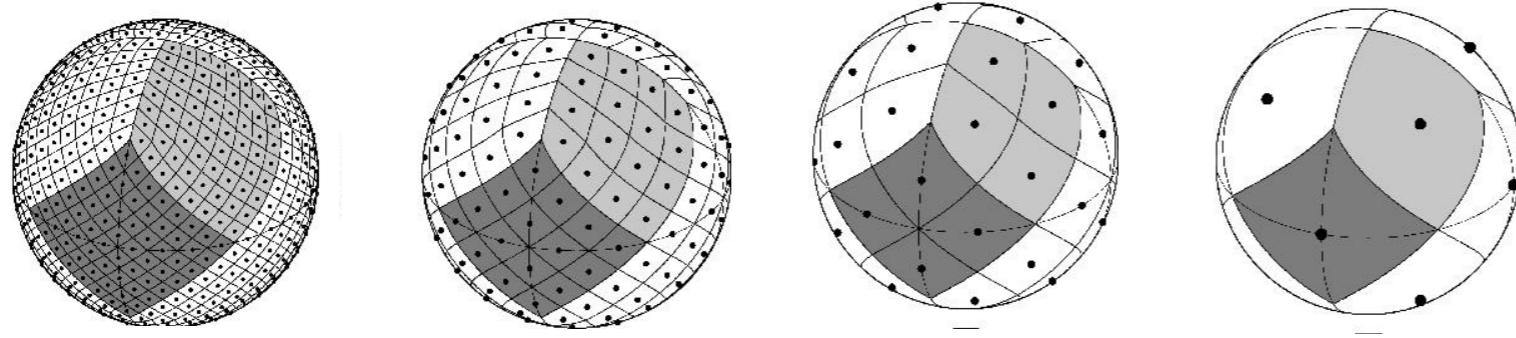
=> Use Wavelets to work at different resolutions:



Undecimated
Wavelet Transform



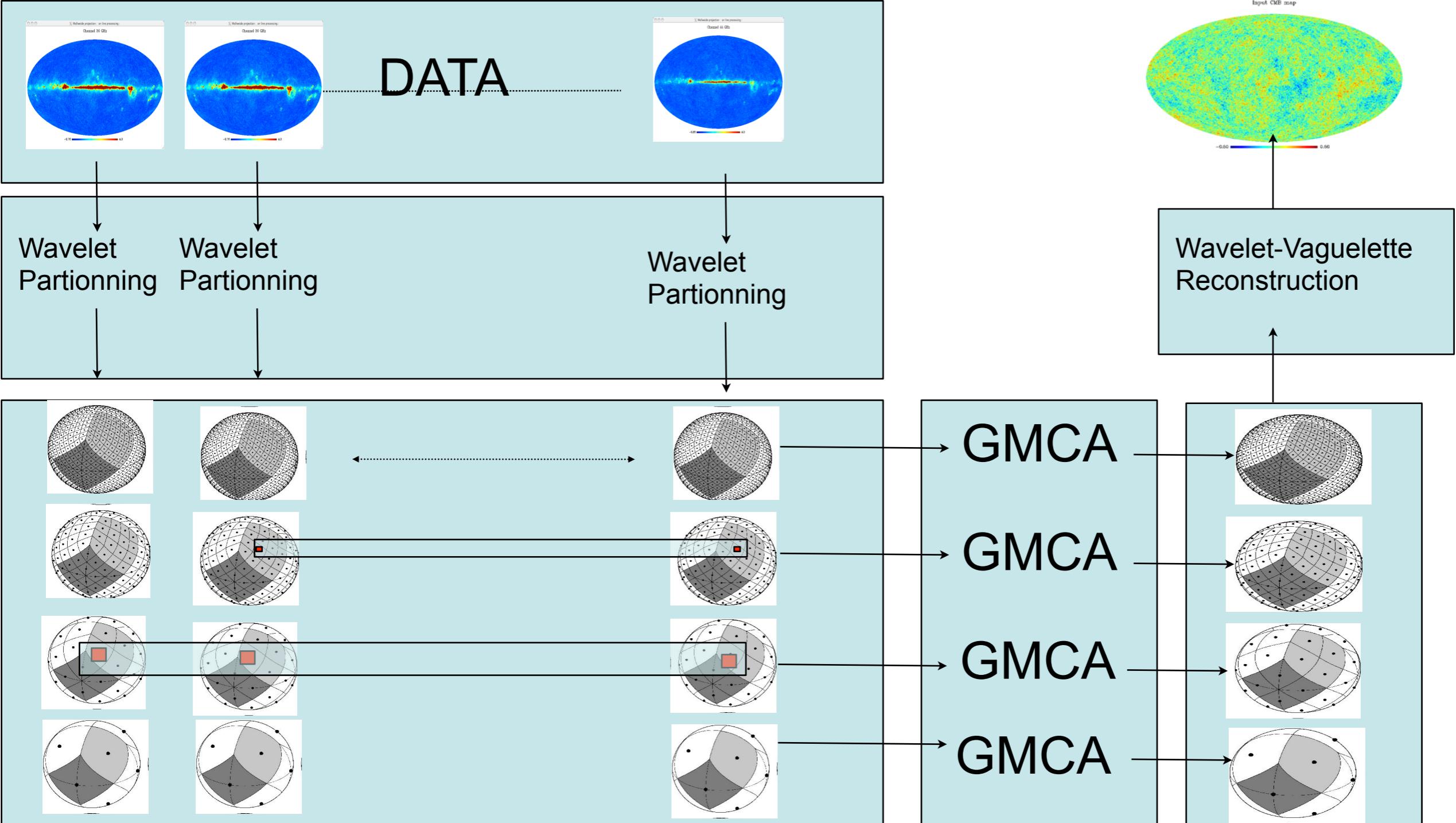
=> Assume the mixing matrix varies smoothly
Partitionning of the Wavelet Scales



Wavelet-Vaguelette GMCA Decomposition

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

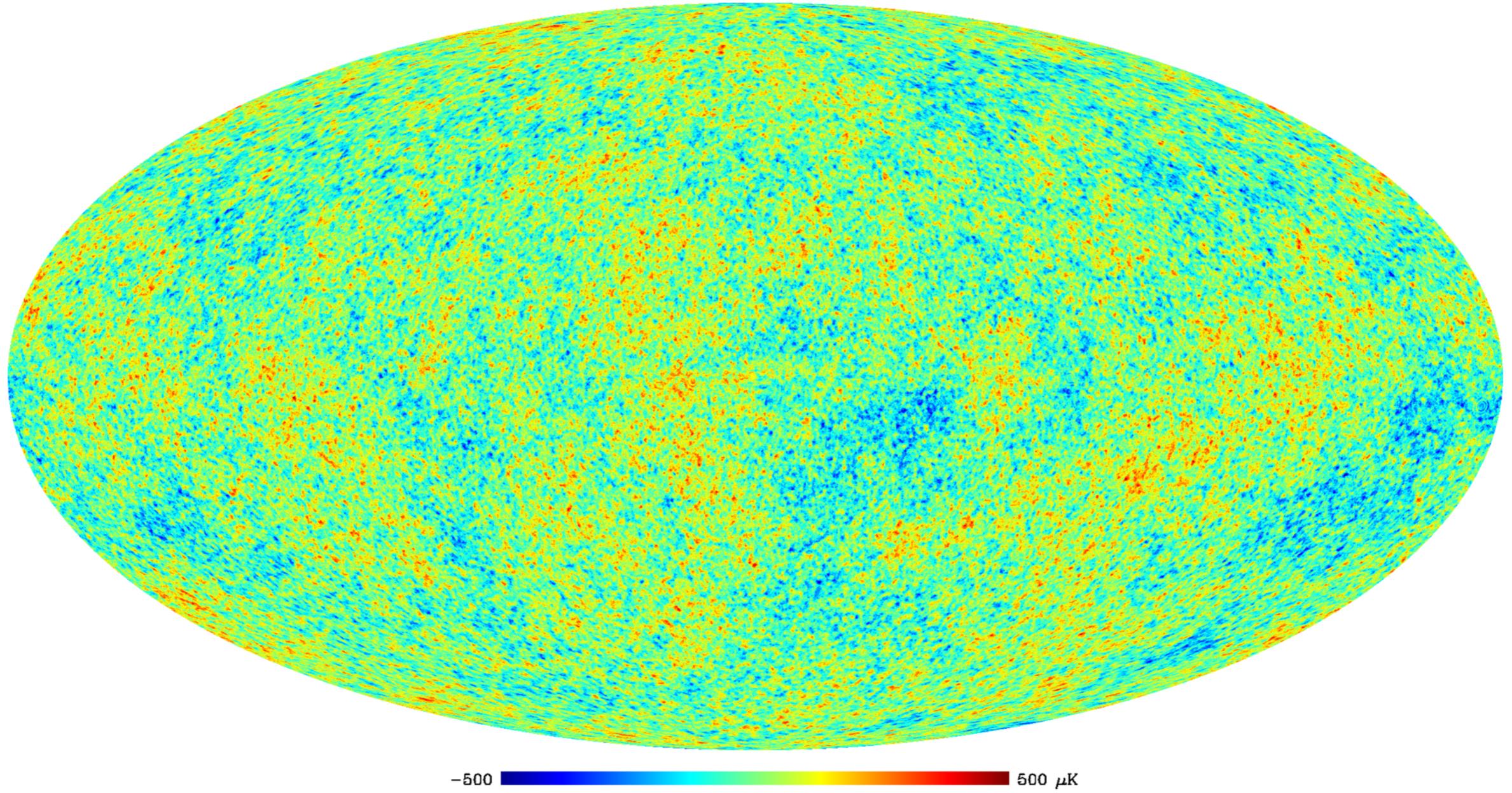
$$\tilde{f} = \sum_j \sum_k \Delta(\langle y, \Psi_{j,k} \rangle) \psi_j, k$$



Full Sky Sparse WMAP + Planck-PR2 Map



CMB map LGMCA_WPR2 at 5 arcmin



Bobin J., Sureau F., Starck J-L, Rassat A. and Paykari P., Joint Planck and WMAP CMB map reconstruction, *A&A*, 563, 2014

Bobin J., Sureau F., Starck, CMB reconstruction from the WMAP and Planck PR2 data, in press, *A&A*, 2016. [arXiv:1511.08690](https://arxiv.org/abs/1511.08690)

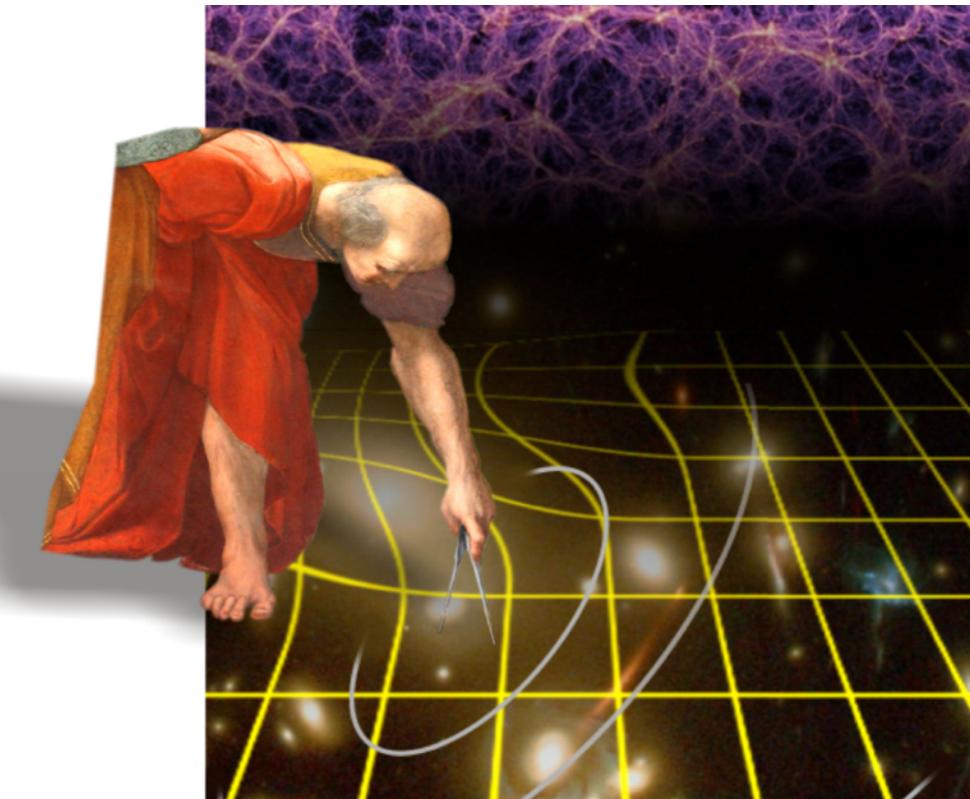
Understand the origin of the Universe's accelerating expansion:

→ probe the properties and nature of *dark energy, dark matter, gravity* and distinguish their effects **decisively**

→ by tracking their observational signatures on the

- geometry of the universe:

Weak Lensing + Galaxy Clustering



- cosmic history of structure formation: WL, z-space distortion, clusters of galaxies

→ **Controlling systematic residuals to an unprecedented level of accuracy, that cannot be reached by any other competing missions/telescopes**

Gains in space:

Stable data: homogeneous data set over the whole sky

→ **Systematics** are small, understood and controlled

→ Homogeneity : Selection function perfectly controlled

Euclid

Mapping the Geometry of the Dark Universe

Euclid, ESA Cosmic Vision: launch in 2020:

- 1274 members
- 125 laboratories
- 15 countries

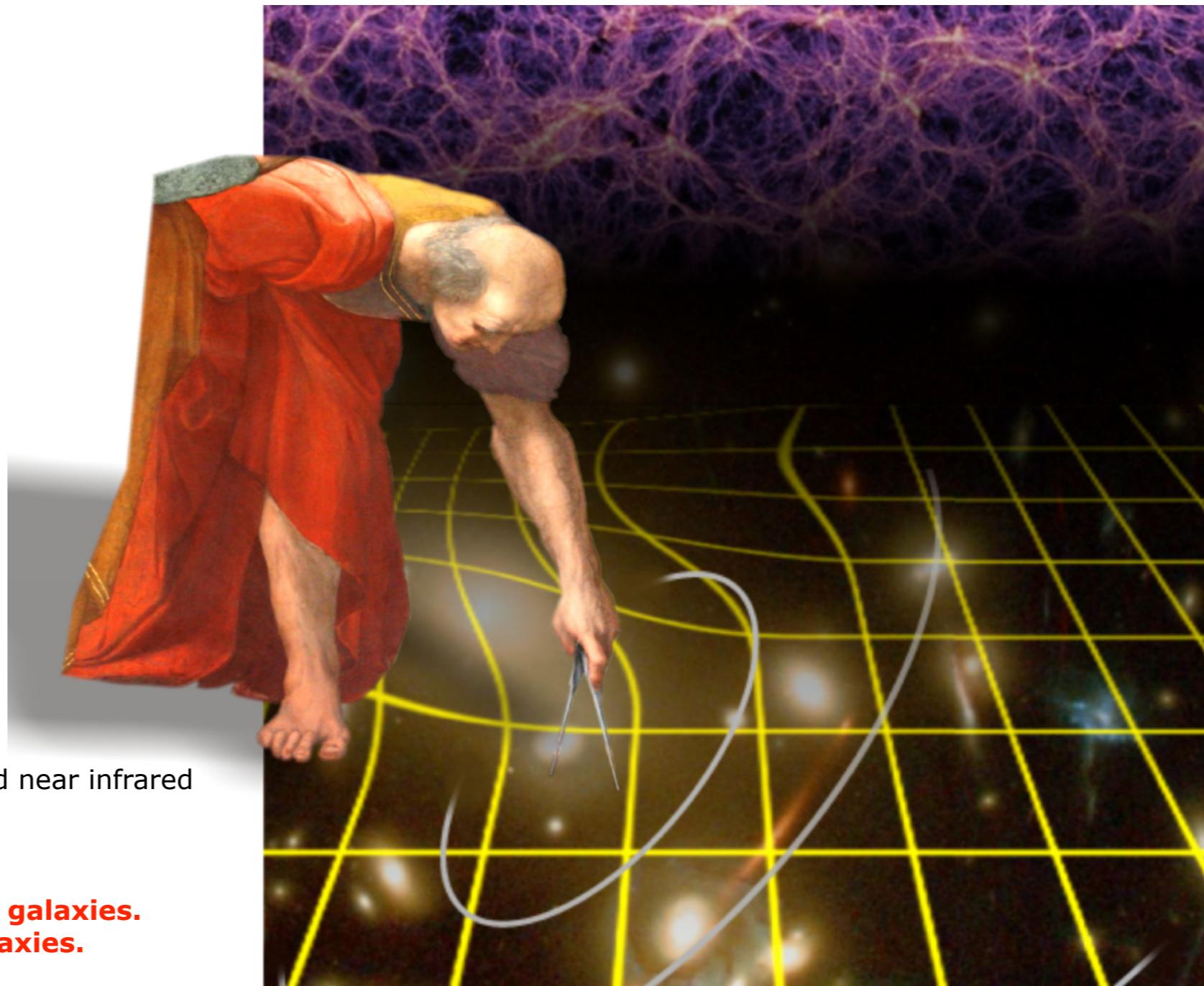
Euclid is the result from the fusion in 2008 of two missions:

- DUNE
- SPACE

Ambitious scientific goals:

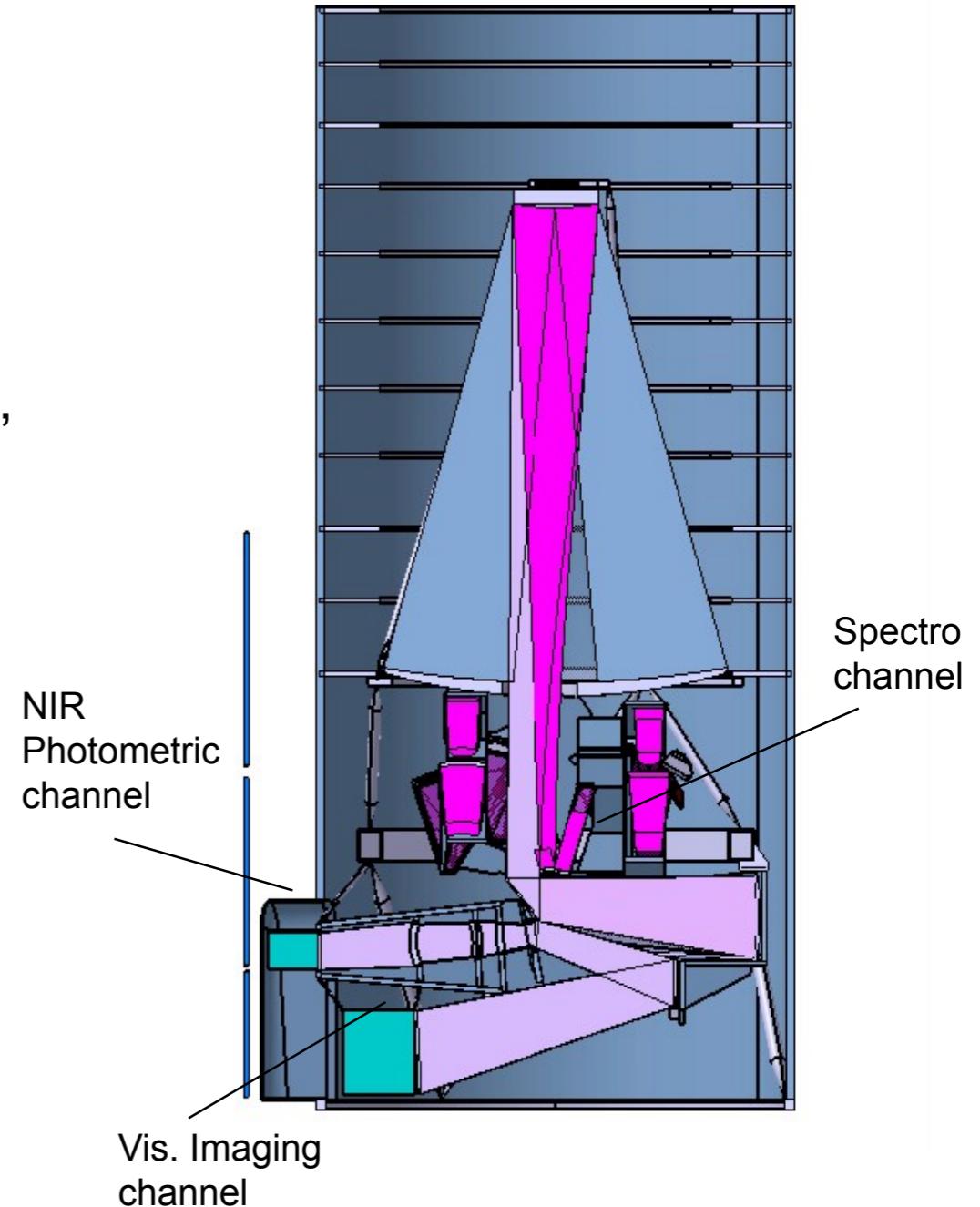
- Observe **15 000 deg²** during 6 years in optical and near infrared wavelength. (forme) dans le visible ET
- Telescope: 1.2m
- 4 bands
- Photometric redshift (distance) for **1 000 000 000 galaxies.**
- Spectroscopic IR measurement of **50 000 000 galaxies.**

==> 850 Gbits of data per day.

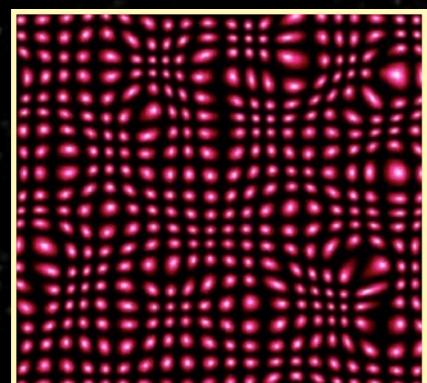


Euclid mission element

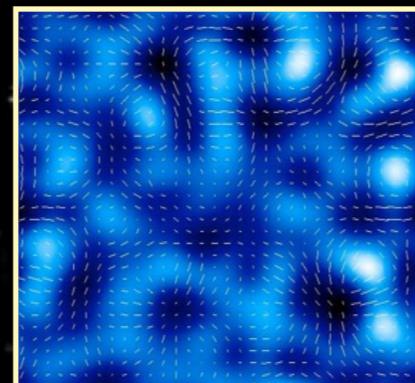
- Launch Soyuz, in 2020, L2 Orbit
- 6 years mission
- Telescope: 1.2 m
- Instruments:
- **VIS**: Visible imaging channel:
 - 0.54 deg², 0.10" pixels, 0.16" PSF FWHM,
 - 1 broad band R+I+Z (0.55-0.92μm),
 - 36 CCD detectors, **galaxy shapes**
- **NISP**: NIR photometry channel:
 - 0.54 deg², 0.3" pixels,
 - 3 bands Y,J,H (1.0-1.7μm),
 - 16 HgCdTe detectors, **photo-z's**
- **NISP**: NIR Spectroscopic channel:
 - 0.54 deg²,
 - R(mean)=250,
 - 0.9-1.7μm, slitless, **spectro redshifts**



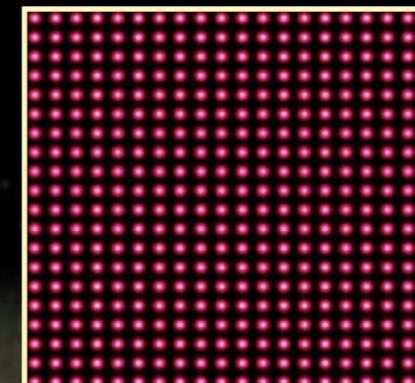
Weak Lensing



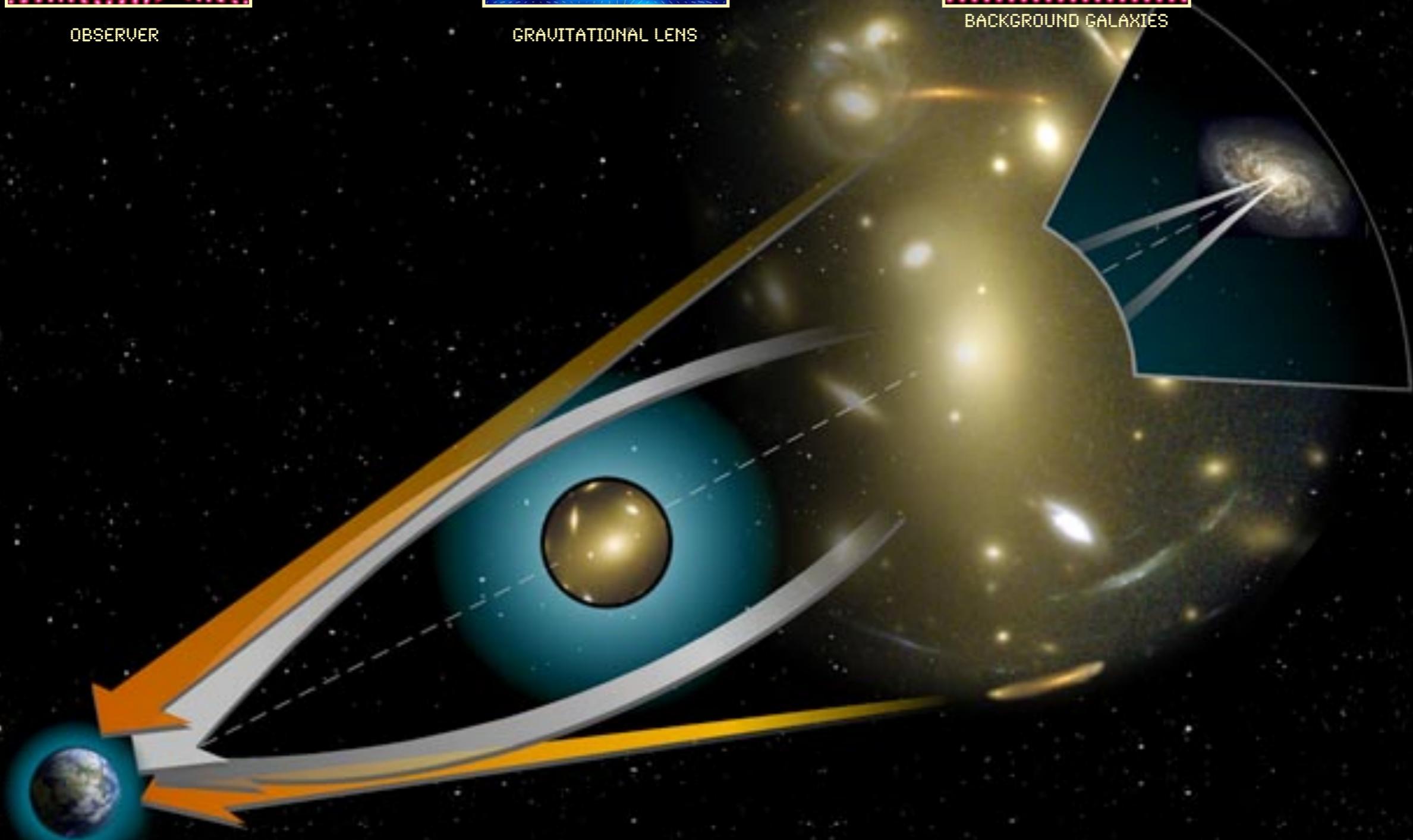
OBSERVER



GRAVITATIONAL LENS



BACKGROUND GALAXIES



Detection + Classification stars/galaxies



Galaxies



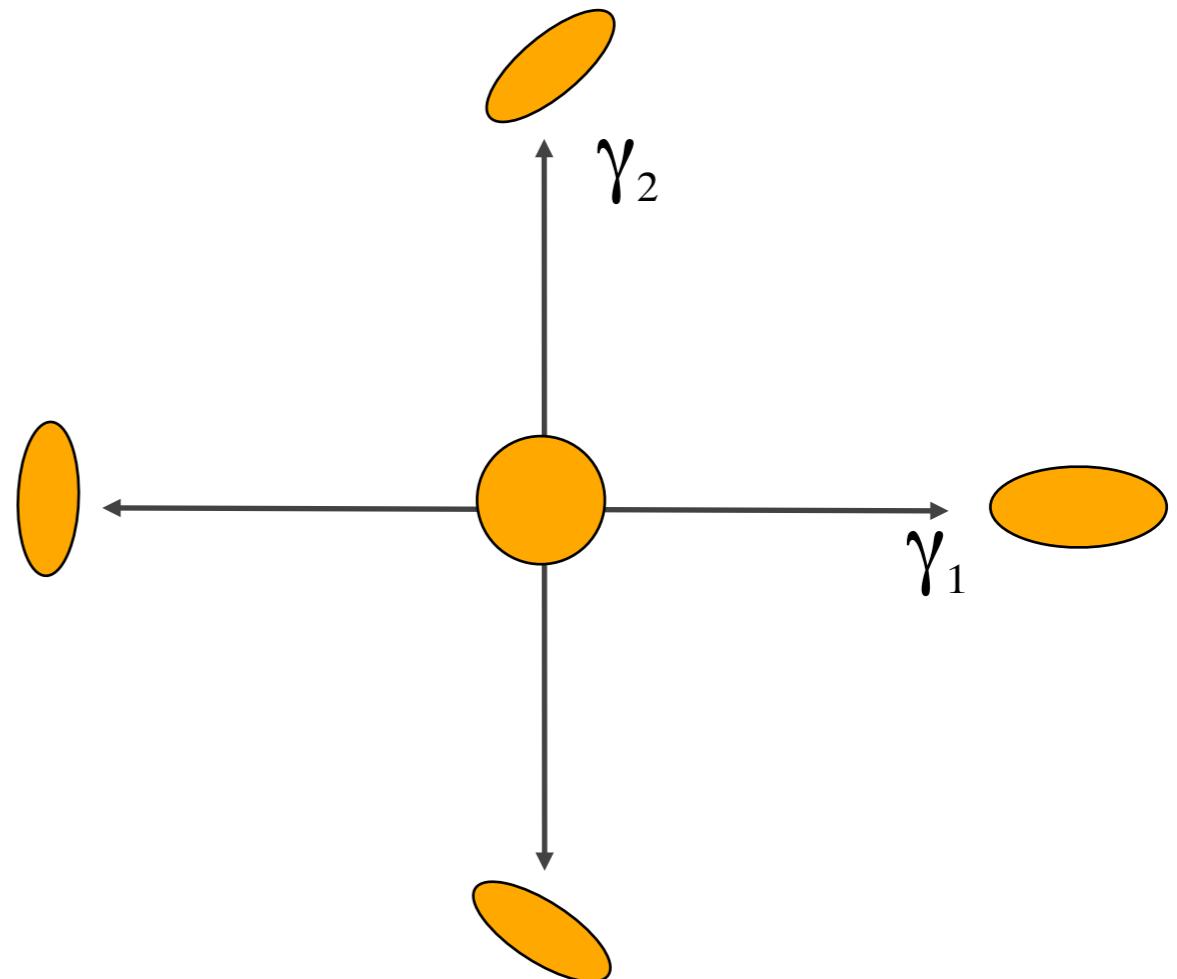
Stars



Shape Parameters

γ_1 = deformation along the x-axis,
and γ_2 at 45 degrees from it.

$$\gamma = \gamma_1 + i\gamma_2 = |\gamma|e^{2i\theta}$$

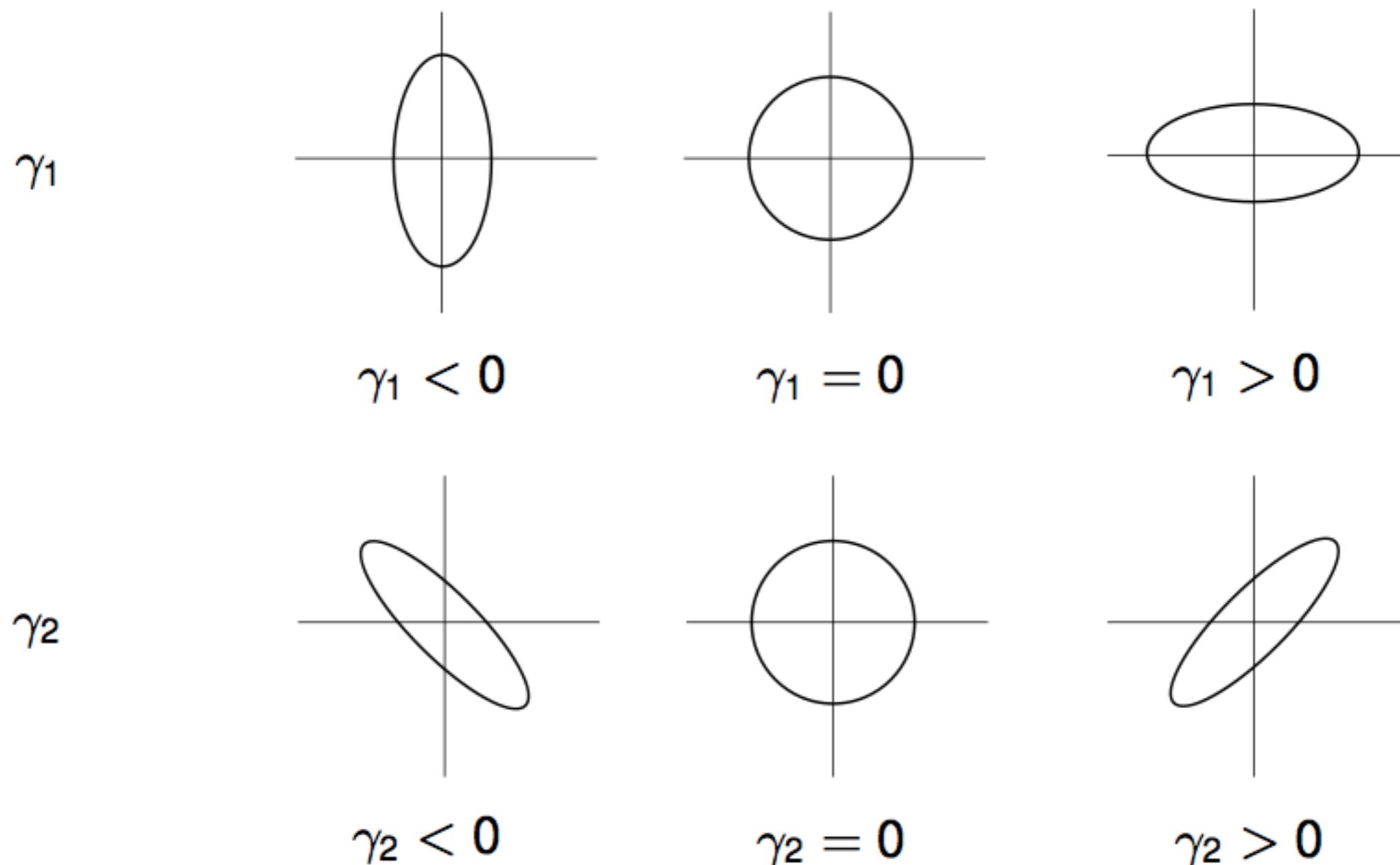


Where the modulus represents the amount of shear and the phase represents its direction.

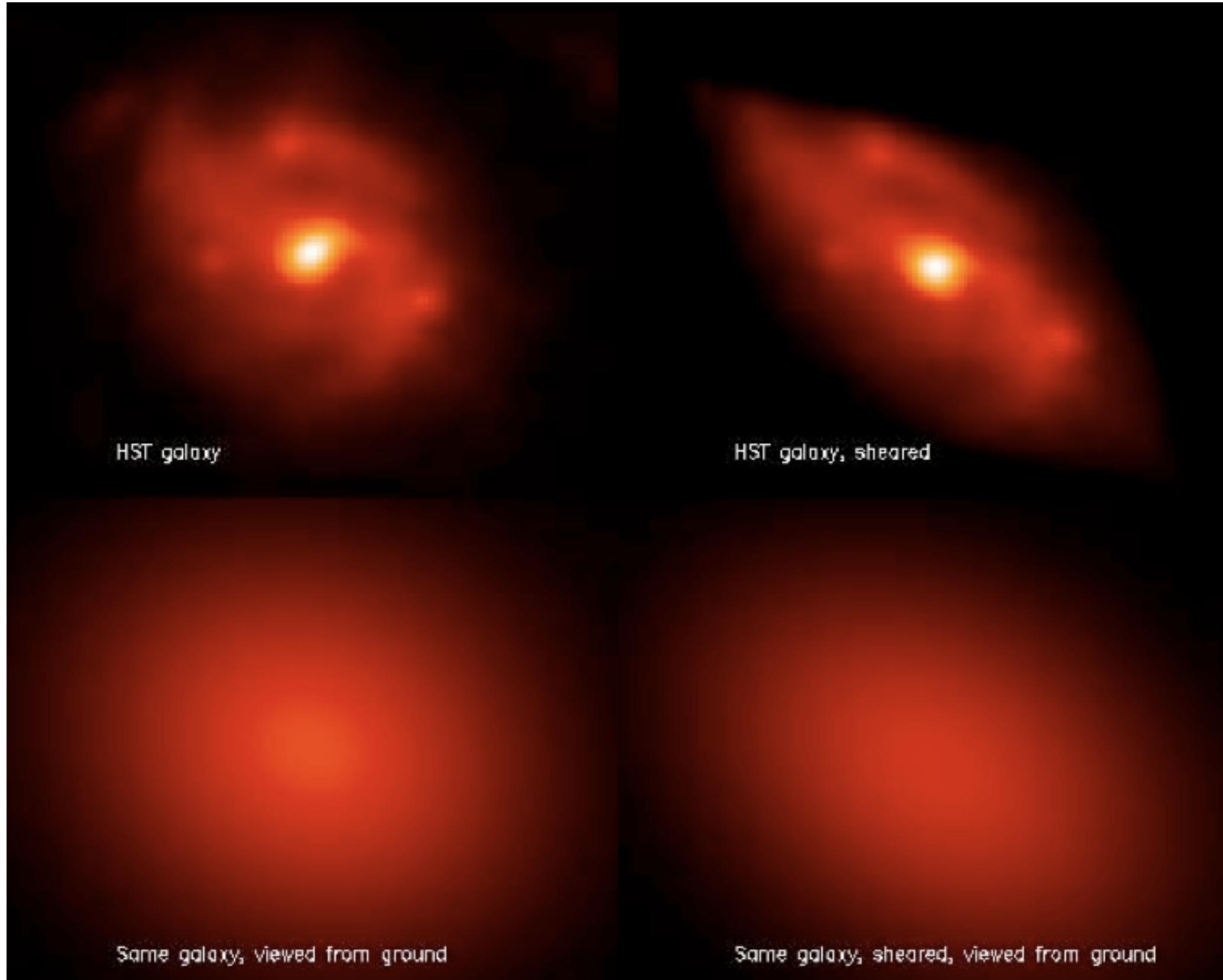
$$\gamma_1 = \frac{M_{1,1} - M_{2,2}}{M_{1,1} + M_{2,2}}, \gamma_2 = \frac{2M_{1,2}}{M_{1,1} + M_{2,2}}, \quad M_{i,j} = \int \theta_i \theta_j S(\theta) w(\theta) d\theta^2$$

$M_{1,1} - M_{2,2}$ and $2M_{1,2}$ correspond respectively to the flattening along the x axis and the 45° axis. $M_{1,1} + M_{2,2}$ is related to the size.

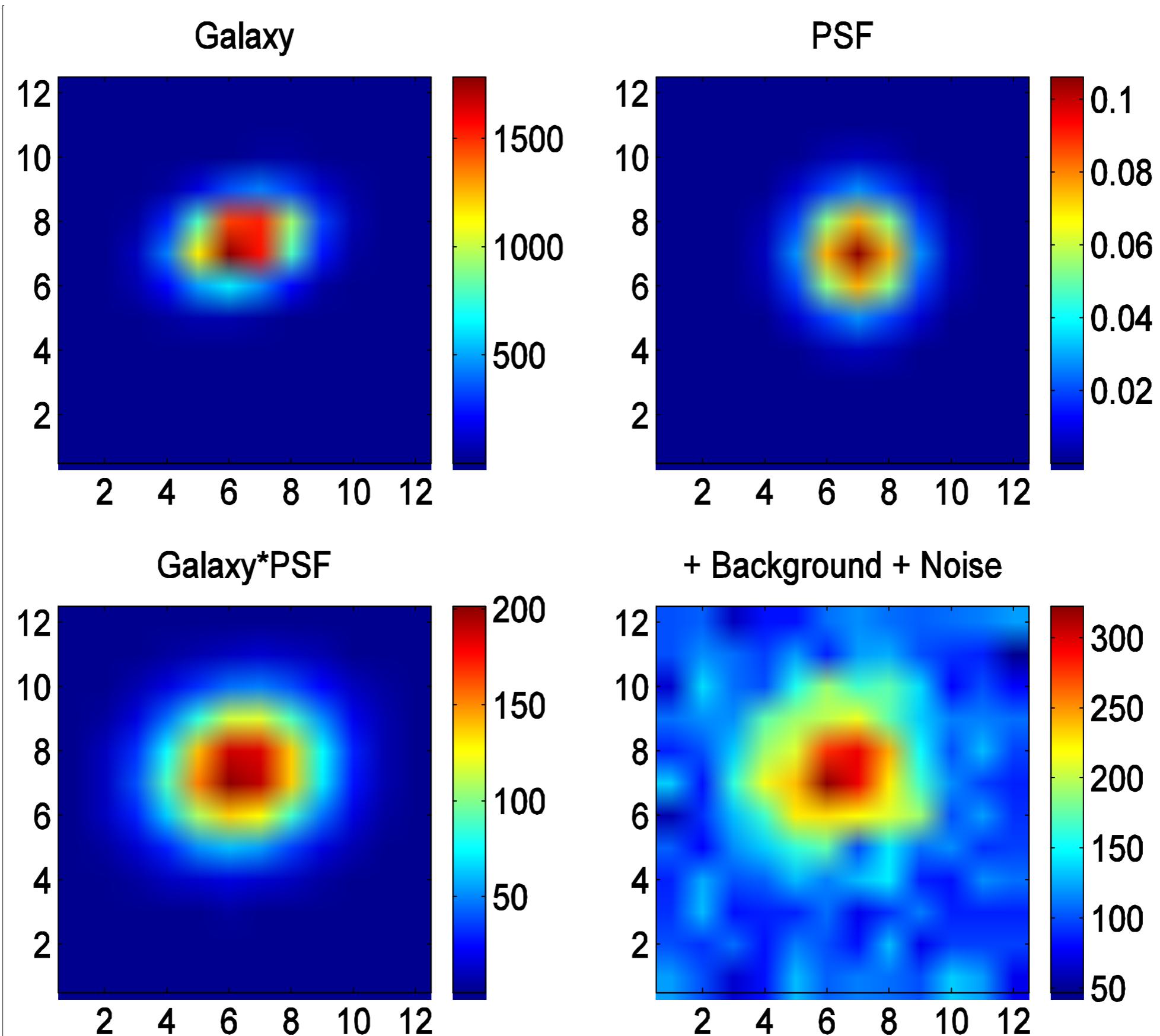
PB 1: We need accurate measurements from noisy data



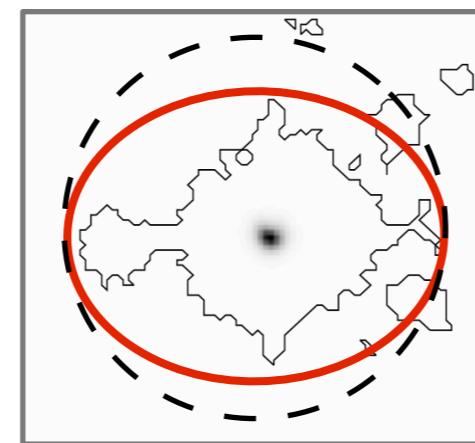
Motivation for spatial observations



Weakly Lensed Galaxies



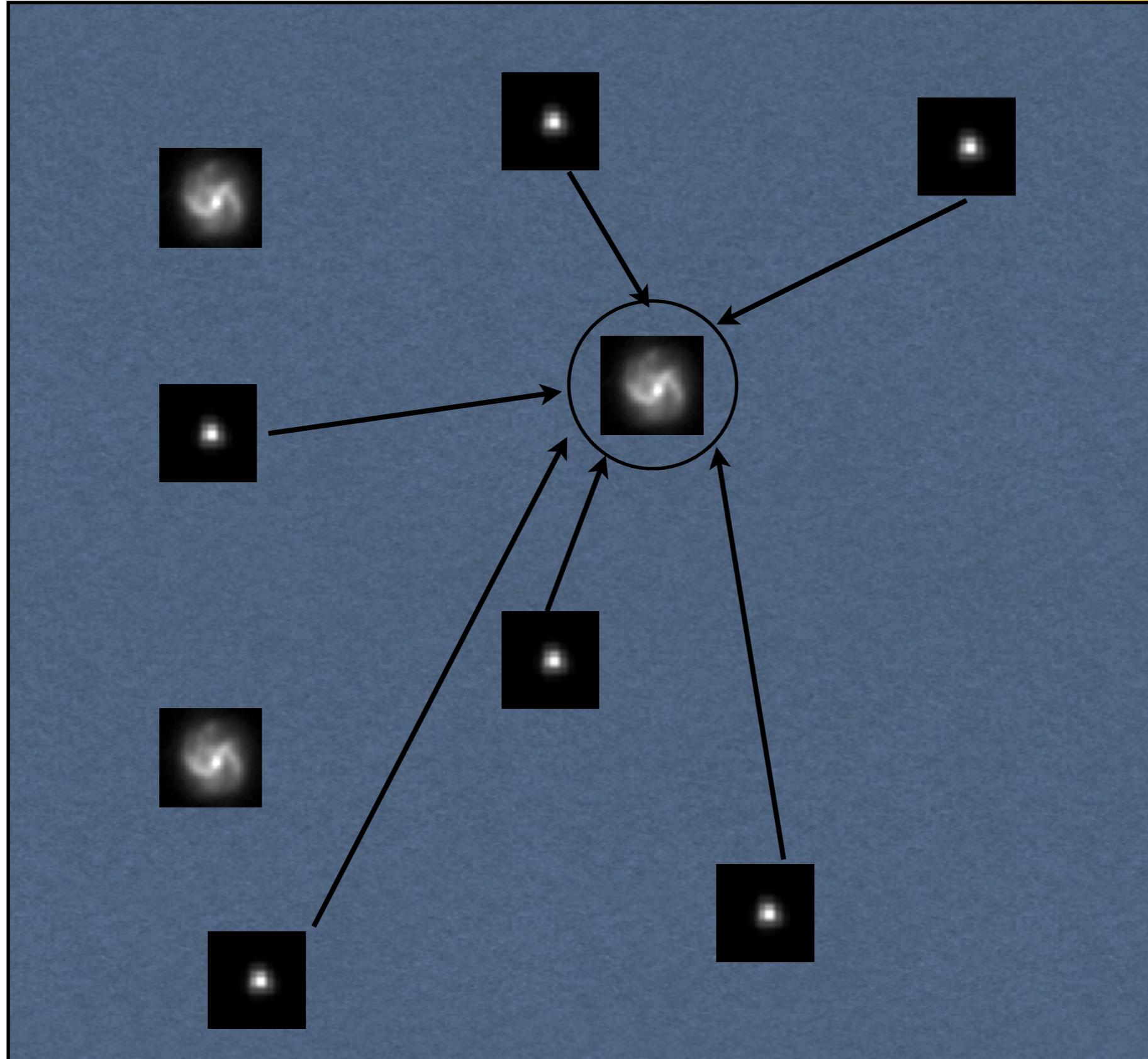
Galaxies are convolved by an asymmetric PSF



PB 2: Shape measurements must be deconvolved

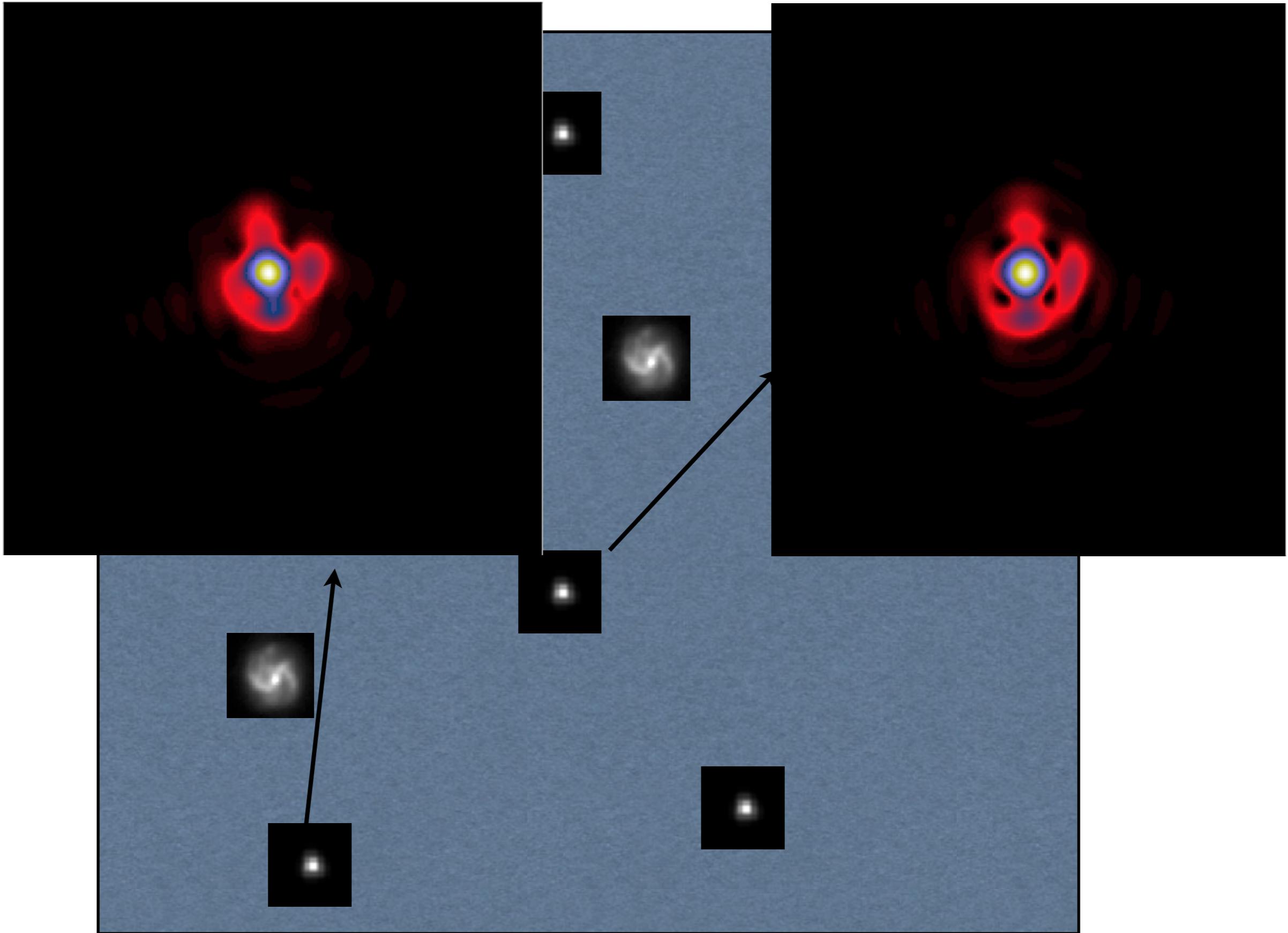
Methods: Moments (KSB), Shapelets, Forward-Fitting,
Bayesian estimation, etc

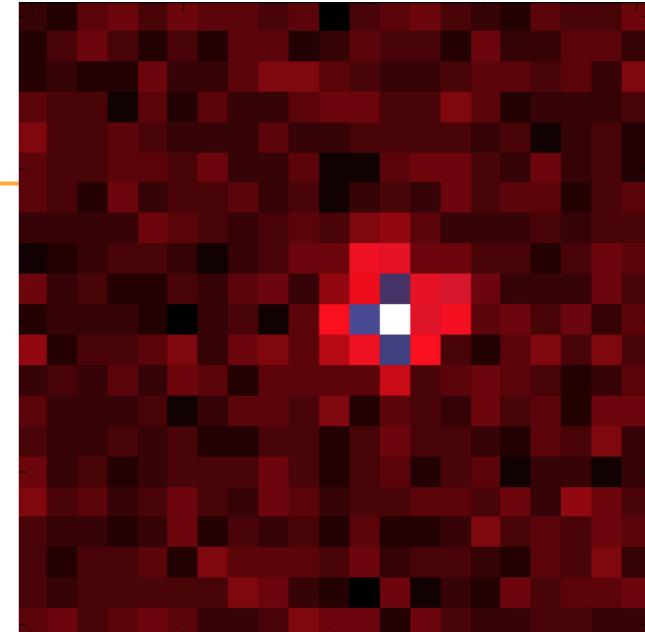
Space Variant PSF



PB 3: We need to interpolate the PSF shape !

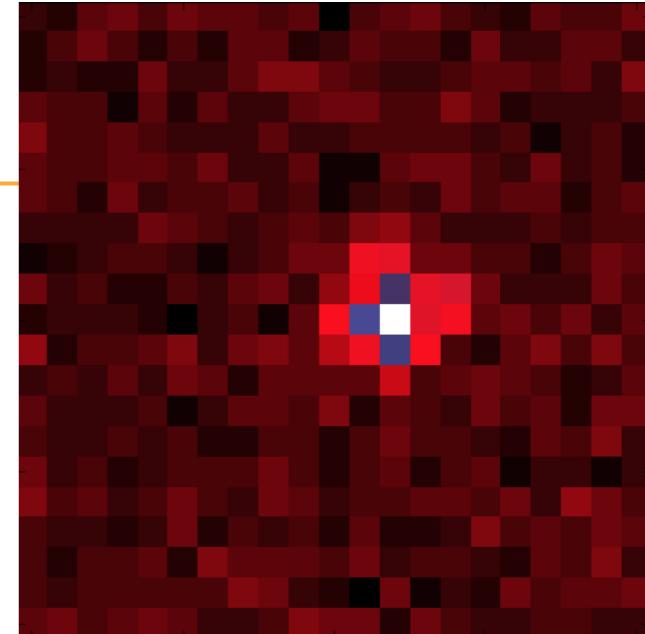
PSF Variability





- PSF Modeling
 - Undersampling
 - Space dependency
 - Time dependency
 - Wavelength dependency

- F. M. Mboula, J.-L. Starck, S. Ronayette, K. Okumura, and J. Amiaux, [Super-resolution method using sparse regularization for point-spread function recovery](#). A&A, 575, id.A86, 2015.
- F. Ngole, J.-L Starck, et al, “Constraint matrix factorization for space variant PSFs field restoration”, submitted, 2016



- PSF Modeling
 - Undersampling
 - Space dependency
 - Time dependency
 - Wavelength dependency

What has been done:

Undersampling + space dependency = **Monochromatic PSF field restoration**

- F. M. Mboula, J.-L. Starck, S. Ronayette, K. Okumura, and J. Amiaux, [Super-resolution method using sparse regularization for point-spread function recovery](#). A&A, 575, id.A86, 2015.
- F. Ngole, J.-L Starck, et al, “Constraint matrix factorization for space variant PSFs field restoration”, submitted, 2016

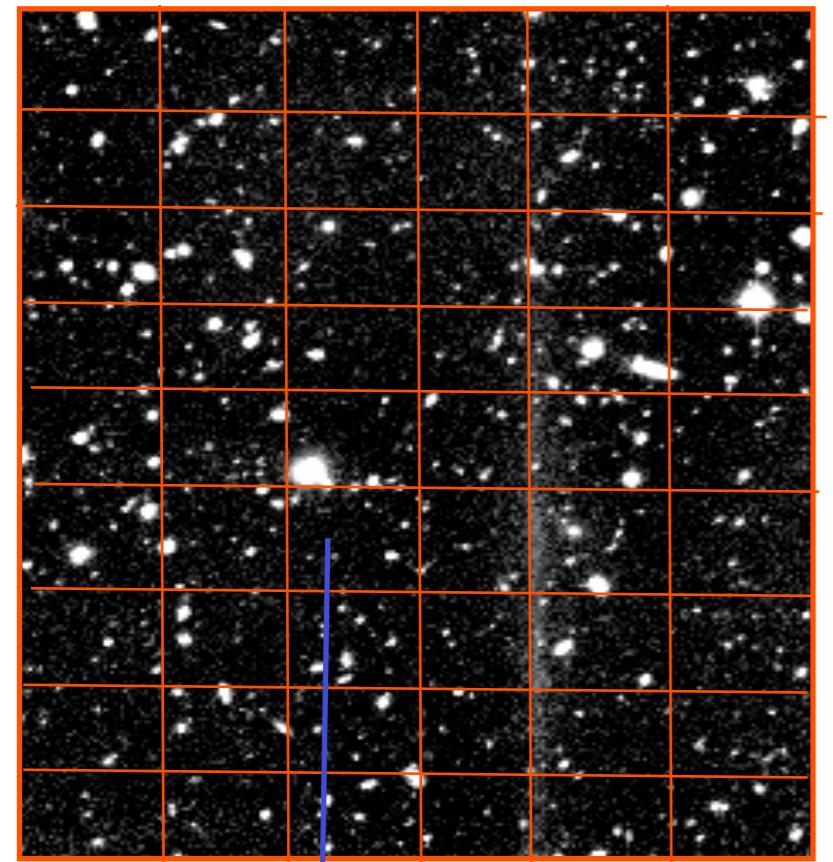
Intrinsic Ellipticities

✓ Galaxies have an intrinsic ellipticity

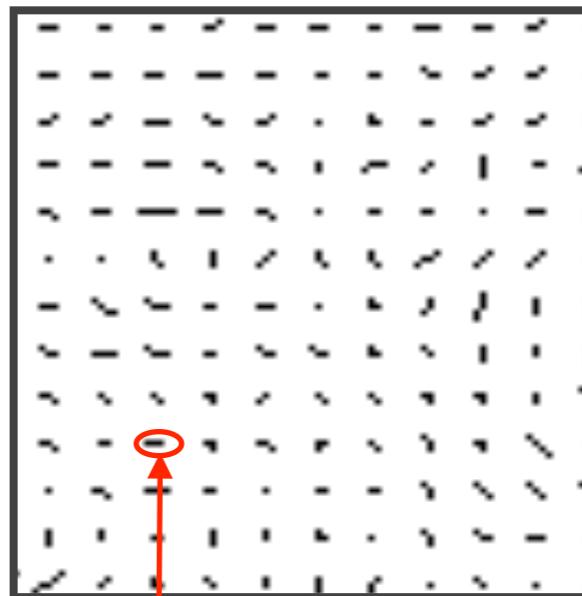
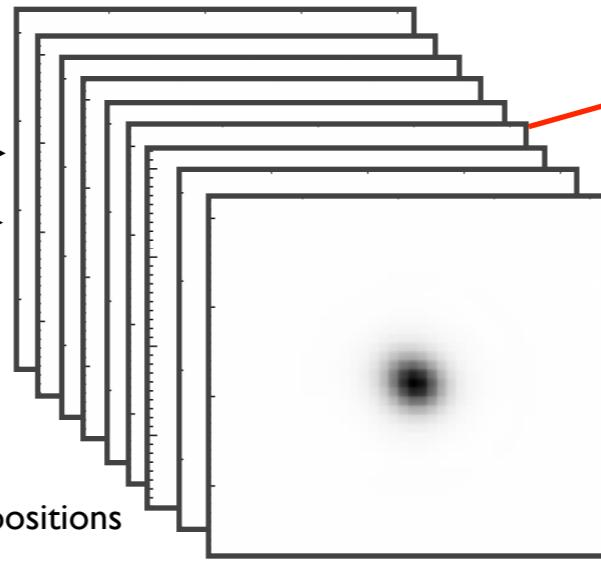
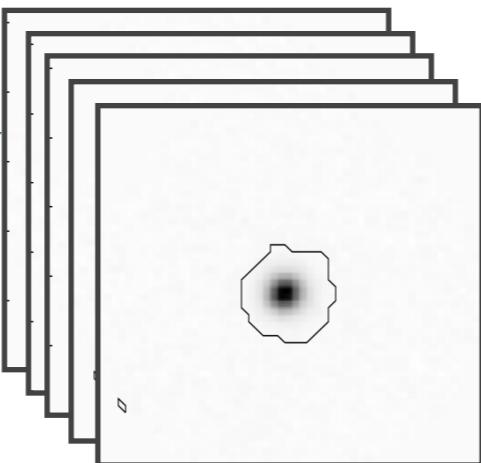
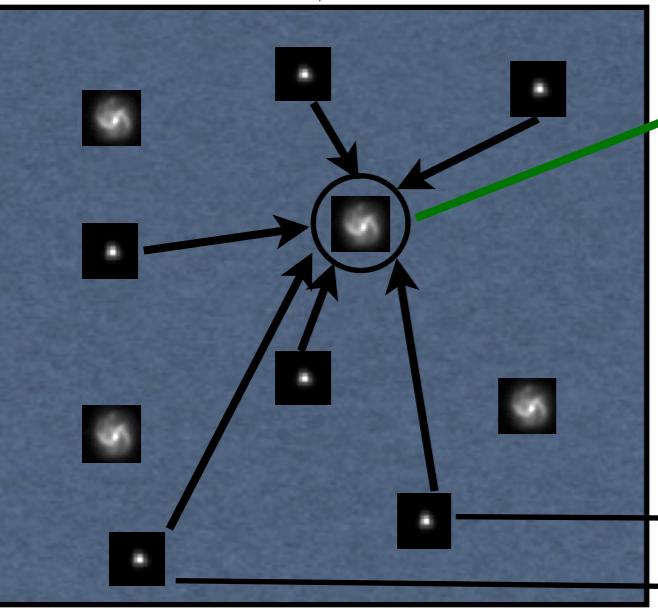


PB 4: We need to correct the measurements from the intrinsic ellipticity

Shear Catalog & Map



Few undersampled images of a given galaxy



PSF superresolution + Interpolation +
Shape Measurement

Many PSF at other positions

$$\mathbf{y}_k = \mathbf{M}_k \mathbf{x}_k + \mathbf{n}_k, \quad k = 1..n$$

- Learning technique as an alternative to PCA to identify the *eigenvectors* PSF (note that PCA cannot handle subsampled data): **Resolved Components Analysis (RCA)**

F. Ngle, J.-L Starck, et al, "Constraint matrix factorization for space variant PSFs field restoration", submitted, 2016

Our model:

$$\text{PSF}^{(k)} = x_k = \sum_{i=1}^r a_{i,k} s_i$$

$a_{i,k}$ = coefficient corresponding to contribution of the i-th vector to the k-th PSF.

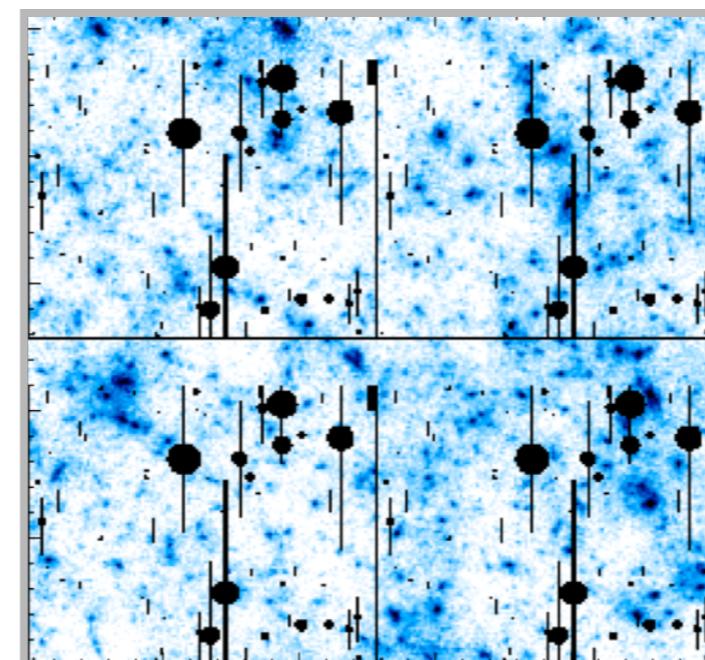
s_i = ith *vector* (2D image)

- **Joint estimations** of super-resolved PSFs at stars positions

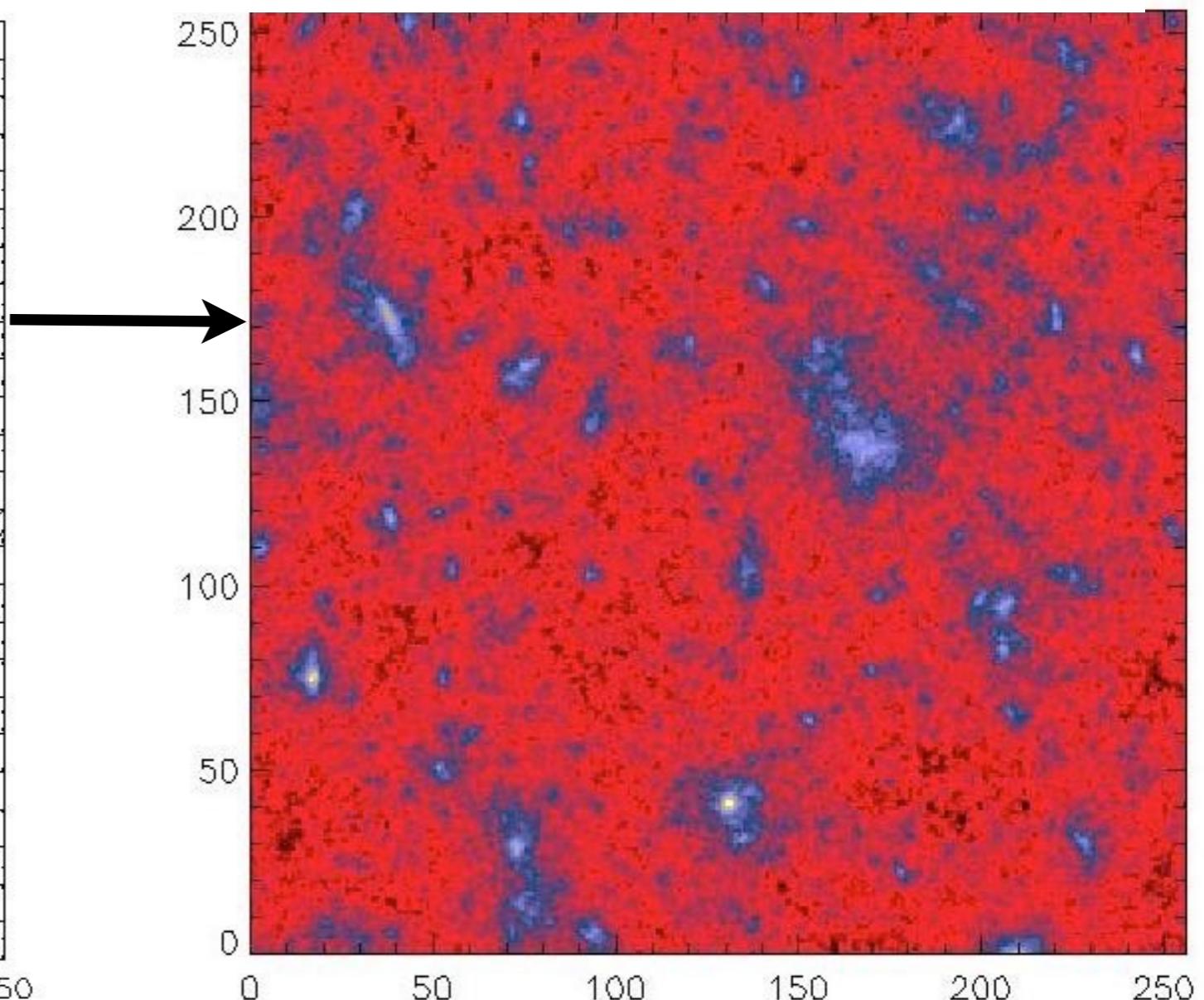
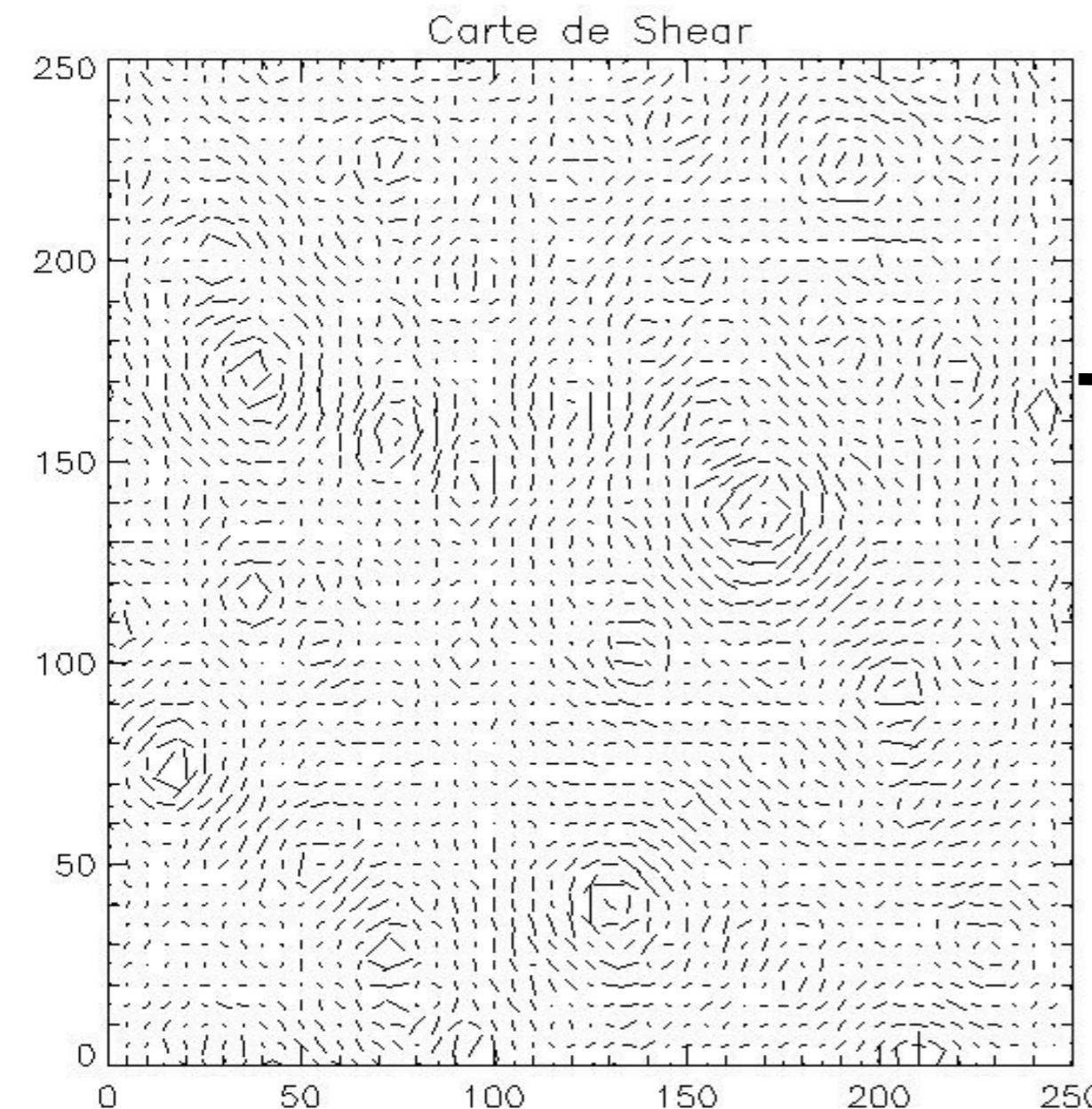
- » Positivity constraint
- » Low rank constraint: Constraint the PSFs to be a linear combination of the *eigenvectors* PSFs
- » Smoothness constraint on each s_i
- » Proximity constraint on the coefficients $\alpha_i^{(*)}$: the closer are the stars, the more the coefficients of the linear combination are similar.

We need to solve a triple inverse problem !!!

- I) Determine the PSF at any position from the measured PSF.
- 2) Measure the galaxy shear and correct it from the PSF.
- 3) Correct the shear from intrinsic ellipticities
+ noise and **missing data!!!**



Missing data

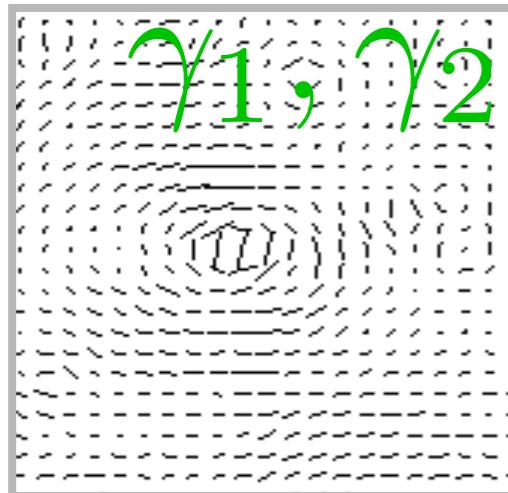


Inversion Equations

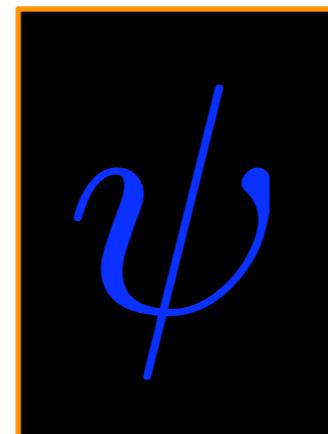
★ J.L. Starck, S. Pires and A. Réfrégier, A&A, Vol. 451, pp 1139-1150, 2006.

★ S. Pires, J.-L. Starck and A. Refregier, "Light on Dark Matter with Weak Gravitational Lensing", IEEE Signal Processing Magazine, 27, 1, pp 76--85, 2010.

SIMULATED SHEAR MAP

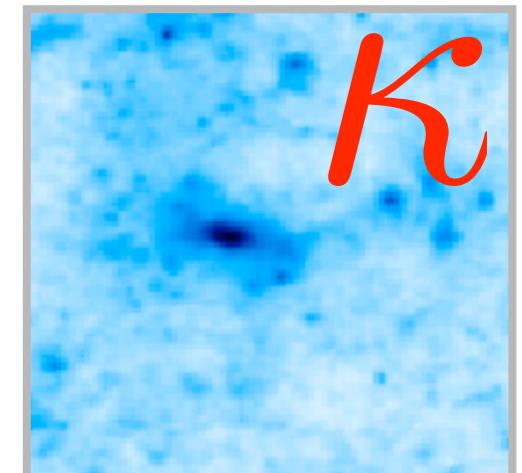


LENSING POTENTIAL



SIMULATED MASS MAP
(Vale & White, 2003)

$$\frac{1}{2} (\partial_1^2 + \partial_2^2) \psi = \kappa$$



From mass to shear:

$$\gamma_i = \hat{P}_i \kappa$$

From shear to mass:

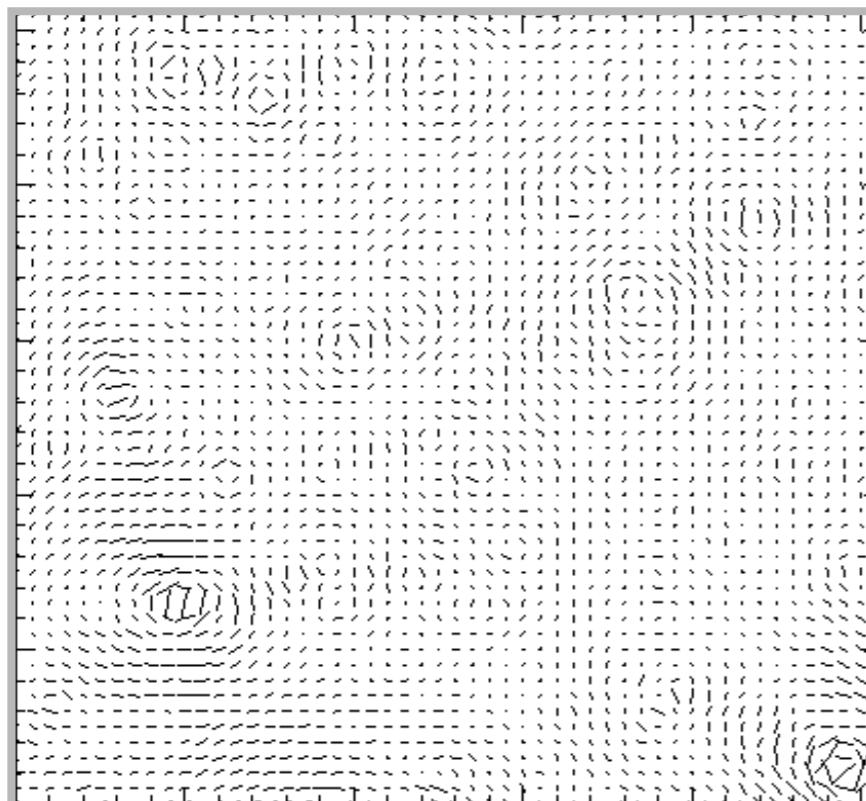
$$\kappa = \hat{P}_1 \gamma_1 + \hat{P}_2 \gamma_2$$

$$\hat{P}_1(k) = \frac{k_1^2 - k_2^2}{k^2}$$

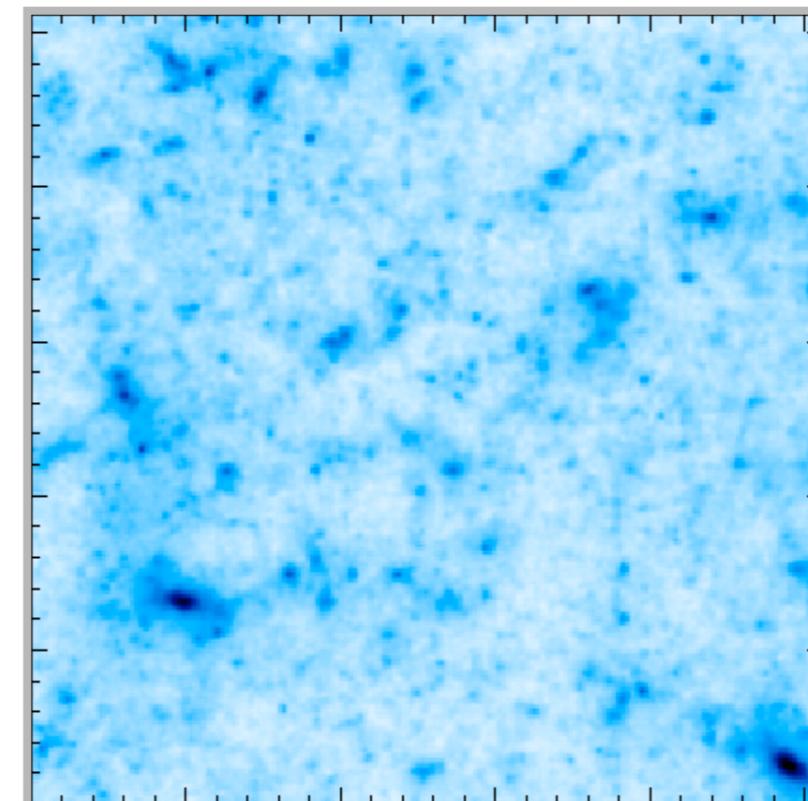
$$\hat{P}_2(k) = \frac{2k_1 k_2}{k^2}$$

$$\begin{pmatrix} \hat{E}(\mathbf{k}) = \hat{\kappa}(\mathbf{k}) \\ \hat{B}(\mathbf{k}) \end{pmatrix} = \underbrace{\frac{1}{|\mathbf{k}|^2} \begin{pmatrix} k_1^2 - k_2^2 & 2k_1 k_2 \\ 2k_1 k_2 & -k_1^2 + k_2^2 \end{pmatrix}}_{A_\kappa} \begin{pmatrix} \hat{\gamma}_1(\mathbf{k}) \\ \hat{\gamma}_2(\mathbf{k}) \end{pmatrix}$$

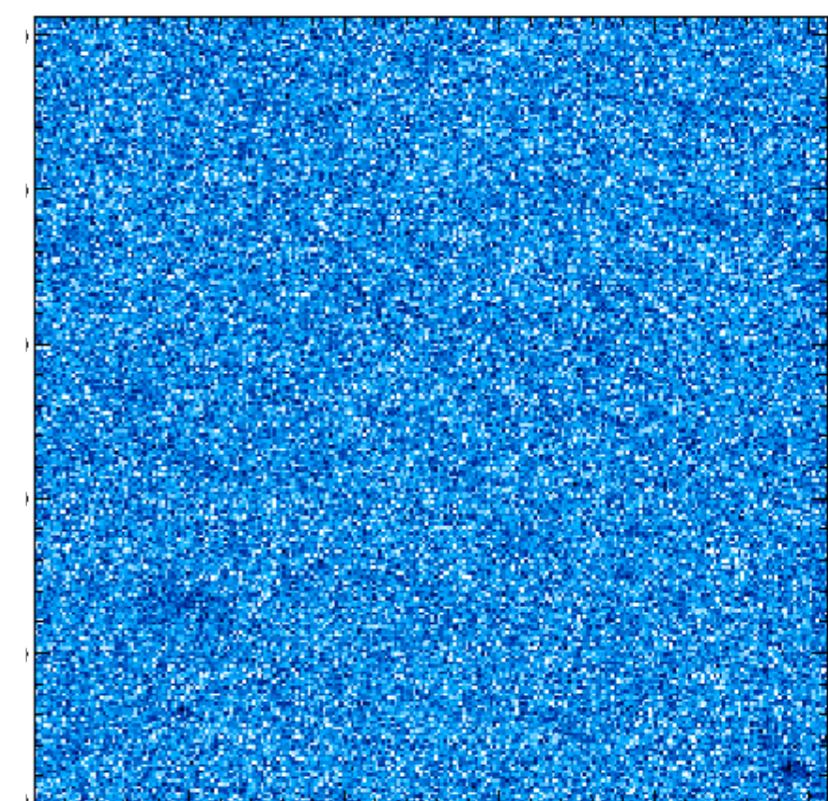
A Simple Reconstruction is Very Noisy



Shear map

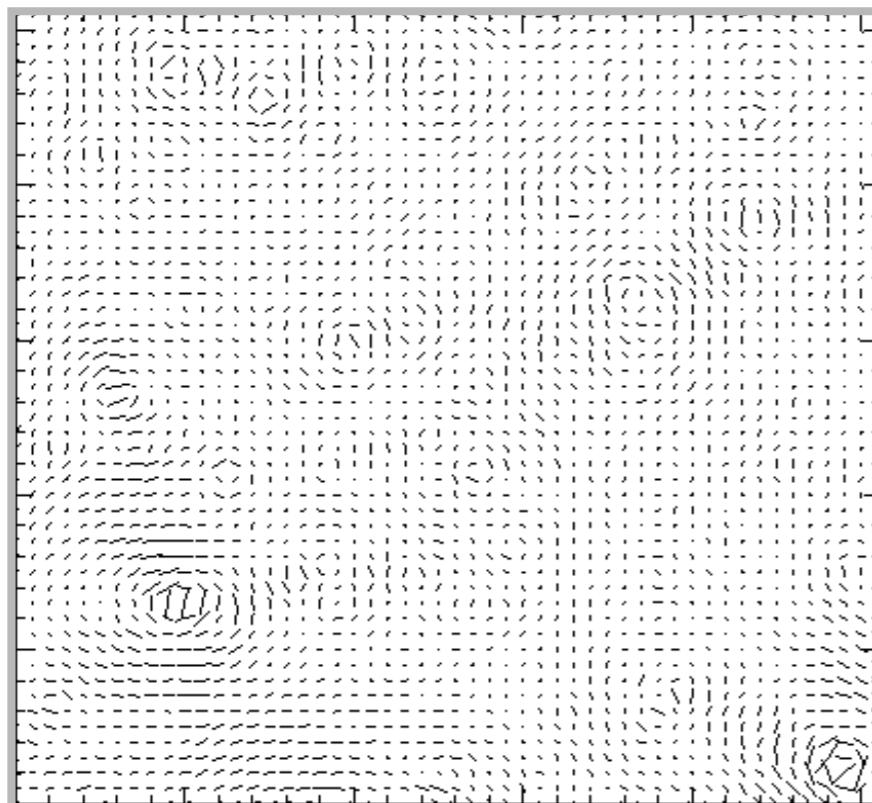


Original mass map

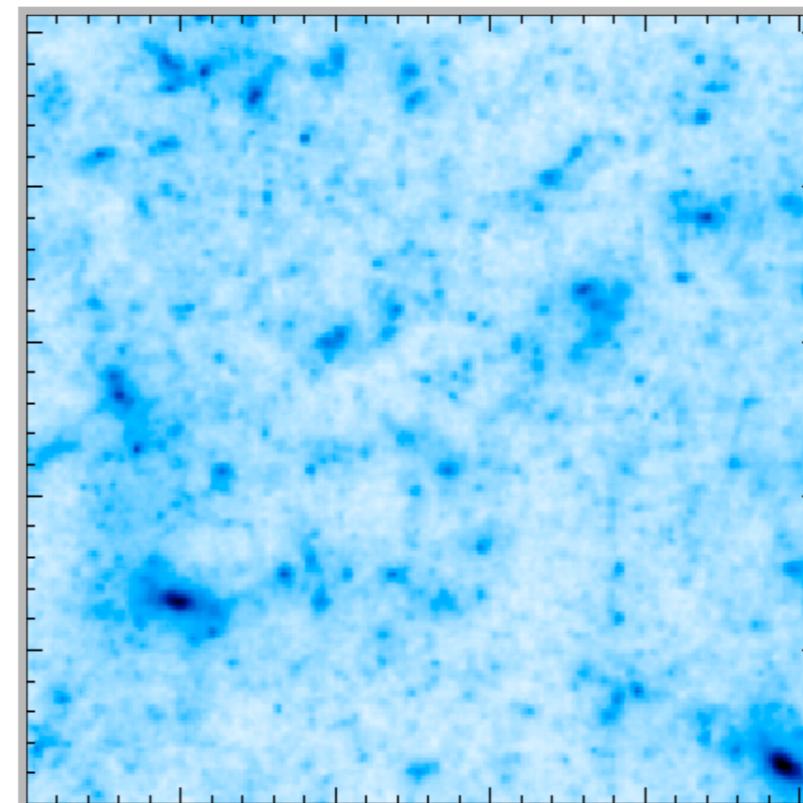


Mass map (space observations)

A Simple Reconstruction is Very Noisy



Shear map

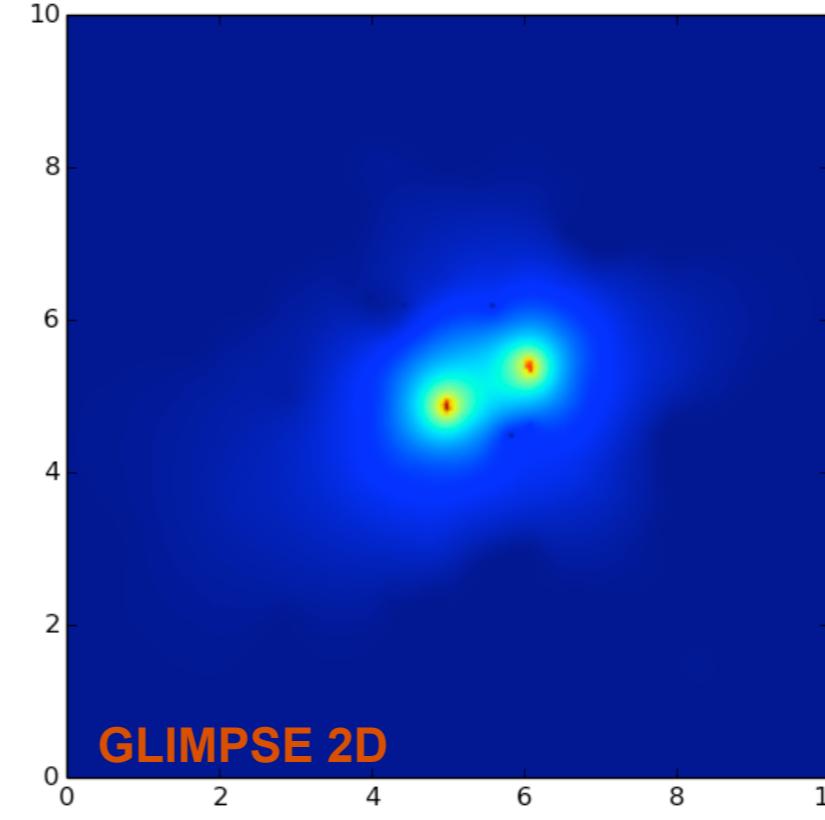
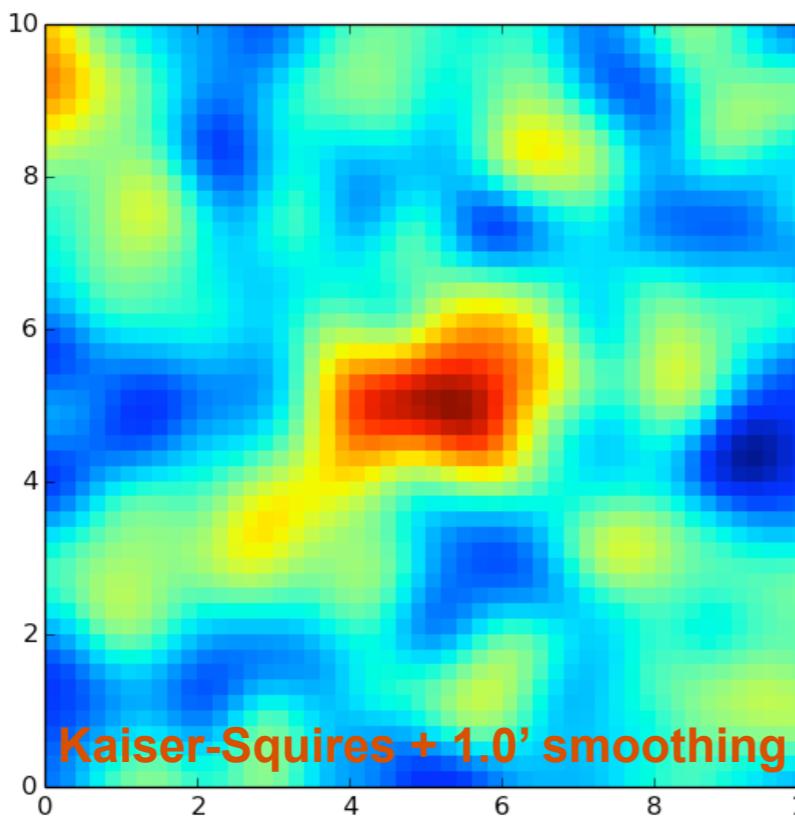
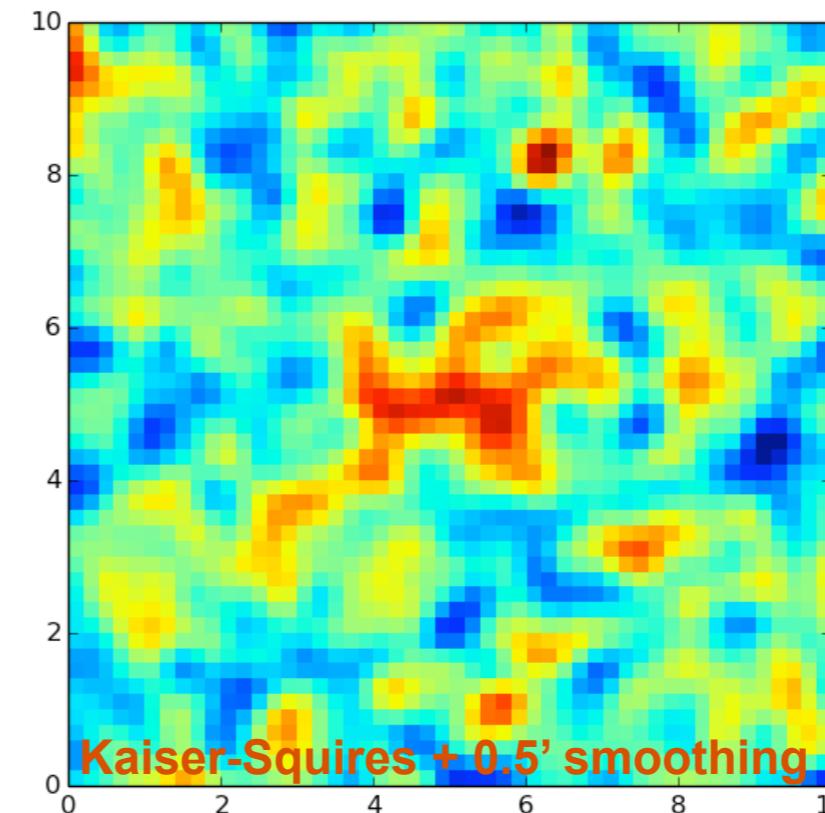
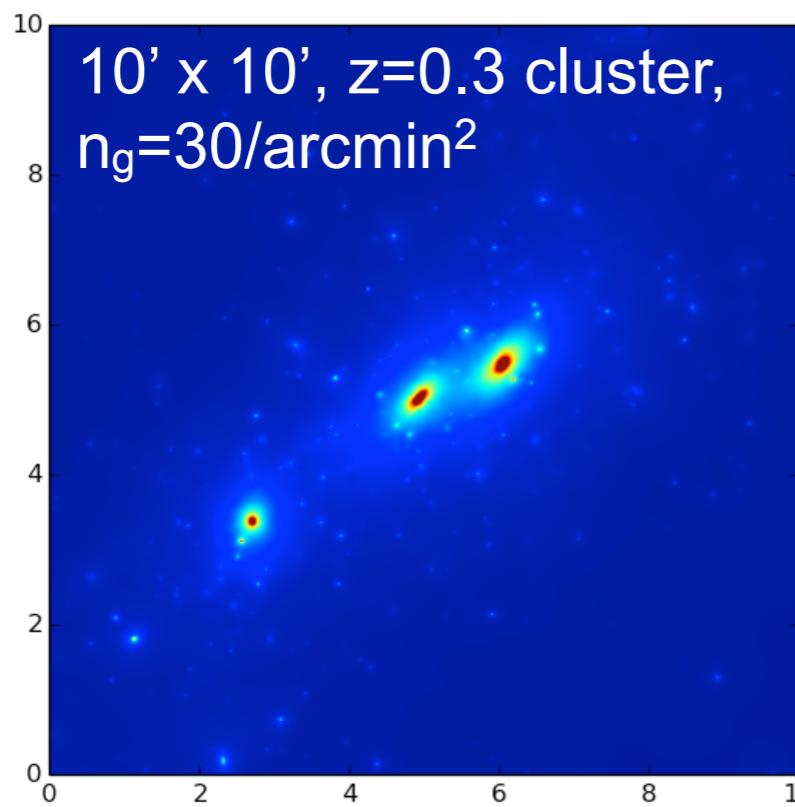


Original mass map

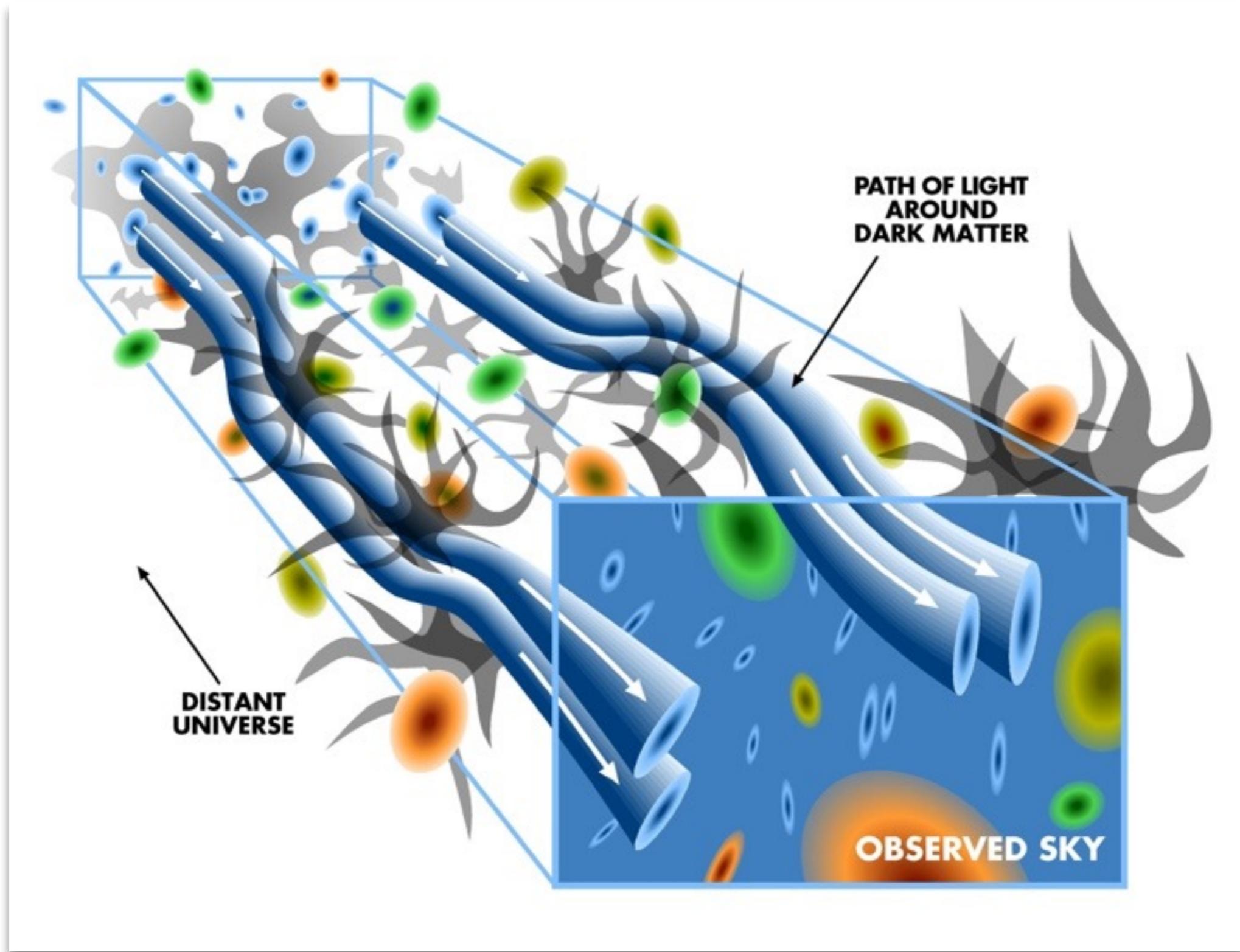


Mass map (space observations)

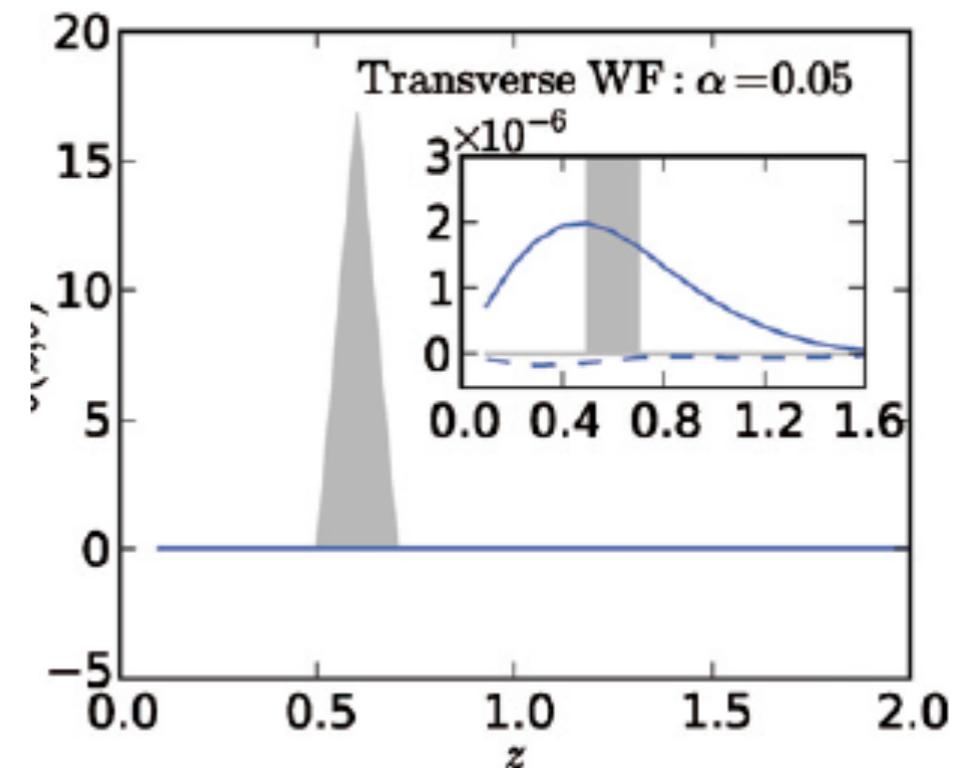
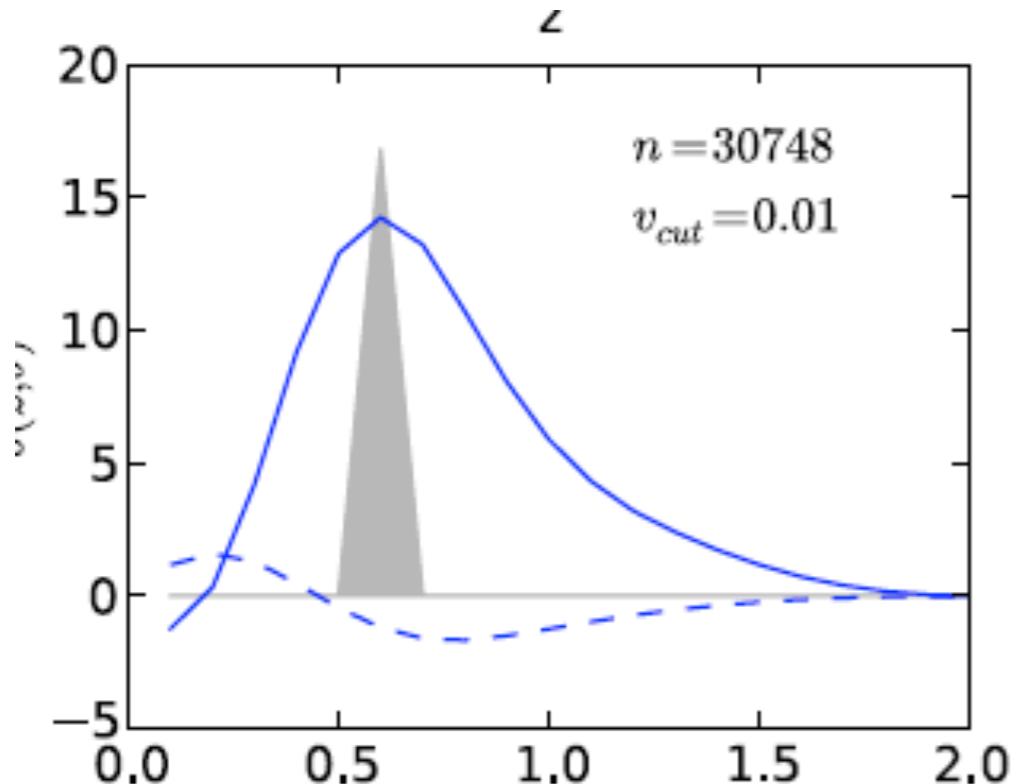
Missing Data + Noise



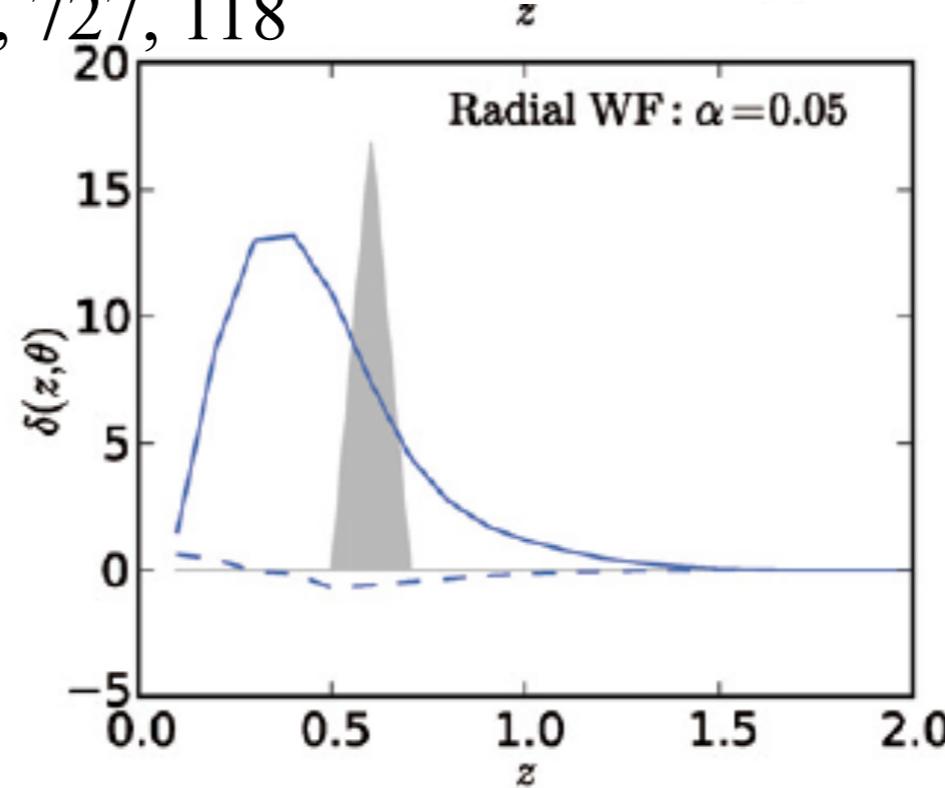
3D Weak Lensing



RESULTS WITH LINEAR METHODS

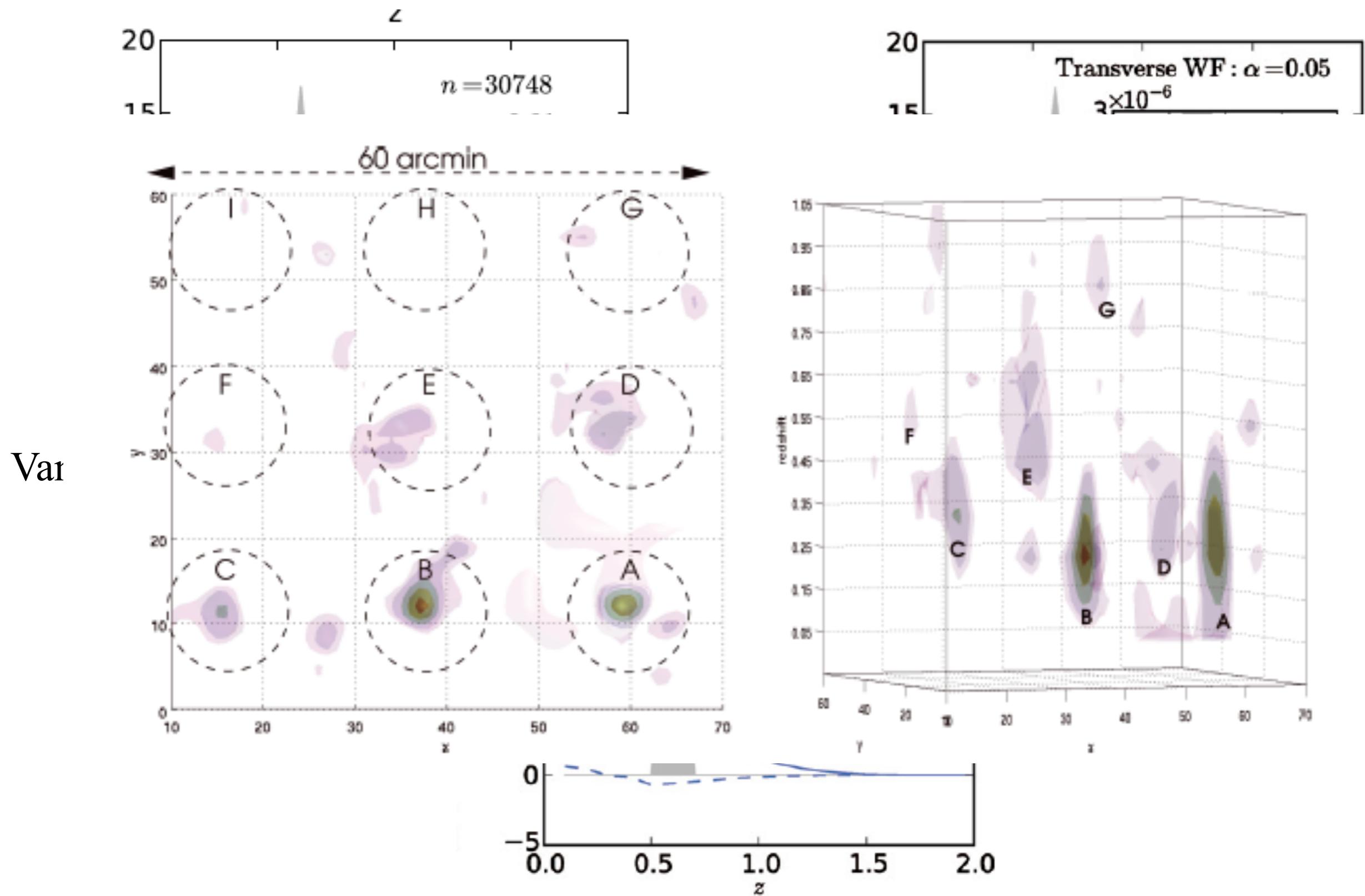


Vanderplas et al. 2011, *ApJ*, 727, 118



Simon et al. 2009, *MNRAS*, 399, 48

RESULTS WITH LINEAR METHODS



Weak Lensing Theory

$$\gamma(\theta) = \frac{1}{\pi} \int d^2\theta' \mathcal{D}(\theta - \theta') \kappa(\theta')$$

Kappa (or convergence) is a dimensionless surface mass density of the lens

$$\kappa(\theta, w) = \frac{3H_0^2 \Omega_M}{2c^2} \int_0^w dw' \frac{f_K(w') f_K(w-w')} {f_K(w)} \frac{\delta[f_K(w')\theta, w']} {a(w')} ,$$

f_K is the angular diameter distance, which is a function of the comoving radial distance r and the curvature K .

$$\begin{aligned}\gamma &= P_{\gamma\kappa} \kappa + n_\gamma, \\ \kappa &= Q\delta + n\end{aligned}$$

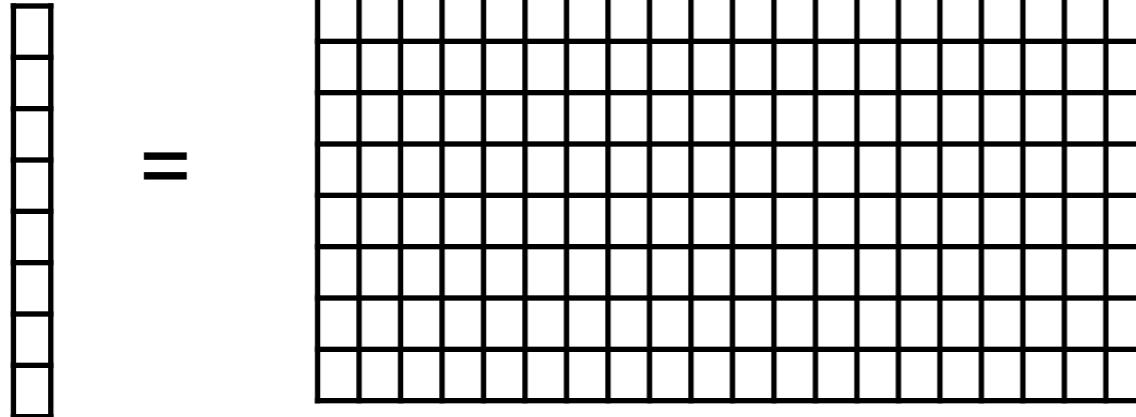
$$\gamma = R\delta + n$$

- Galaxies are not intrinsically circular: intrinsic ellipticity $\sim 0.2\text{-}0.3$; gravitational shear ~ 0.02
- Reconstructions require knowledge of distances to galaxies

3D Weak Lensing

κ

Q



δ

N

+



δ is sparse.

Q spreads out the information in δ along κ bins.

More unknown than measurements

Full 3D Weak Lensing

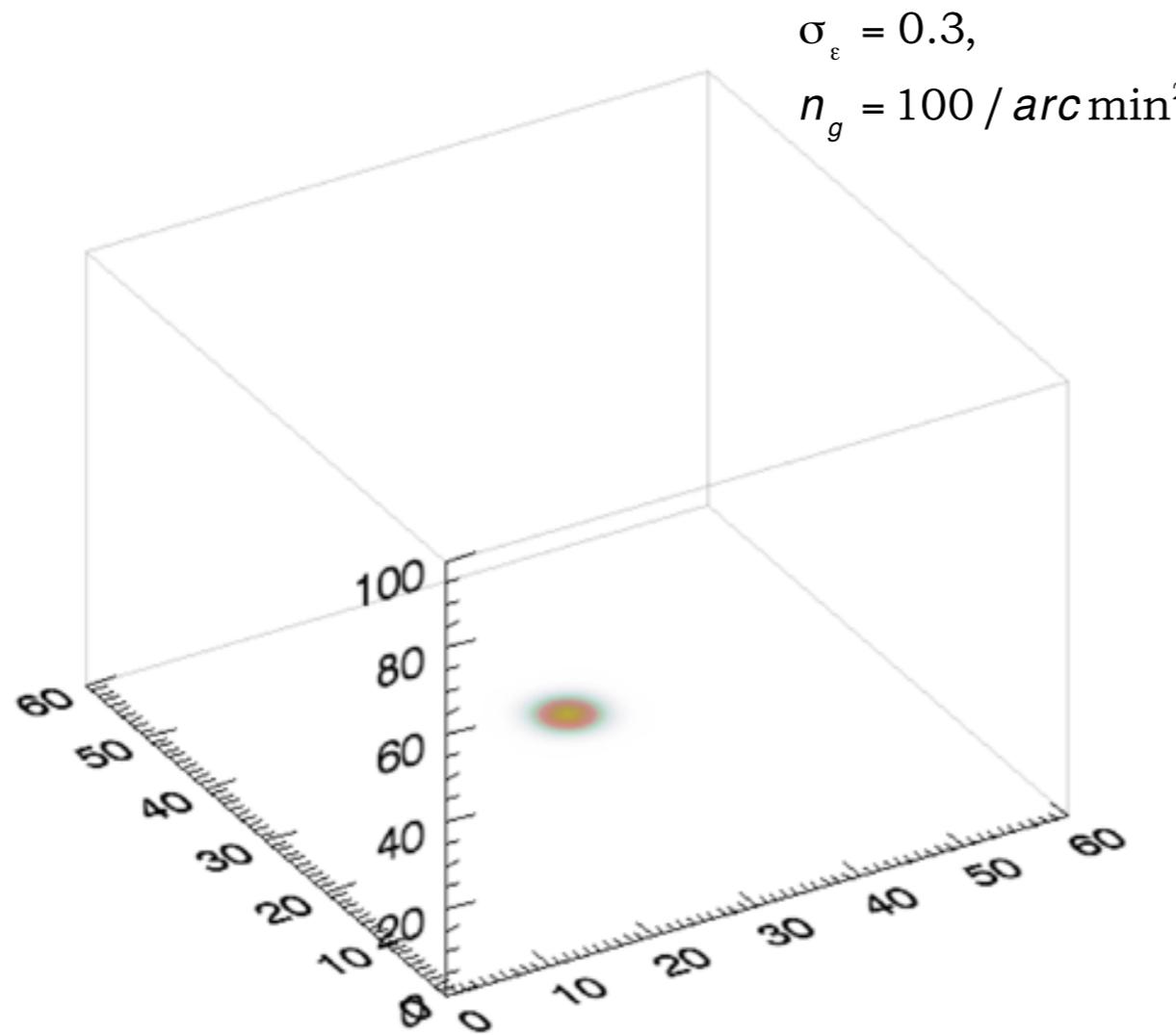
A. Leonard, F.X. Dupe, and J.-L. Starck, "[A Compressed Sensing Approach to 3D Weak Lensing](#)", *Astronomy and Astrophysics*, 539, A85, 2012.

A. Leonard, F. Lanusse, J-L. Starck, GLIMPSE: Accurate 3D weak lensing reconstruction using sparsity, *Astronomy and Astrophysics*, *Astronomy and Astrophysics*, 2014

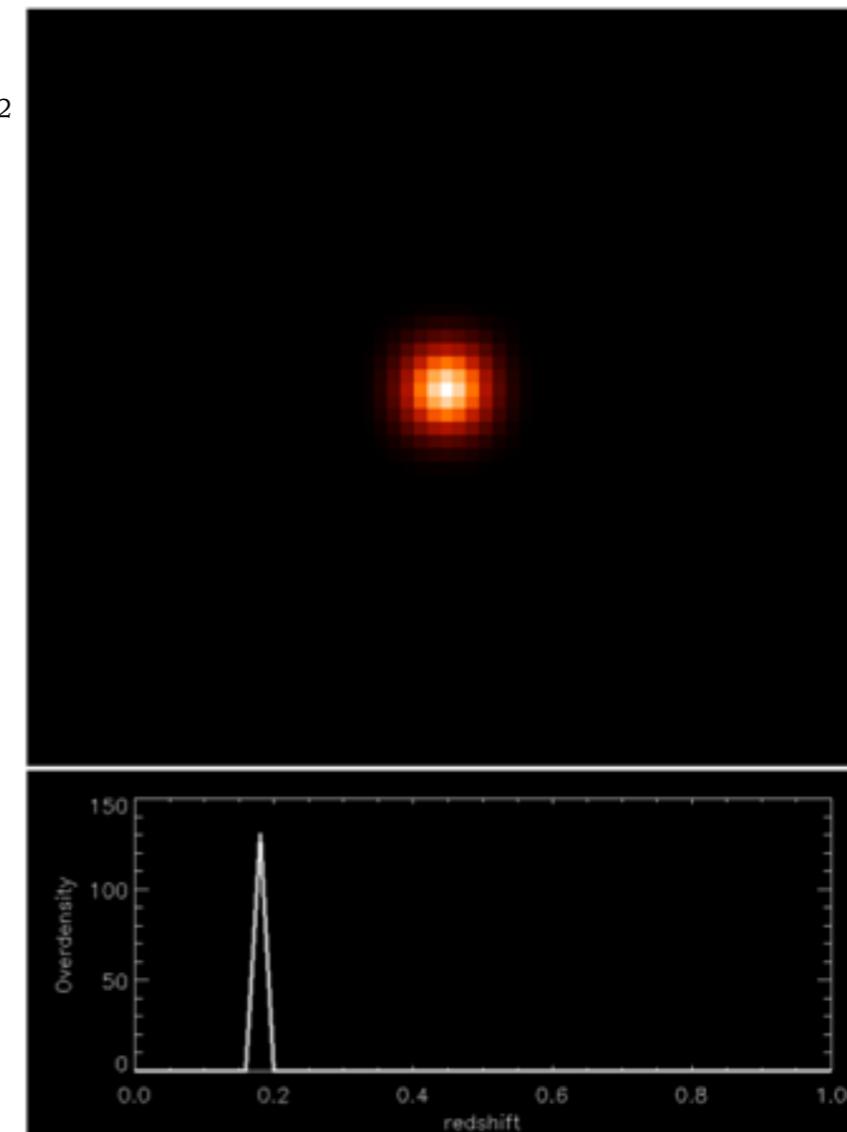
$$\min_{\alpha} \| \alpha \|_1 \text{ s.t. } \frac{1}{2} \| \gamma - R\Phi\alpha \|_{\Sigma^{-1}}^2 \leq \epsilon$$

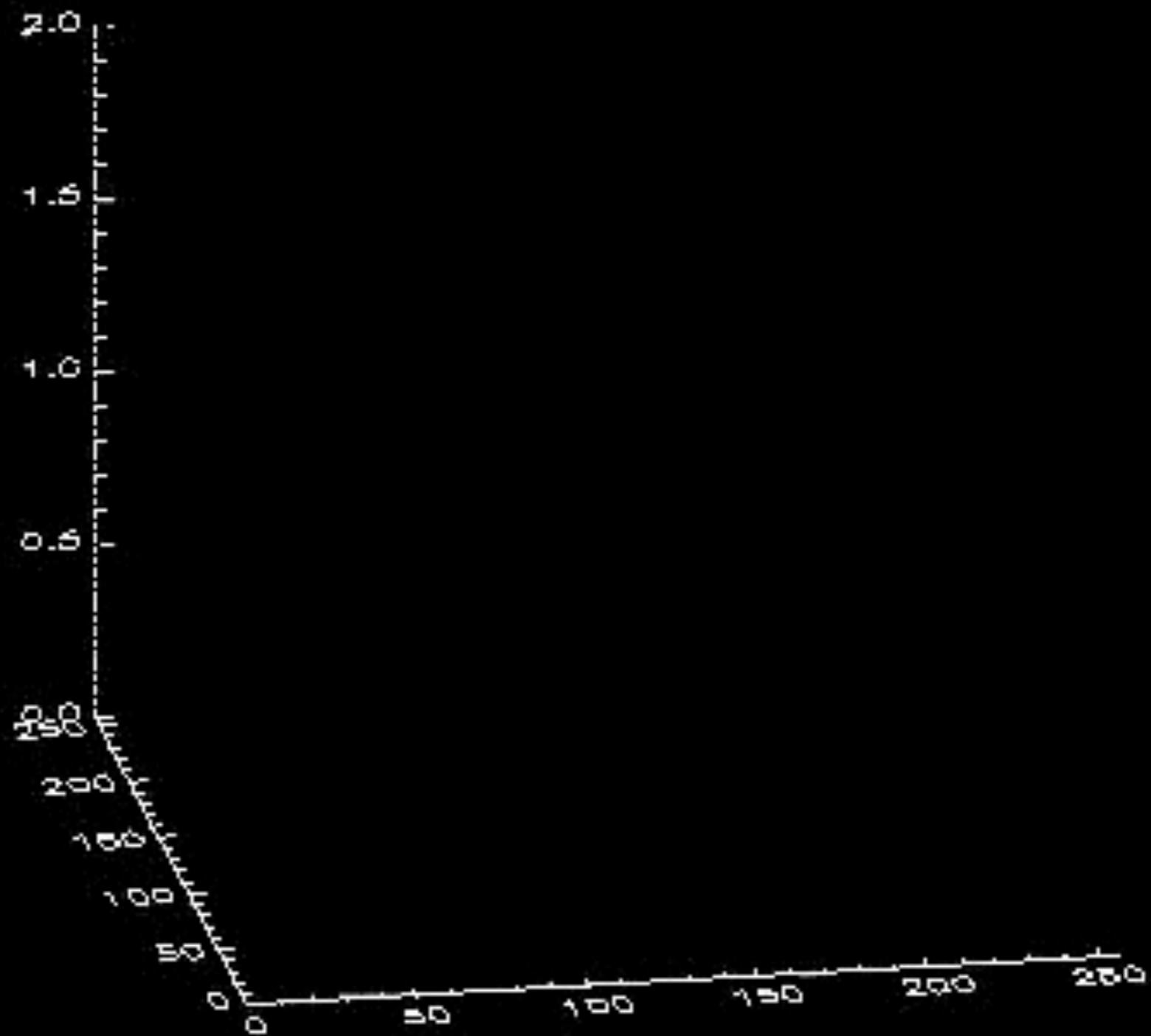
$\delta = \Phi\alpha$ Φ = 2D Wavelet Transform on each redshift bin

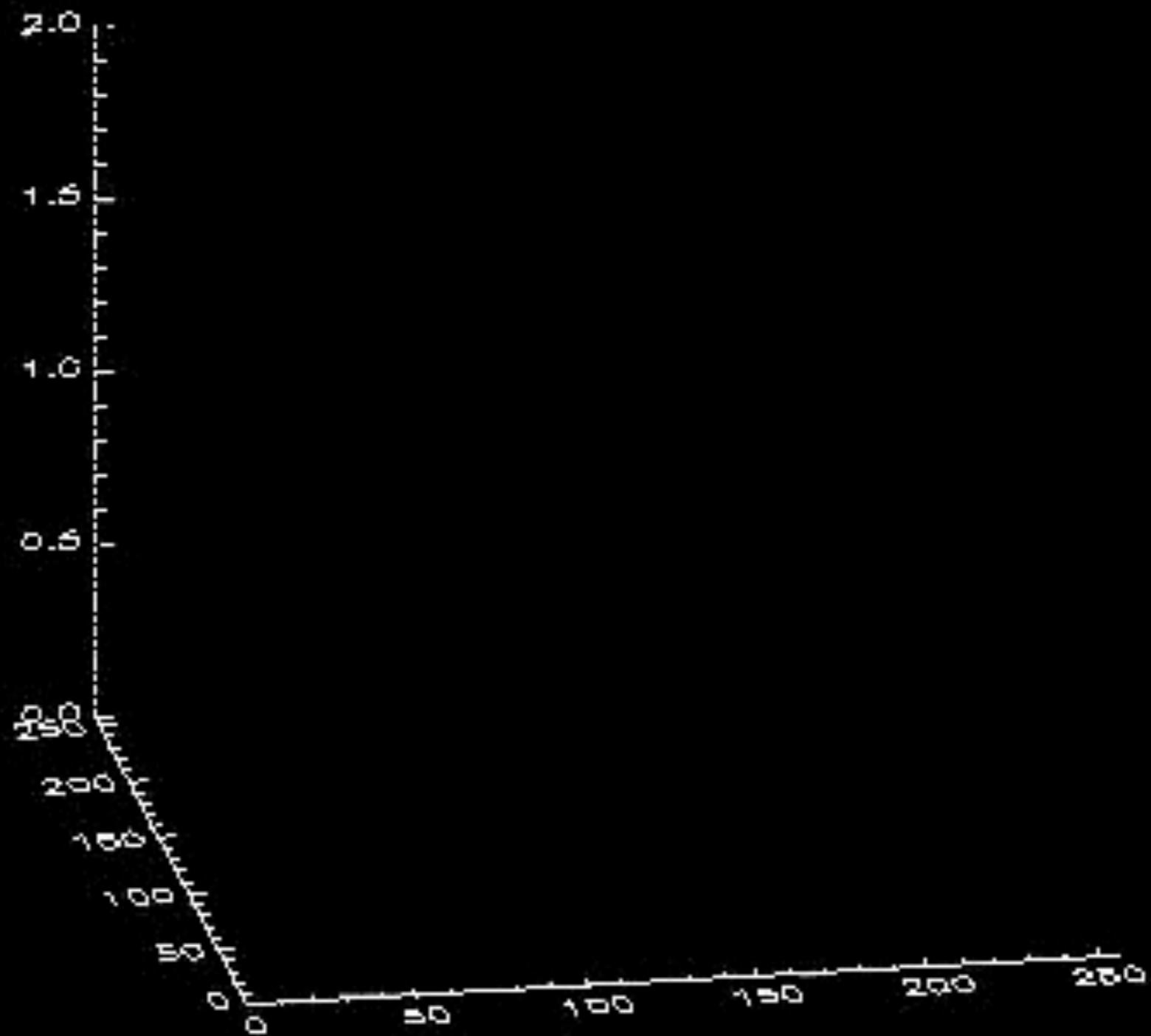
Example of the reconstruction of a cluster at redshift $z=0.2$ using 5 times more reconstruction bin:

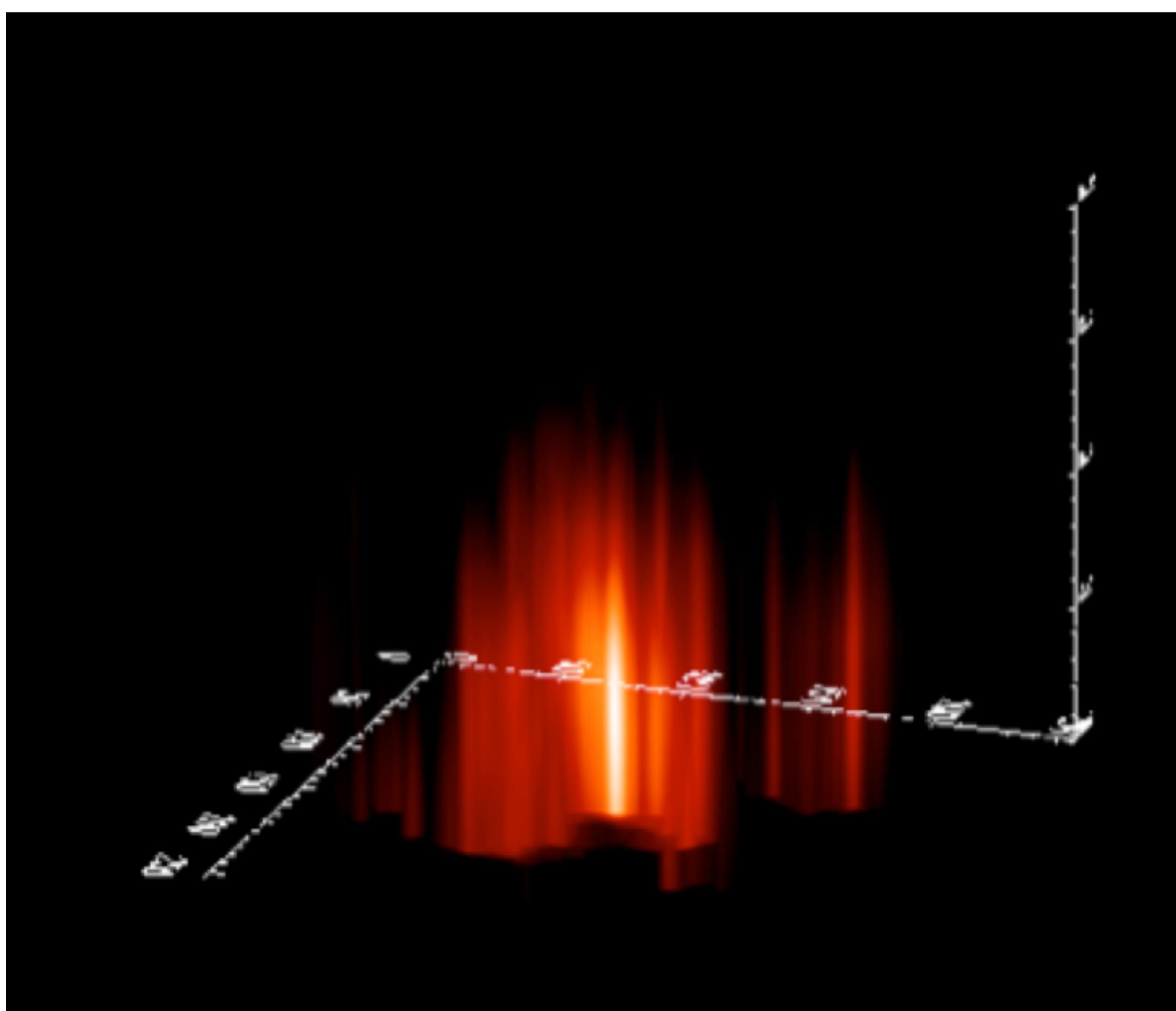


$$\sigma_{\varepsilon} = 0.3, \\ n_g = 100 / \text{arc min}^2$$









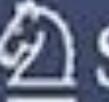


Jean-Luc Starck
Fionn Murtagh

Astronomical Image and Data Analysis

Second Edition



 Springer

Copyrighted Material

Jean-Luc Starck • Fionn Murtagh • Jalal M. Fadili

Sparse Image and Signal Processing



Wavelets
and Related
Geometric Multiscale
Analysis
Second Edition

Copyrighted Material