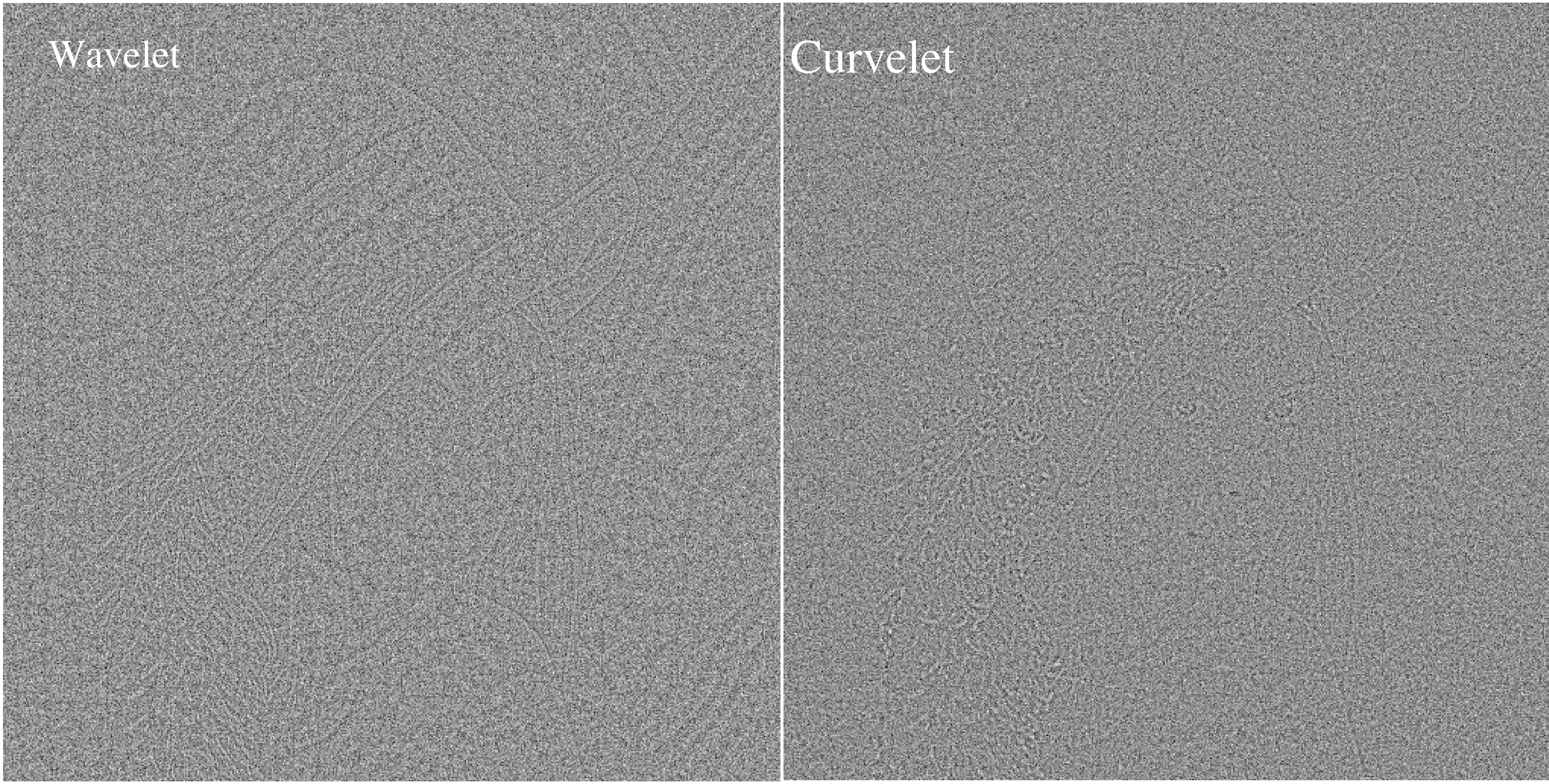
Data Analysis Using a Combination of Redundant Multiscale Transforms

Jean-Luc Starck

Service d'Astrophysique, CEA-Saclay, France.
jstarck@cea.fr, http://jstarck.free.fr

Collaborators:

D.L. Donoho, *Department of Statistics, Stanford*
E. Candès, *Department of Applied Mathematics, California Institute of Technology*
M. Elad, *The Technion, Israel Institute of Technology*


Wavelet
Curvelet

RESTORATION: HOW TO COMBINE SEVERAL MULTISCALE TRANSFORMS ?

The problem we need to solve for image restoration is to make sure that our reconstruction will incorporate information judged as significant by any of our representations.

<u>Very High Quality Image Restoration</u>*, in Signal and Image Processing IX, San Diego, 1-4 August, 2001, Eds Laine, Andrew F.; Unser, Michael A.; Aldroubi, Akram, Vol. 4478, pp 9-19, 2001.*

Notations:

Consider K linear transforms $T_{1}, ..., T_{K}$ and α_k the coefficients of x after applying $T_K : \alpha_k = T_k s, \quad s = T_k^{-1} \alpha_k$.

We propose solving the following optimization problem:

$$\min \quad \text{Complexity\_penalty}(\tilde{s}), \qquad \text{subject to } \tilde{s} \in C$$

Where C is the set of vectors which obey the linear constraints:

$$\tilde{s} > 0, \qquad \text{positivity constraint}$$

$$\left|\left(T_k \tilde{s} - T_k s\right)_l\right| \le e, \quad if \left(T_k s\right)_l \qquad \text{is significant}$$

The second constraint guarantees that the reconstruction will take into account any pattern which is detected by any of the K transforms.

We use an ℓ_1 penalty on the coefficient sequence.

$$\min \|T\tilde{s}\|_{\ell_1}, \quad \text{subject to} \quad s \in C,$$

There are other possible choices of complexity penalties; for instance, an alternative would be

$$\min \|\tilde{s}\|_{TV}, \quad \text{subject to} \quad s \in C.$$

where $\| \cdot \|_{TV}$ is the Total Variation norm.

Hybrid Steepest Descent

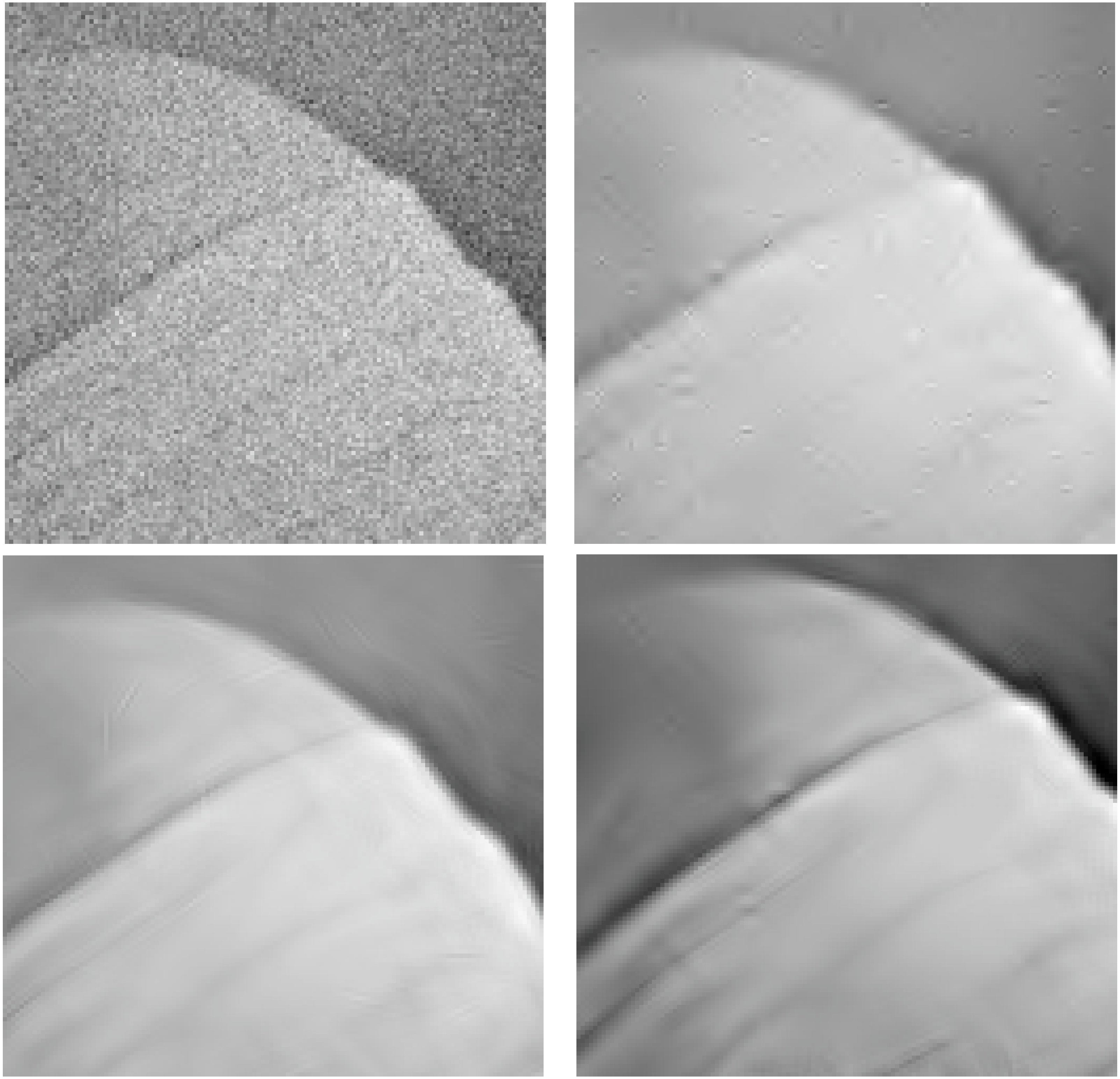
We propose solving this problem using the method of hybrid steepest descent (HSD) (Yamada, 2001). HSD consists of building the sequence

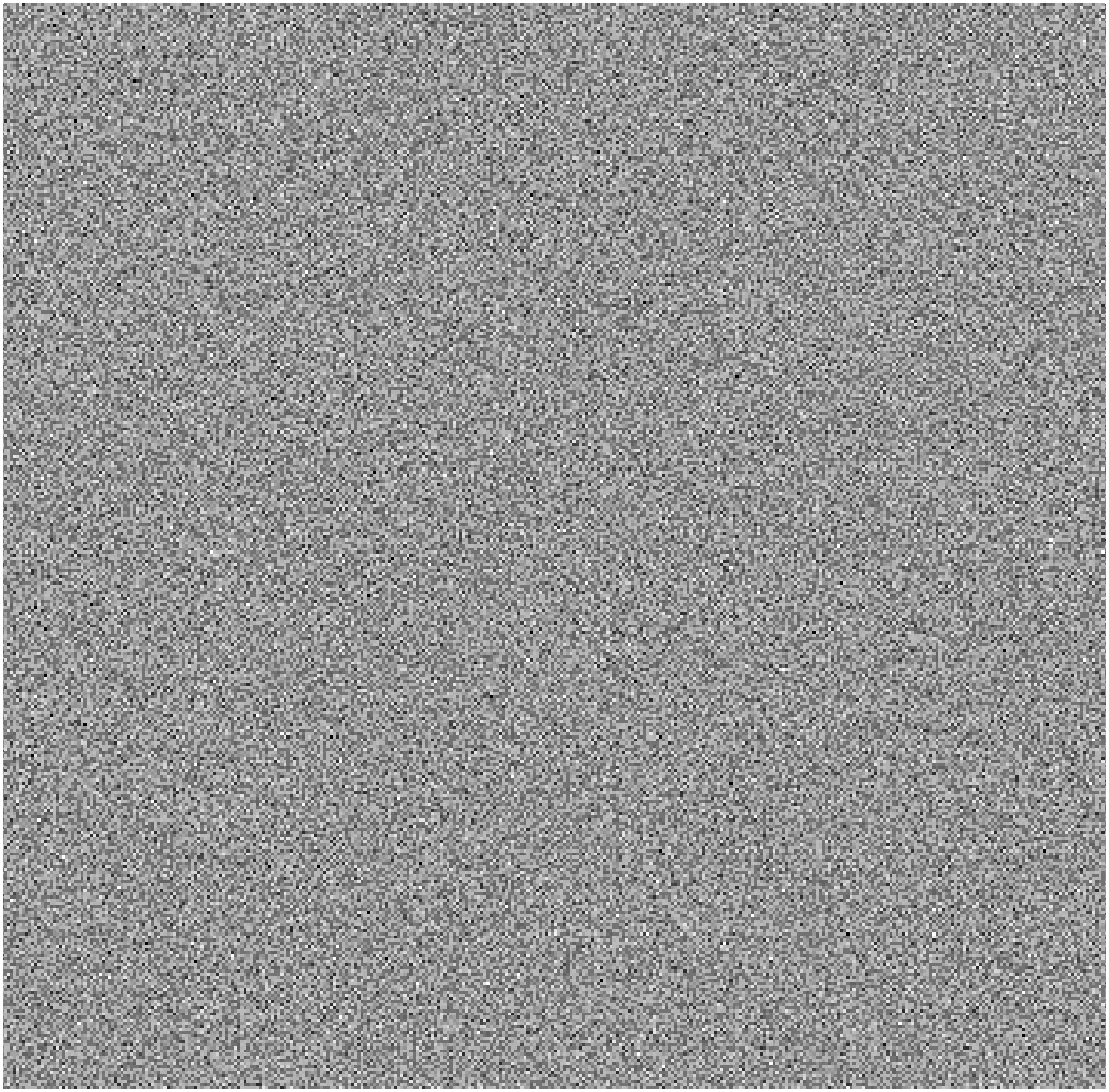
$$s^{n+1} = P(s^n) - \lambda_{n+1}\nabla_J(P(s^n));$$

Here, P is the ℓ_2 projection operator onto the feasible set C, ∇_J is the gradient of the function to be minimized, and $(\lambda_n)_{n\geq 1}$ is a sequence obeying $(\lambda_n)_{n\geq 1} \in [0,1]$ and $\lim_{n\to+\infty} \lambda_n = 0$.

The combined filtering algorithm

1. Initialize $L_{\max} = 1$, the number of iterations N_i, and $\delta_\lambda = \frac{L_{\max}}{N_i}$.
2. Estimate the noise standard deviation σ, and set $e_k = \frac{\sigma}{2}$.
3. For k = 1, .., K calculate the transform: $\alpha_k^{(s)} = T_k s$.
4. Set $\lambda = L_{\max}$, $n = 0$, and $\tilde{s}^n$ to $\mathbf{0}$.
5. While $\lambda >= 0$ do
   - $u = \tilde{s}^n$.
   - For k = 1, .., K do
     - Calculate the transform $\alpha_k = T_k u$.
     - For all coefficients $\alpha_{k,l}$ do
       * Calculate the residual $r_{k,l} = \alpha_{k,l}^{(s)} - \alpha_{k,l}$
       * if $\alpha_{k,l}^{(s)}$ is significant and $\mid r_{k,l} \mid > e_{k,l}$ then $\alpha_{k,l} = \alpha_{k,l}^{(s)}$
       * $\alpha_{k,l} = sgn(\alpha_{k,l})(\mid \alpha_{k,l} \mid - \lambda)_+$.
     - $u = T_k^{-1} \alpha_k$
   - Threshold negative values in u and $\tilde{s}^{n+1} = u$.
   - $n = n + 1$, $\lambda = \lambda - \delta_\lambda$, and goto 5.





DECONVOLUTION:

$$s = P * \tilde{s} + N$$

We propose solving the following optimization problem:

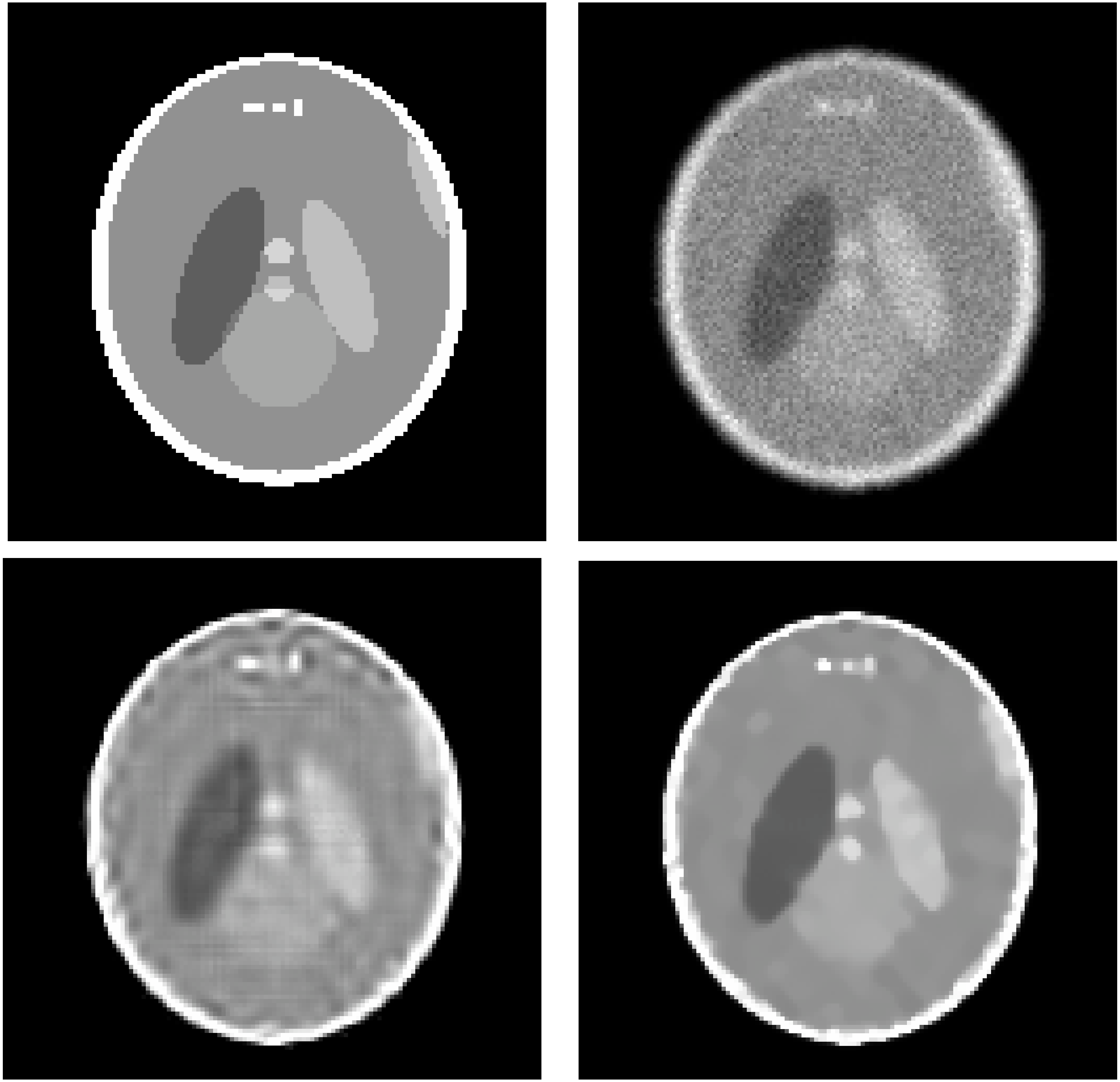
min	Complexity_penalty $(\bar{s})$,	subject to	$\bar{s}$	C

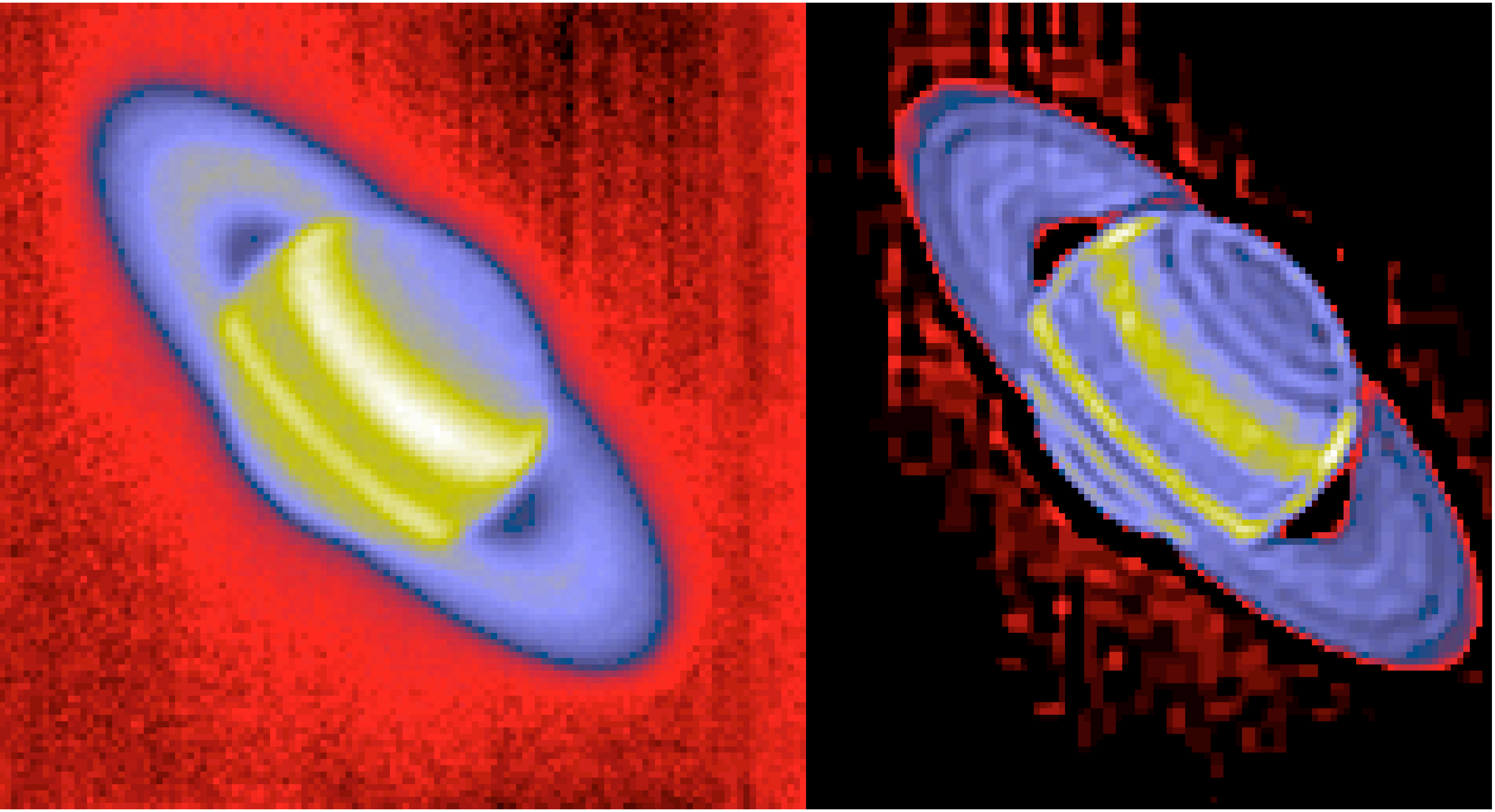
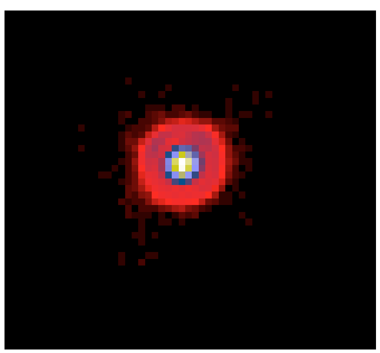
Where C is the set of vectors which obey the linear constraints:

$\tilde{s} > 0,$ positivity constraint

$\left|\left(T_k \tilde{s} - T_k P * s\right)_l\right| \le e, \quad if \left(T_k s\right)_l$ is significant

The second constraint guarantees that the reconstruction will take into account any pattern which is detected by any of the K transforms.



Morphological Component Analysis (MCA)

Given a signal s, we assume that it is the result of a sparse linear combination of atoms from a known dictionary D.

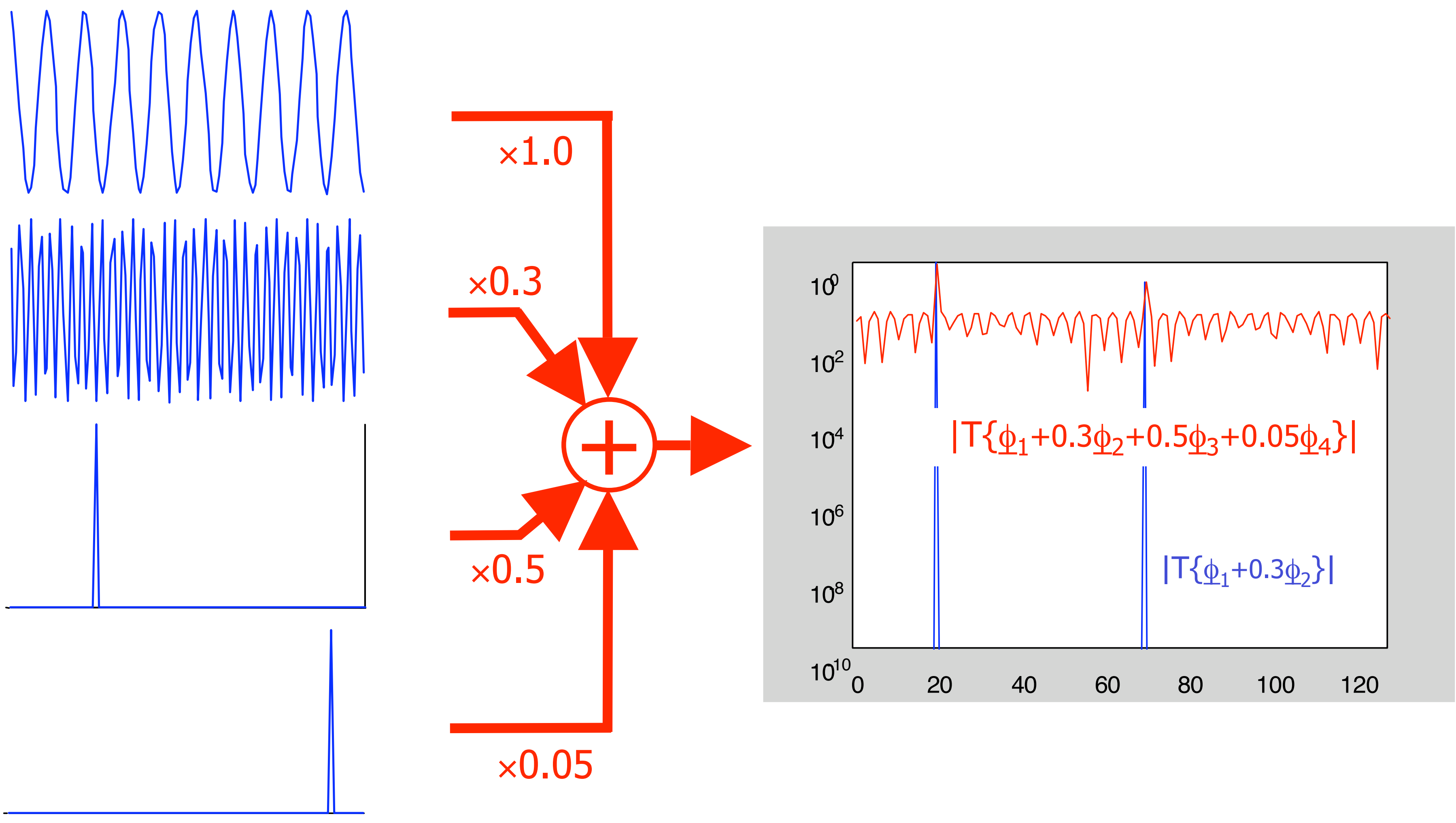
A dictionary D is defined as a collection of waveforms $\left(\phi_\gamma\right)_{\gamma\in\Gamma}$, and the goal is to obtain a representation of a signal s with a linear combination of a small number of basis such that:

$$s = \sum_\gamma \alpha_\gamma \phi_\gamma$$

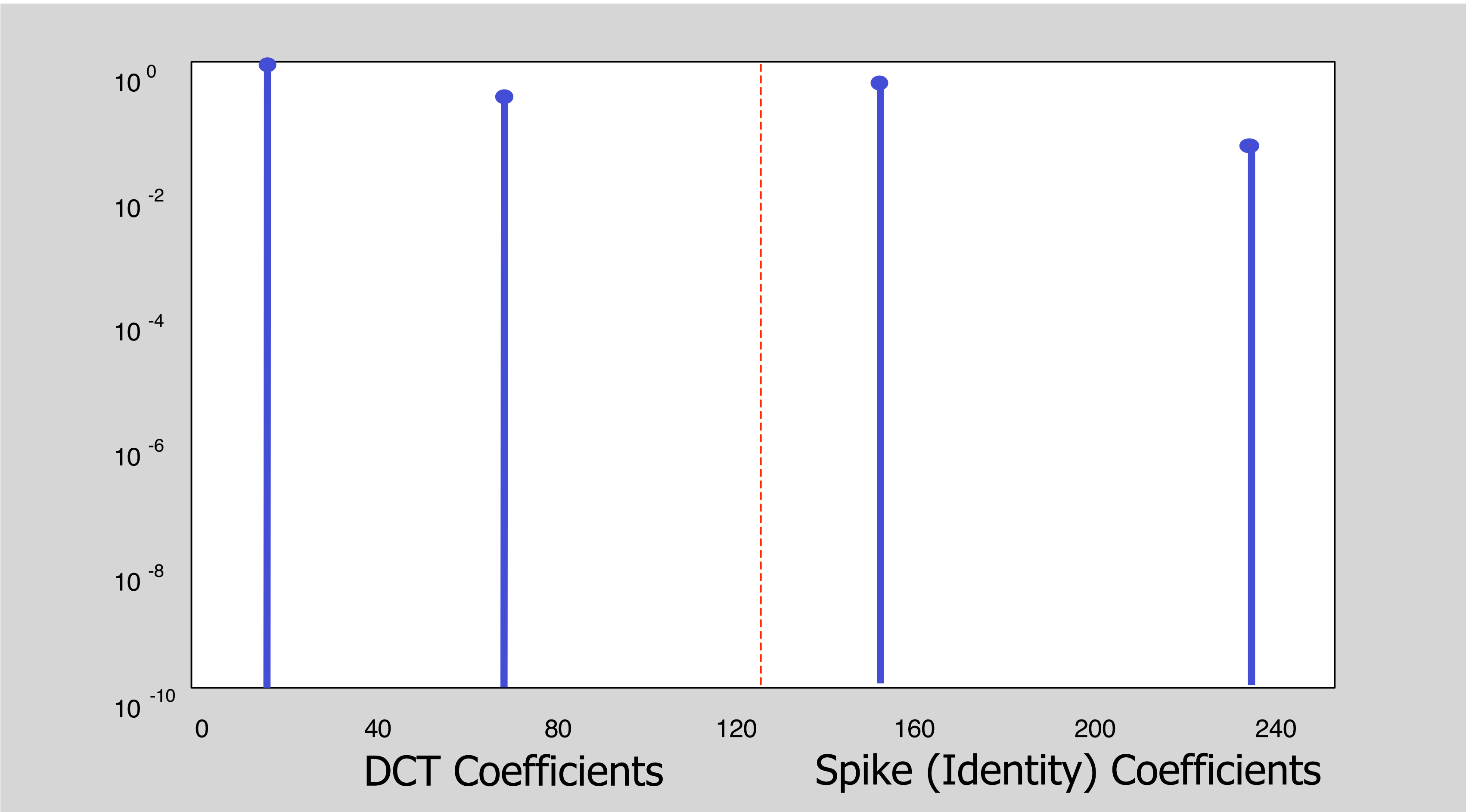
Or an approximate decomposition:

$$s = \sum_\gamma \alpha_\gamma \phi_\gamma + R$$

Example – Composed Signal



Example – Desired Decomposition



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

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

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

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

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

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

We consider now that the dictionary is built of a set of L dictionaries related to multiscale transforms, such wavelets, ridgelet, or curvelets.

Considering L transforms, and α_k the coefficients relative to the kth transform:

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

Noting $T_1, \ldots T_L$ the L transform operators, we have:

$$\alpha_k = T_k s_k, \qquad s_k = T_k^{-1} \alpha_k, \qquad s = \sum_{k=1}^{L} s_k$$

A solution α is obtained by minimizing a functional of the form:

$$J(\alpha) = \left\| s - \sum_{k=1}^{L} T_k^{-1} \alpha_k \right\|_2^2 + \|\alpha\|_p$$

Different Problem Formulation

$$J(s_1,\ldots,s_L) = \left\| s - \sum_{k=1}^{L} s_k \right\|_2^2 + \lambda \sum_{k=1}^{L} \left\| T_k s_k \right\|_p$$

.We do not need to keep all transforms in memory.

.We can easily add some constraints on a given component

An efficient algorithm is the Block-Coordinate Relaxation Algorithm (Sardy, Bruce and Tseng, 1998):

. Initialize all s_k to zero

. Iterate j=1,...,M

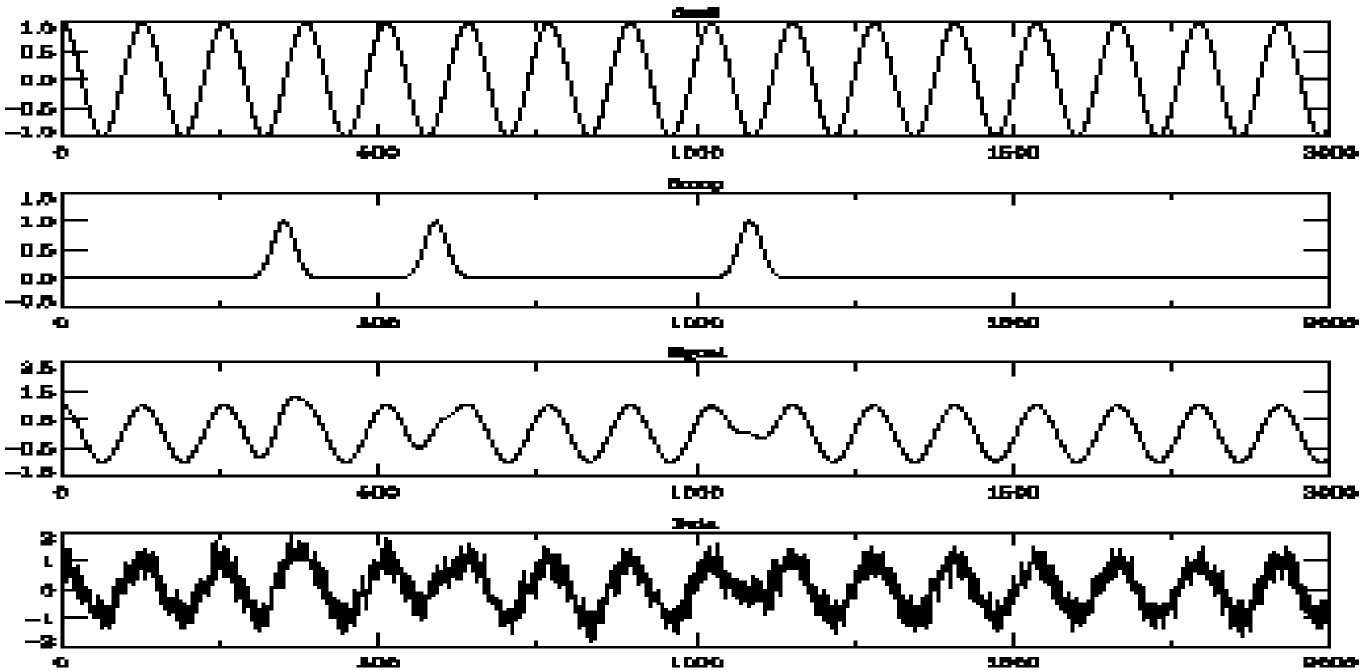
- Iterate k=1,..,L

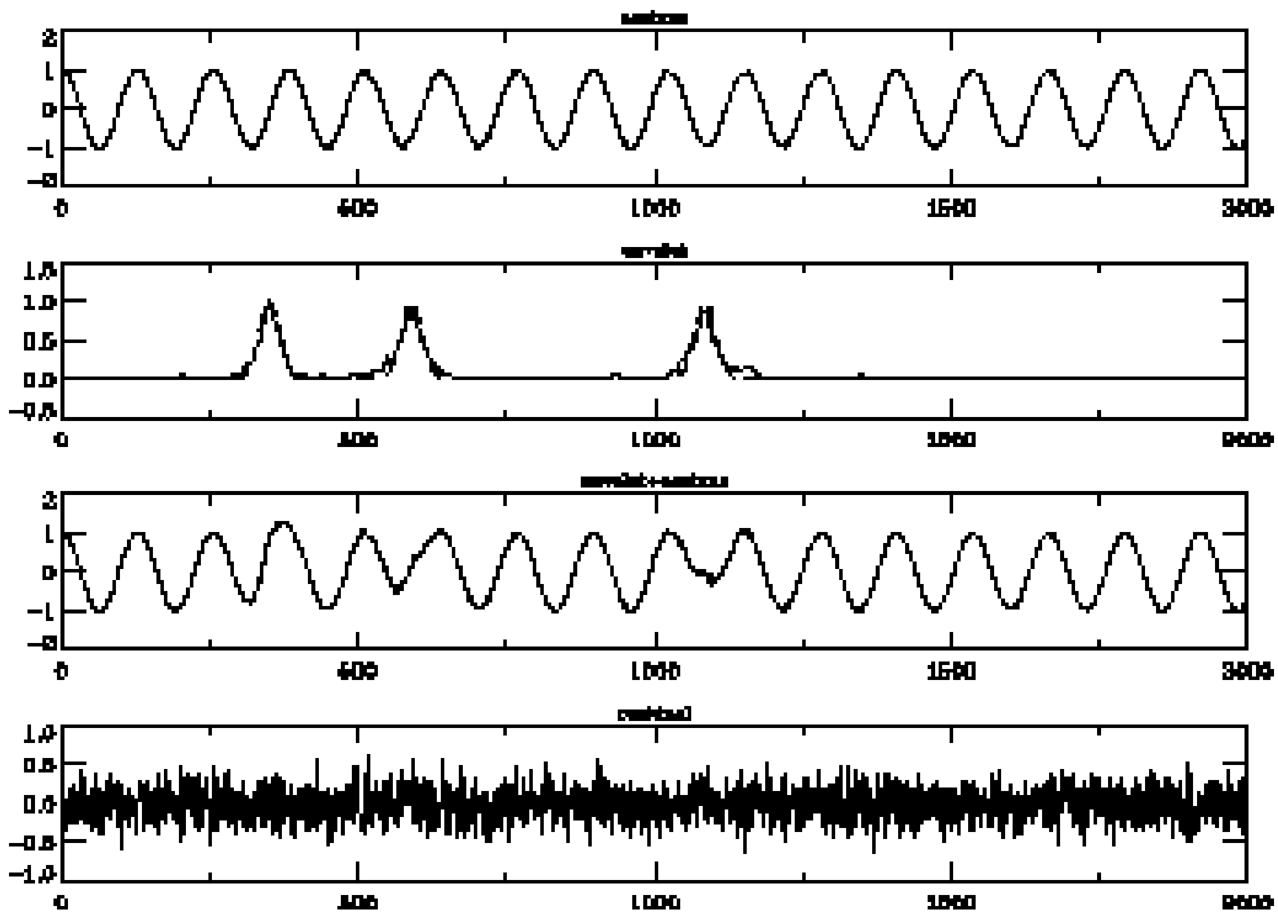
  - Update the kth part of the current solution by fixing all other parts and minimizing:

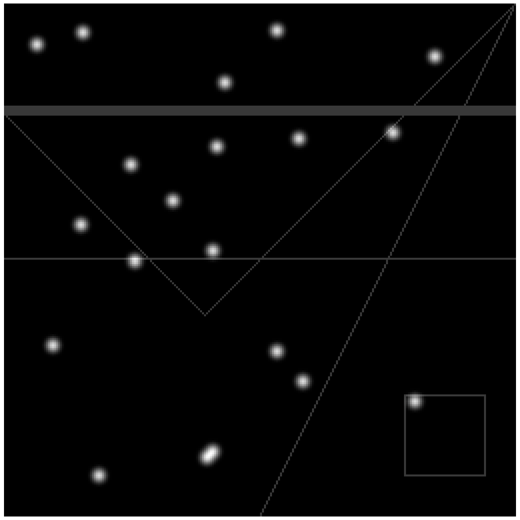
$$J(s_k) = \left\| s - \sum_{i=1, i\neq k}^{L} s_i - s_k \right\|_2^2 + \lambda \left\| T_k s_k \right\|_1$$

Which is obtained by a simple soft thresholding of :

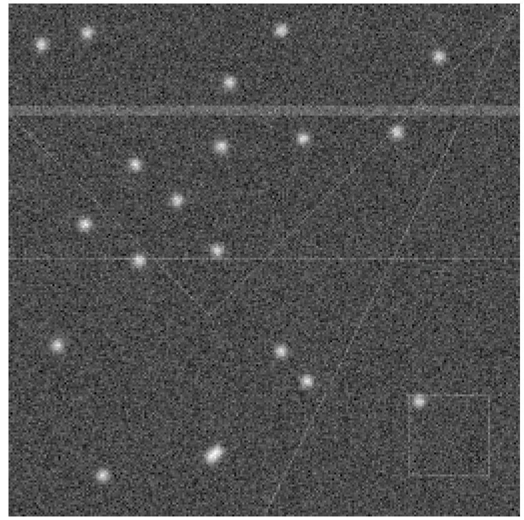
$$s_r = s - \sum_{i=1, i\neq k}^{L} s_i$$



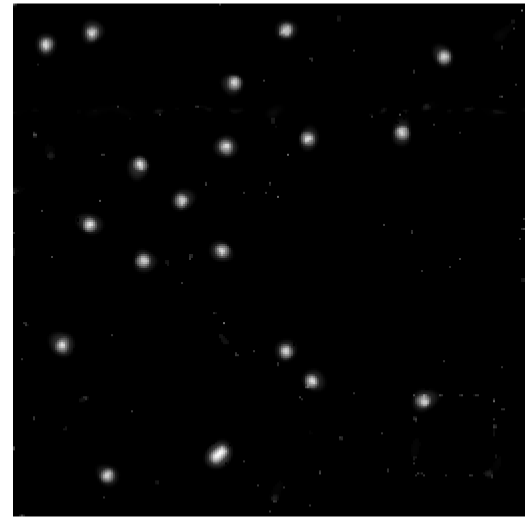




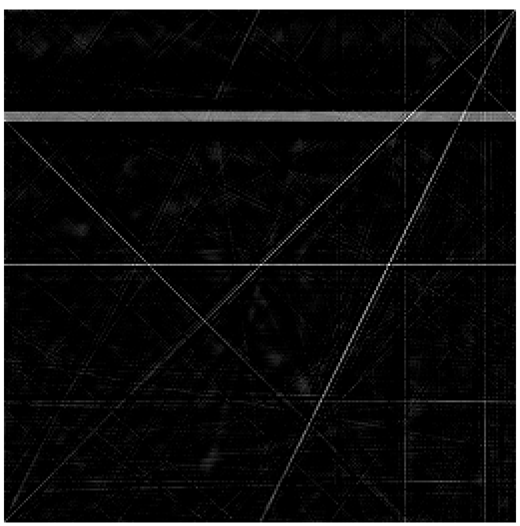
a) Simulated image (Gaussians+lines)



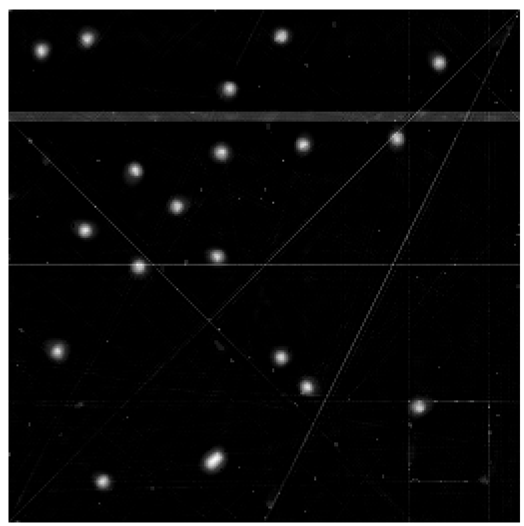
b) Simulated image + noise



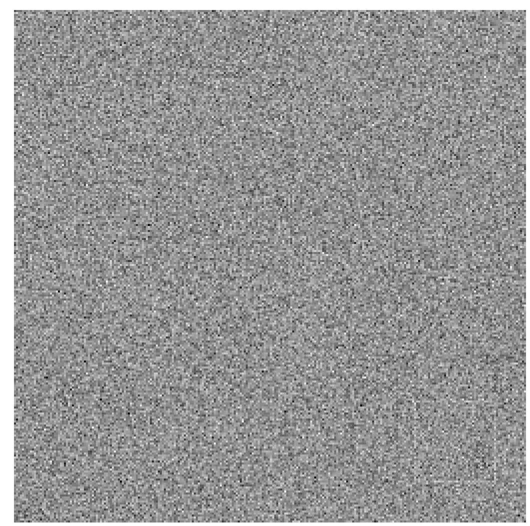
c) A trous algorithm



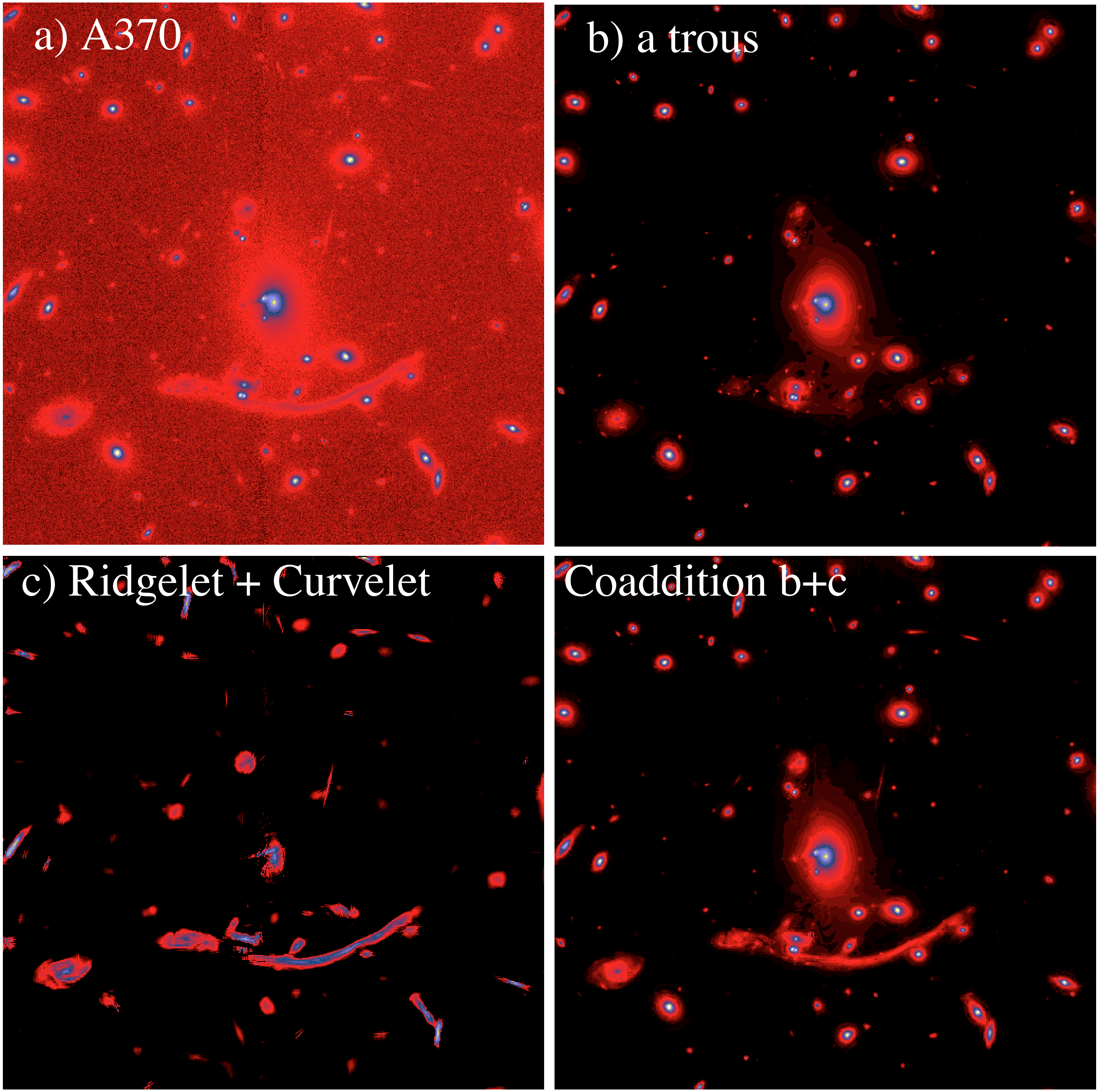
d) Curvelet transform

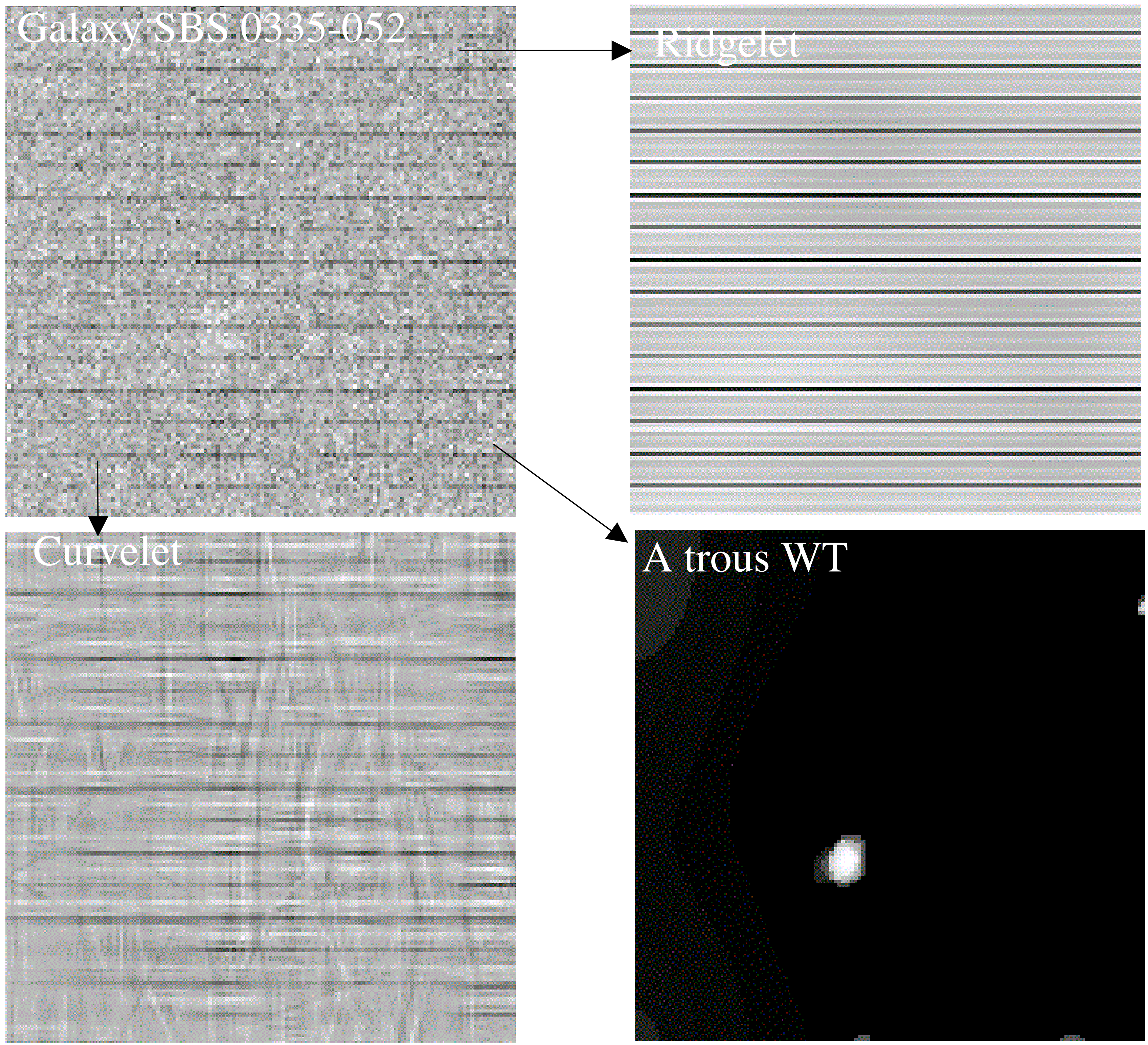


e) coaddition c+d

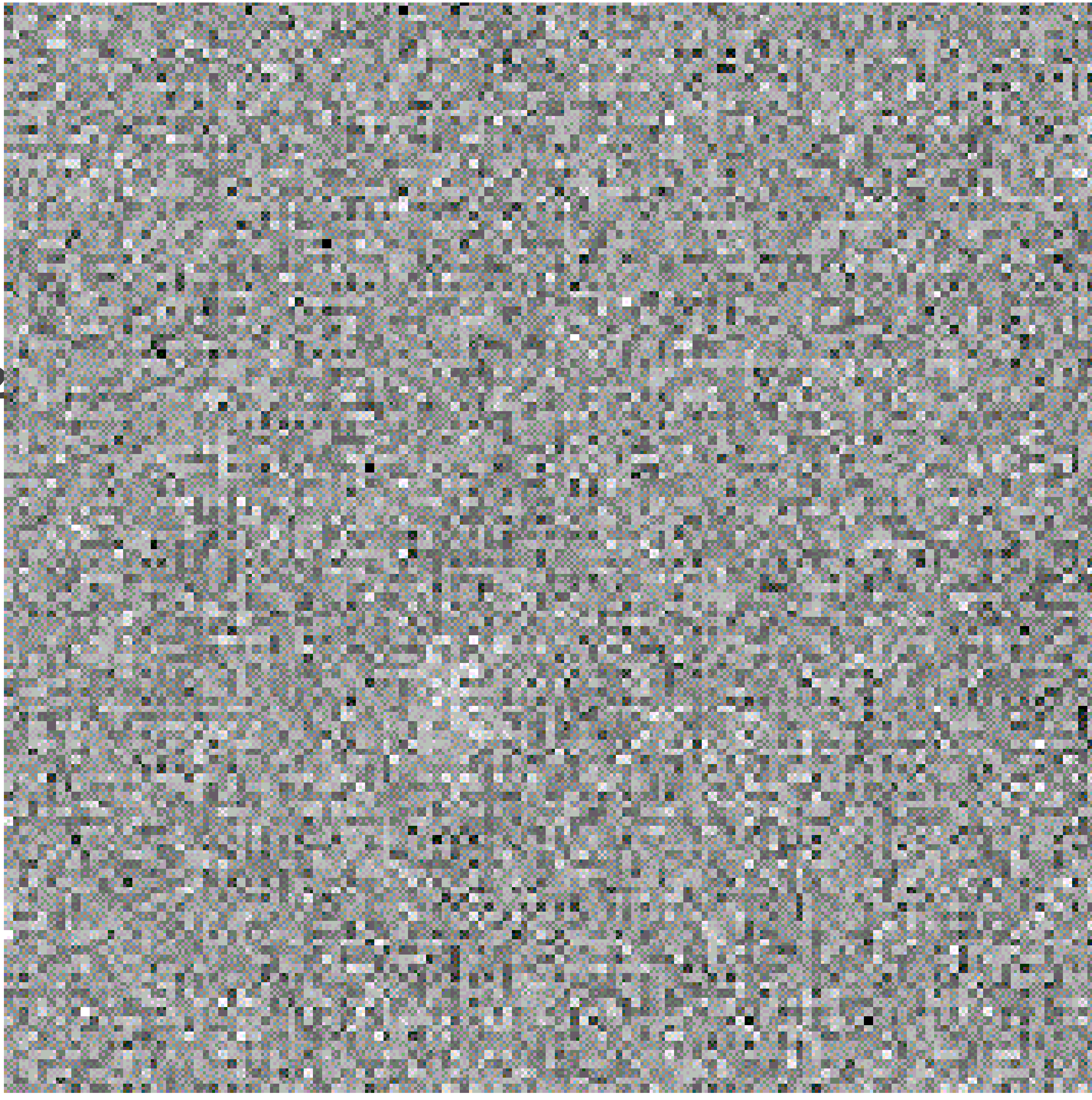


f) residual = e-b


a) A370
b) a trous
c) Ridgelet + Curvelet
Coaddition b+c


Galaxy SBS 0335-052
Ridgelet
Curvelet
A trous WT

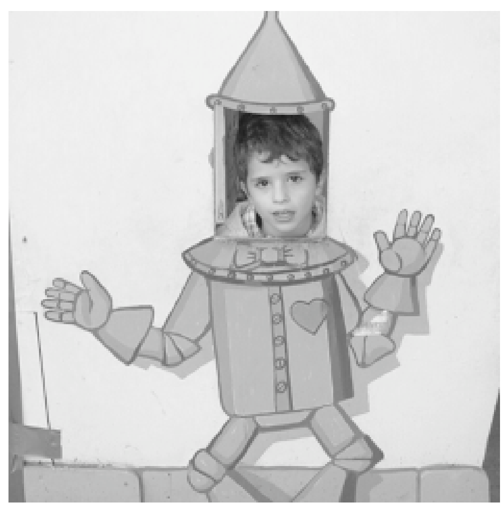
Galaxy SBS 0335-052
10 micron
GEMINI-OSCIR



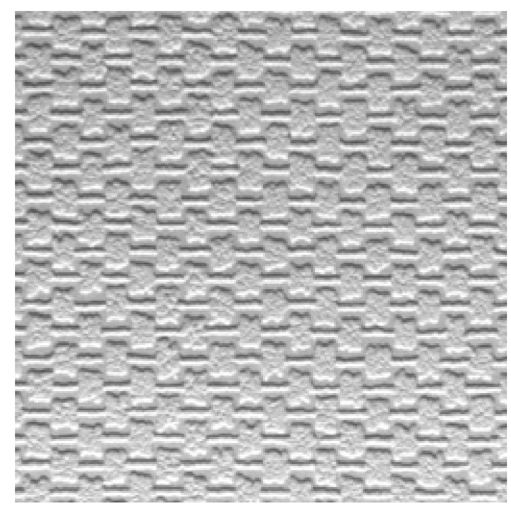
Separation of Texture from Piecewise Smooth Content

<u>The separation task</u>: decomposition of an image into a texture and a natural (piecewise smooth) scene part.

=



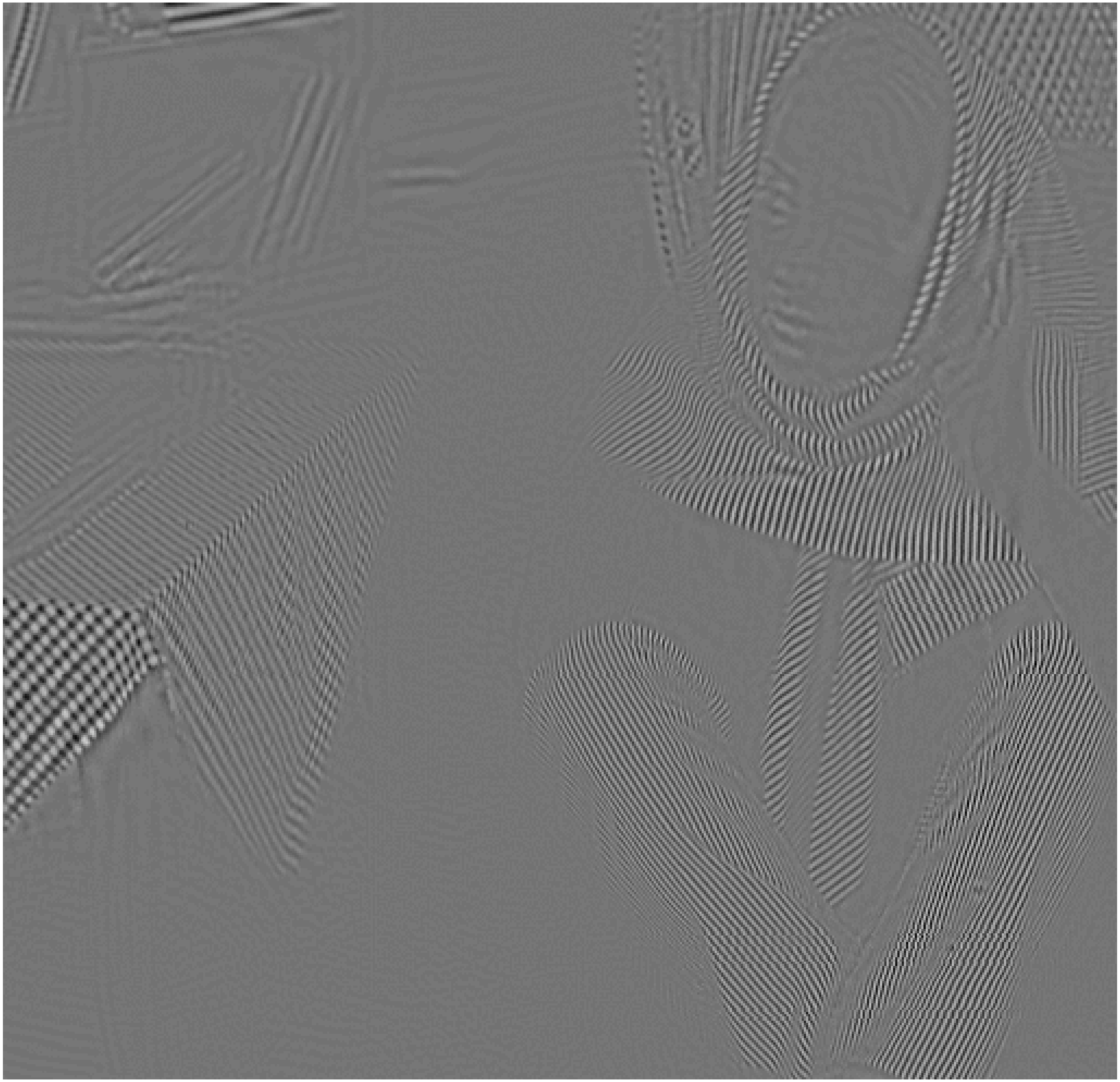
+

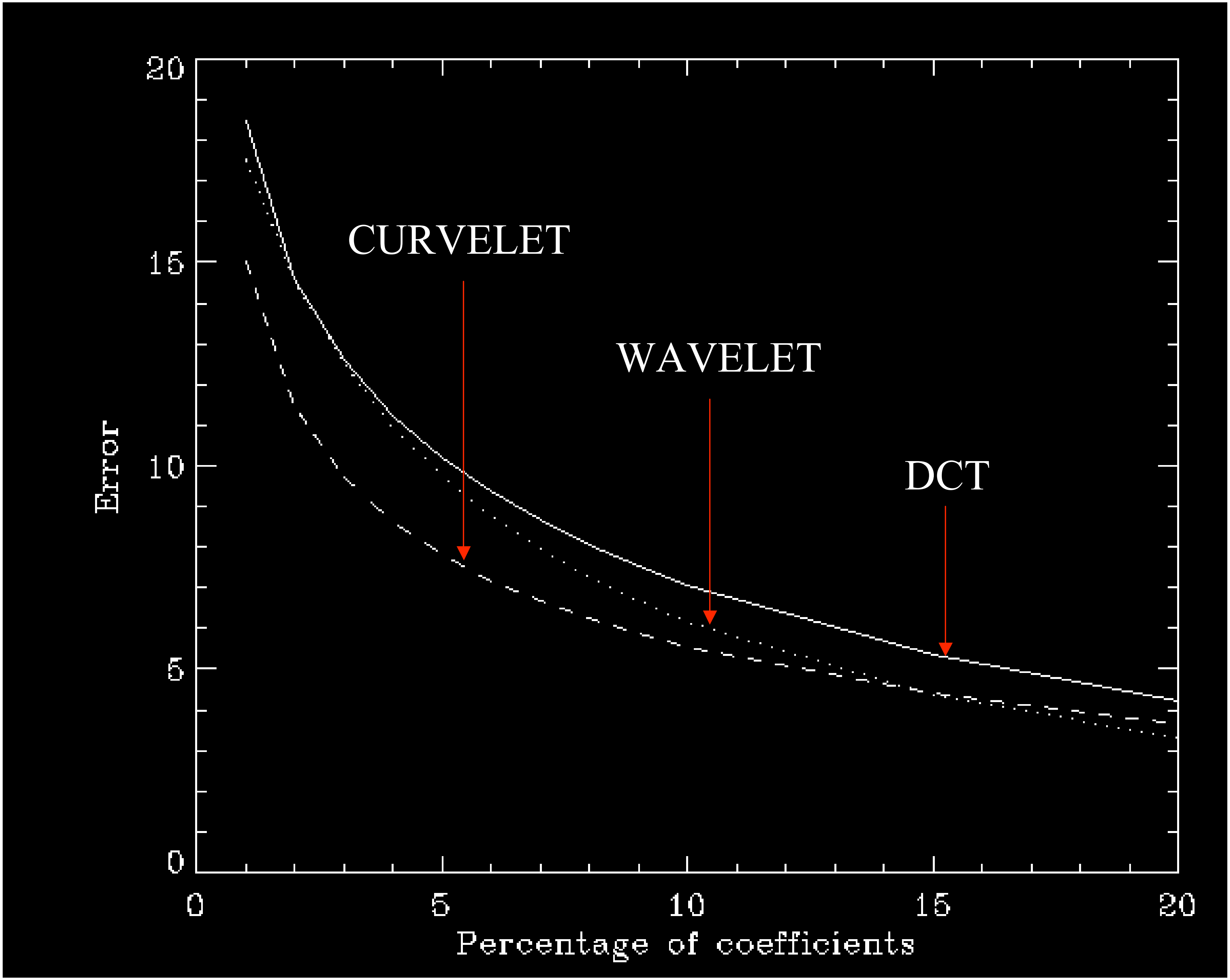


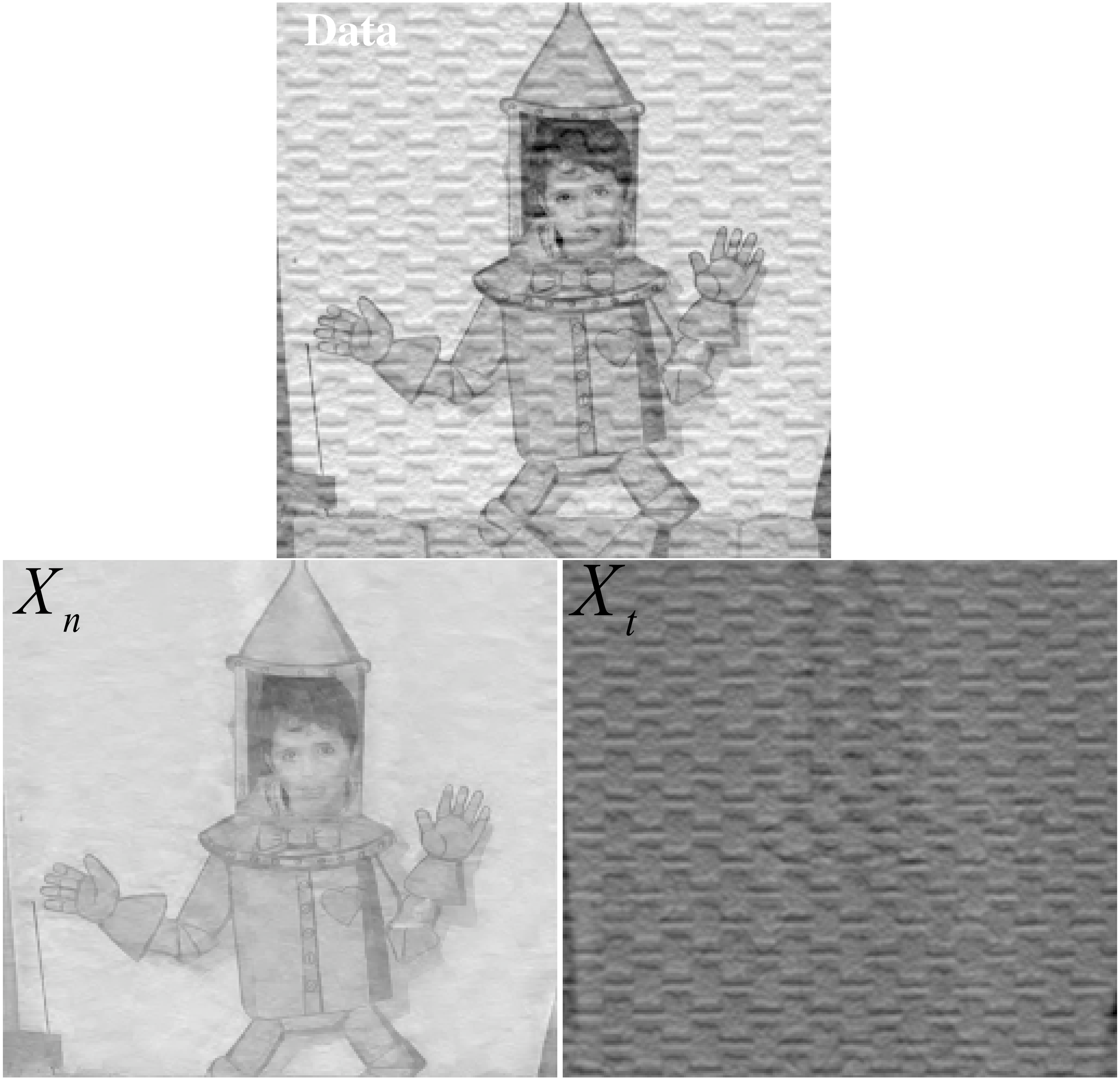
Numerical Consideration

The DCT is denoted $\mathcal{D}$ and its inverse by $\mathcal{D}^{-1}$ (with a clear abuse of notations). The curvelet transform is denoted it by $\mathcal{C}$ and its inverse by $\mathcal{C}^{-1}$. We have two unknowns - $\underline{X}_t$ and $\underline{X}_n$ - the texture and the piecewise smooth images. The optimization problem to be solved is

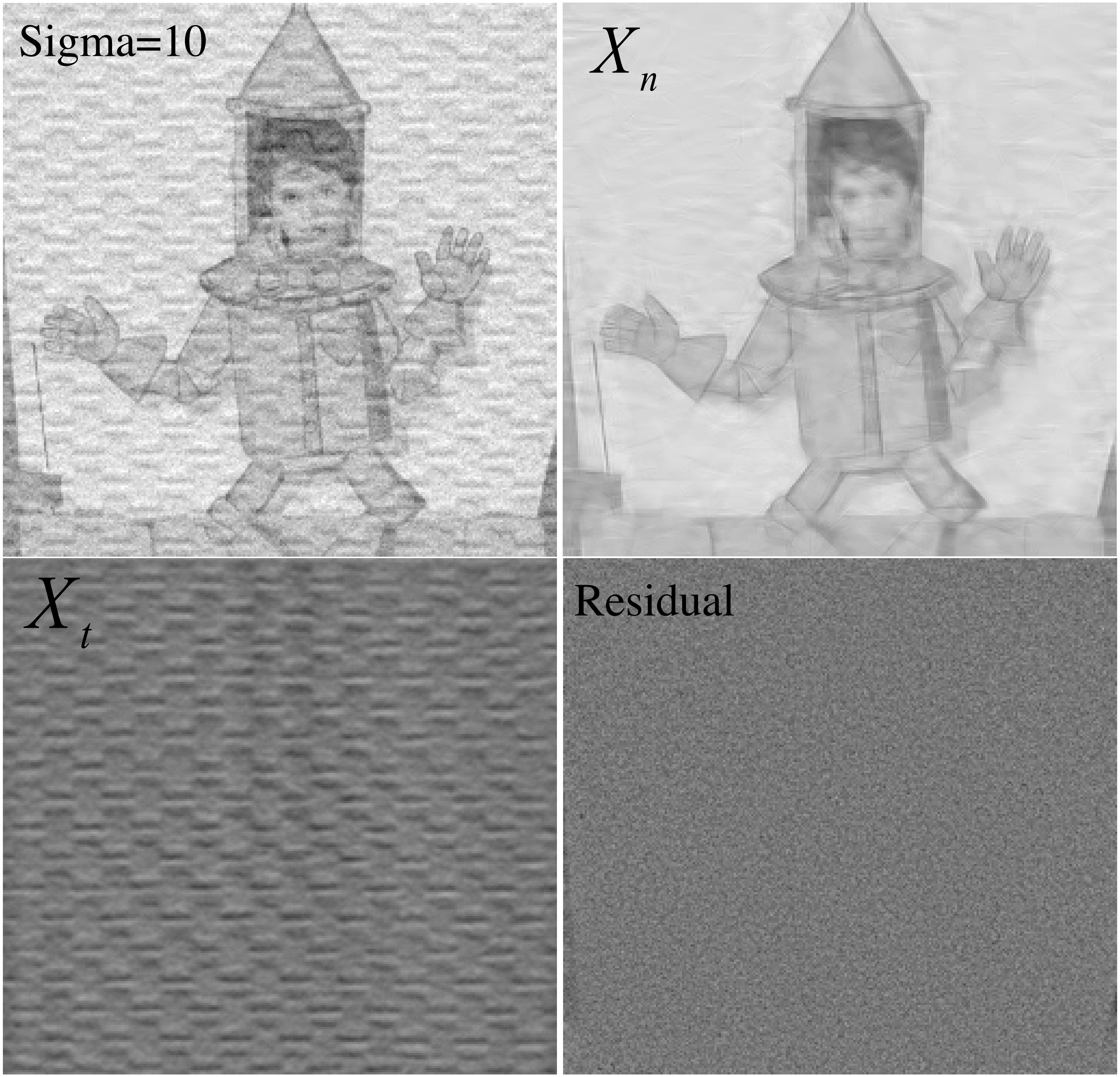
$$\min_{\{\underline{X}_t,\, \underline{X}_n\}} \quad \|\mathcal{D}\underline{X}_t\|_1 + \|\mathcal{C}\underline{X}_n\|_1 + \lambda \left\|\underline{X} - \underline{X}_t - \underline{X}_n\right\|_2^2 + \gamma TV\left\{\underline{X}_n\right\}.$$




CURVELET
WAVELET
DCT
Error
Percentage of coefficients
0
5
10
15
20


Data
X_n
X_t

Edge Detection


Sigma=10
X_n
X_t
Residual

Interpolation of Missing Data

$$J(s_1,\ldots,s_L) = \left\| M(s - \sum_{k=1}^{L} s_k) \right\|_2^2 + \lambda \sum_{k=1}^{L} \left\| T_k s_k \right\|_p$$

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

If the data are composed of a piecewise smooth component + texture

$$J(X_t, X_n) = \left\| M(X - X_t - X_n) \right\|_2^2 + \lambda (\left\| \mathbf{C} X_n \right\|_1 + \left\| \mathbf{D} X_t \right\|_1) + \gamma \, \mathrm{TV}(X_n)$$

. Initialize all s_k to zero

. Iterate j=1,...,M

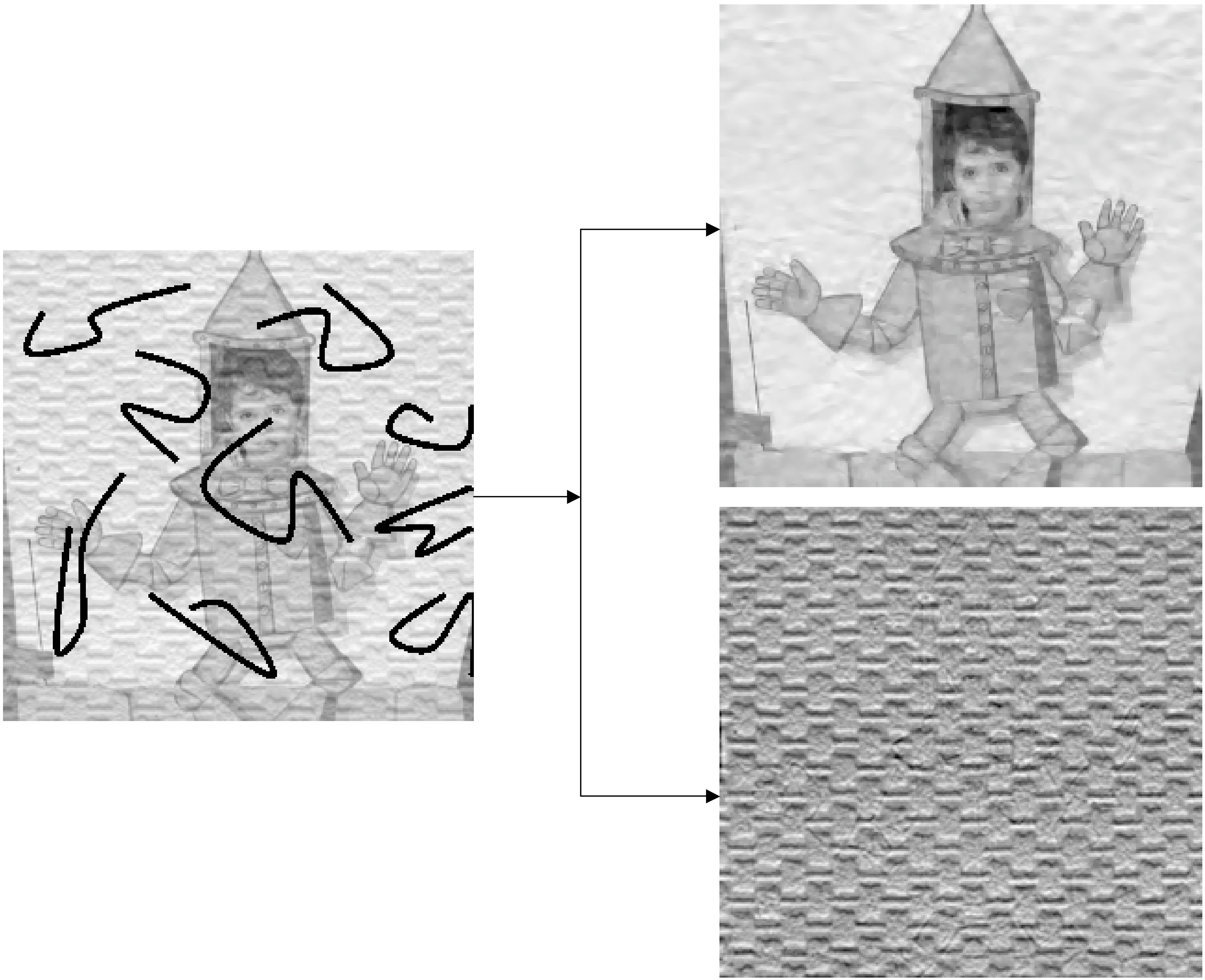
- Iterate k=1,..,L

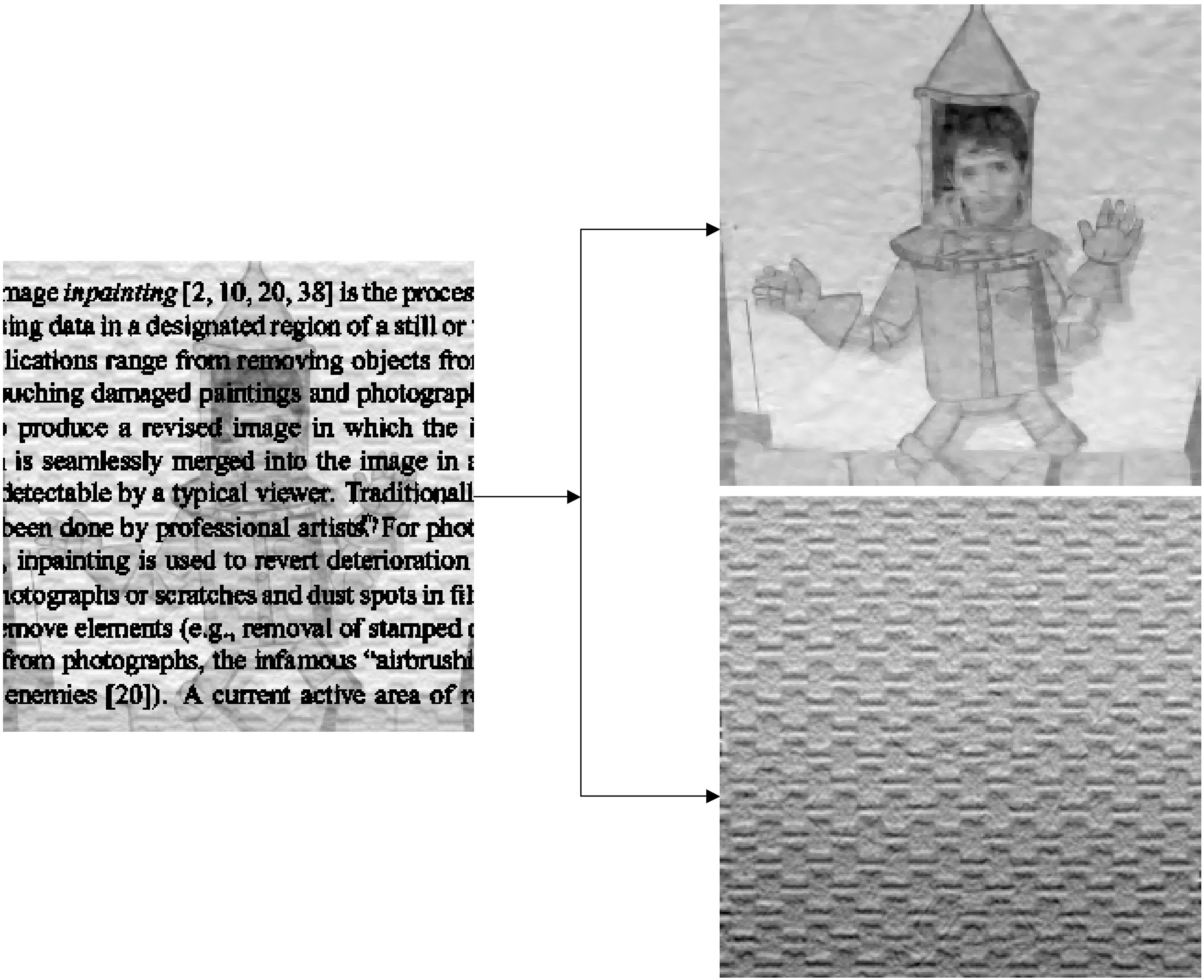
- Update the kth part of the current solution by fixing all other parts and minimizing:

$$J(s_k) = \left\| M(s - \sum_{i=1, i \neq k}^{L} s_i - s_k) \right\|_2^2 + \lambda \| T_k s_k \|_1$$

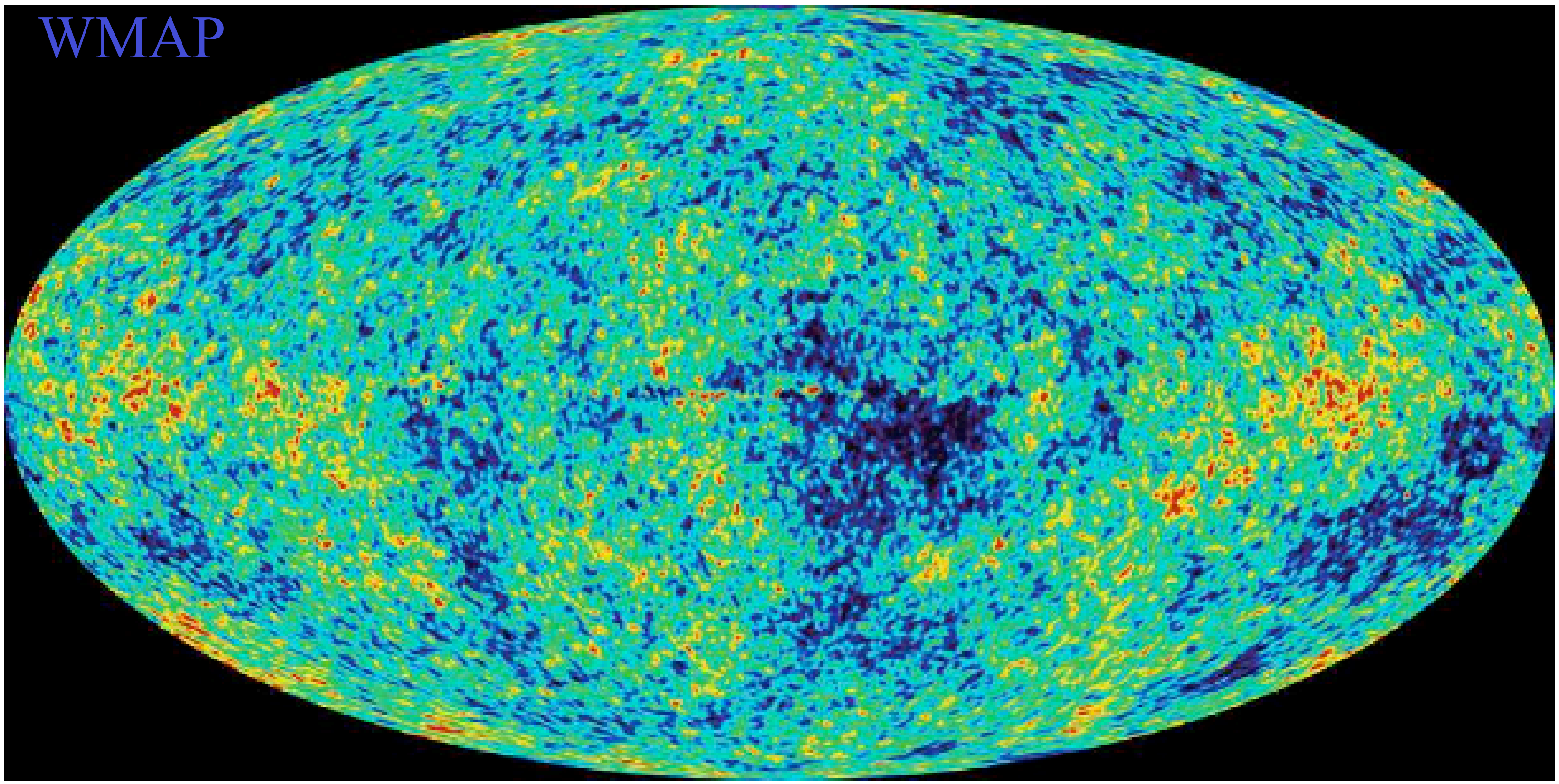
Which is obtained by a simple soft thresholding of :

$$s_r = M(s - \sum_{i=1, i \neq k}^{L} s_i)$$

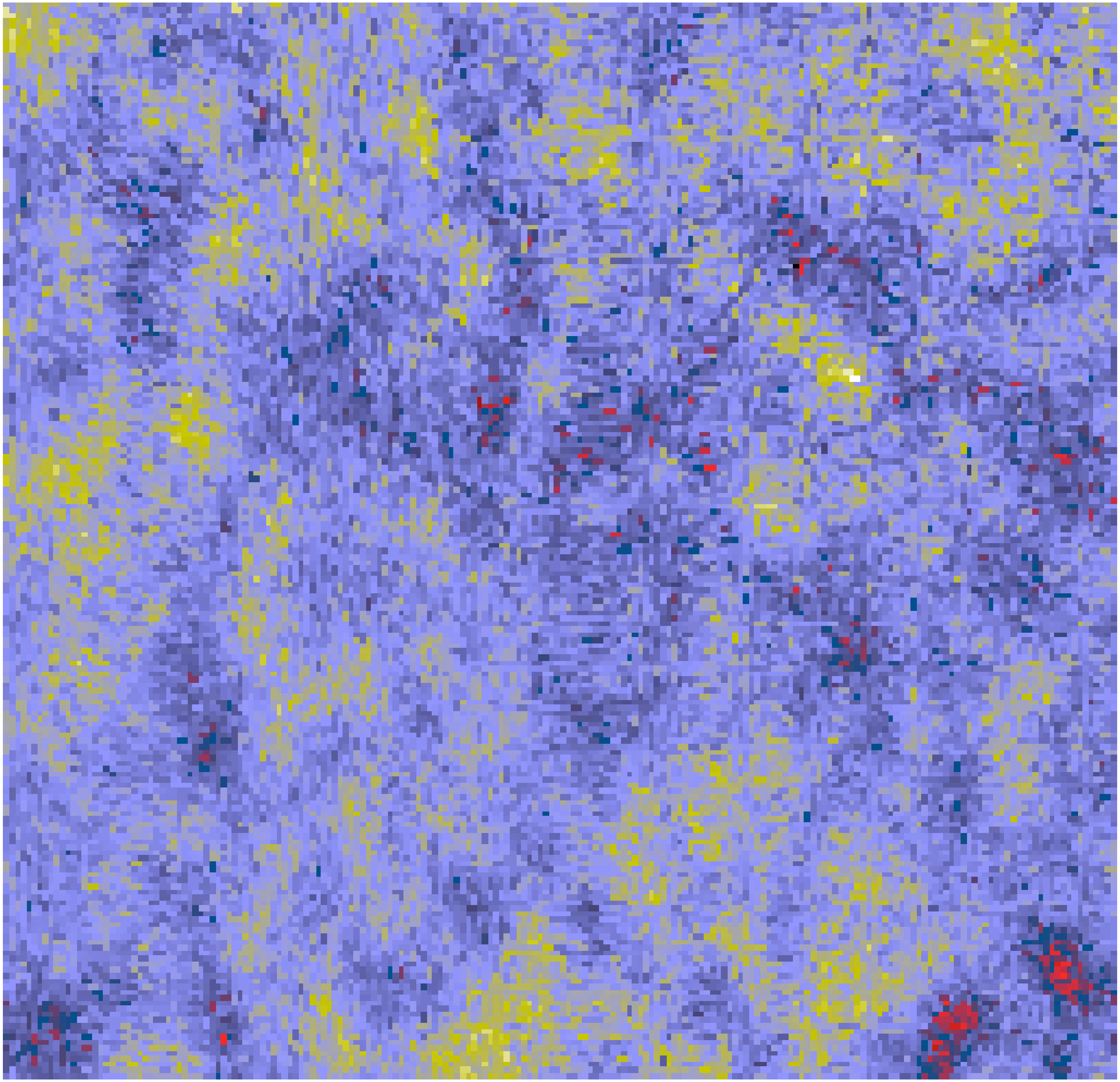


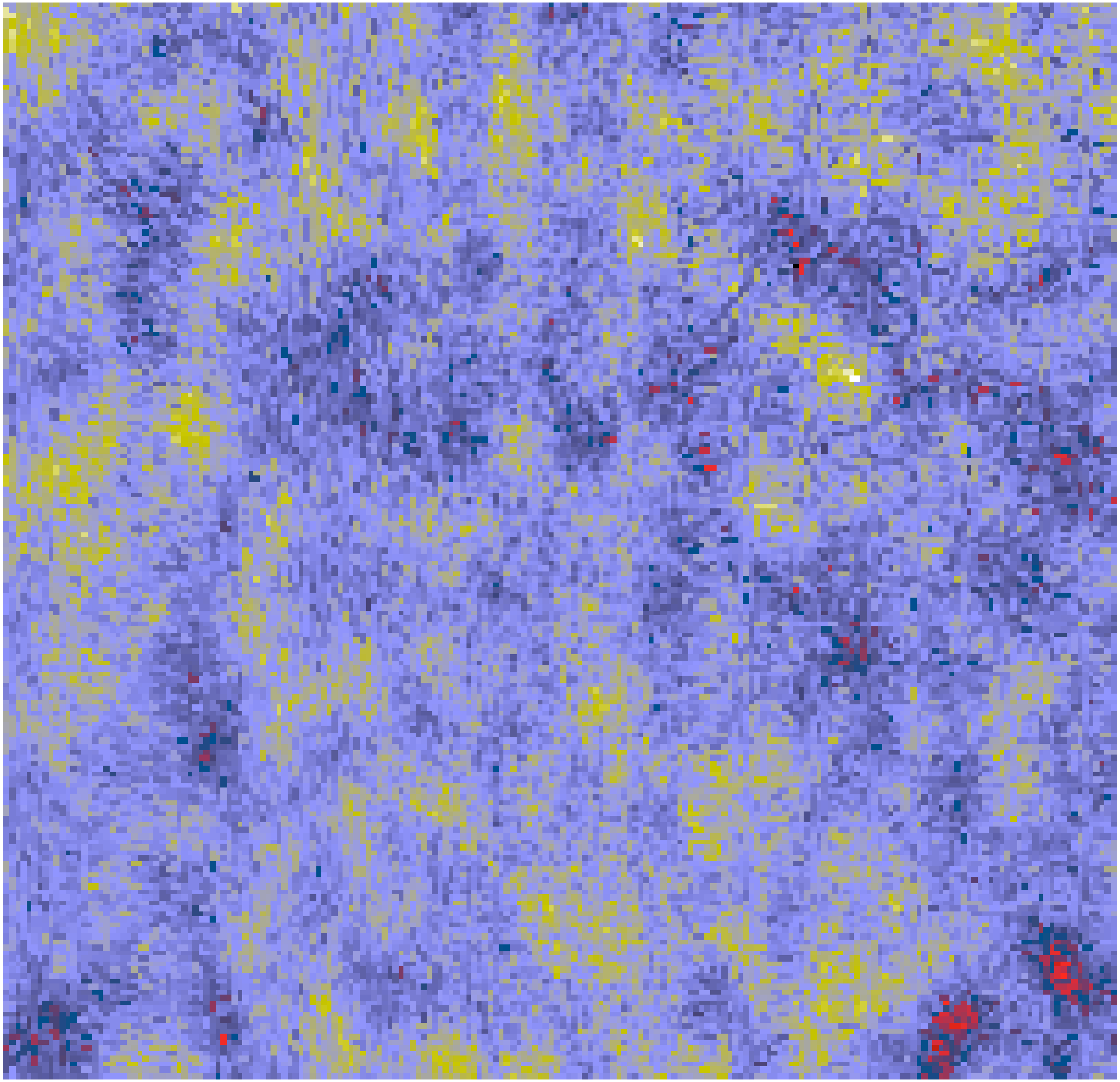

nage inpainting [2, 10, 20, 38] is the proces
ing data in a designated region of a still or
lications range from removing objects fro
uching damaged paintings and photograp
produce a revised image in which the
is seamlessly merged into the image in
detectable by a typical viewer. Traditionall
been done by professional artists. For phot
, inpainting is used to revert deterioration
otographs or scratches and dust spots in fil
move elements (e.g., removal of stamped
from photographs, the infamous "airbrushi
enemies [20]). A current active area of r

Application in Cosmology

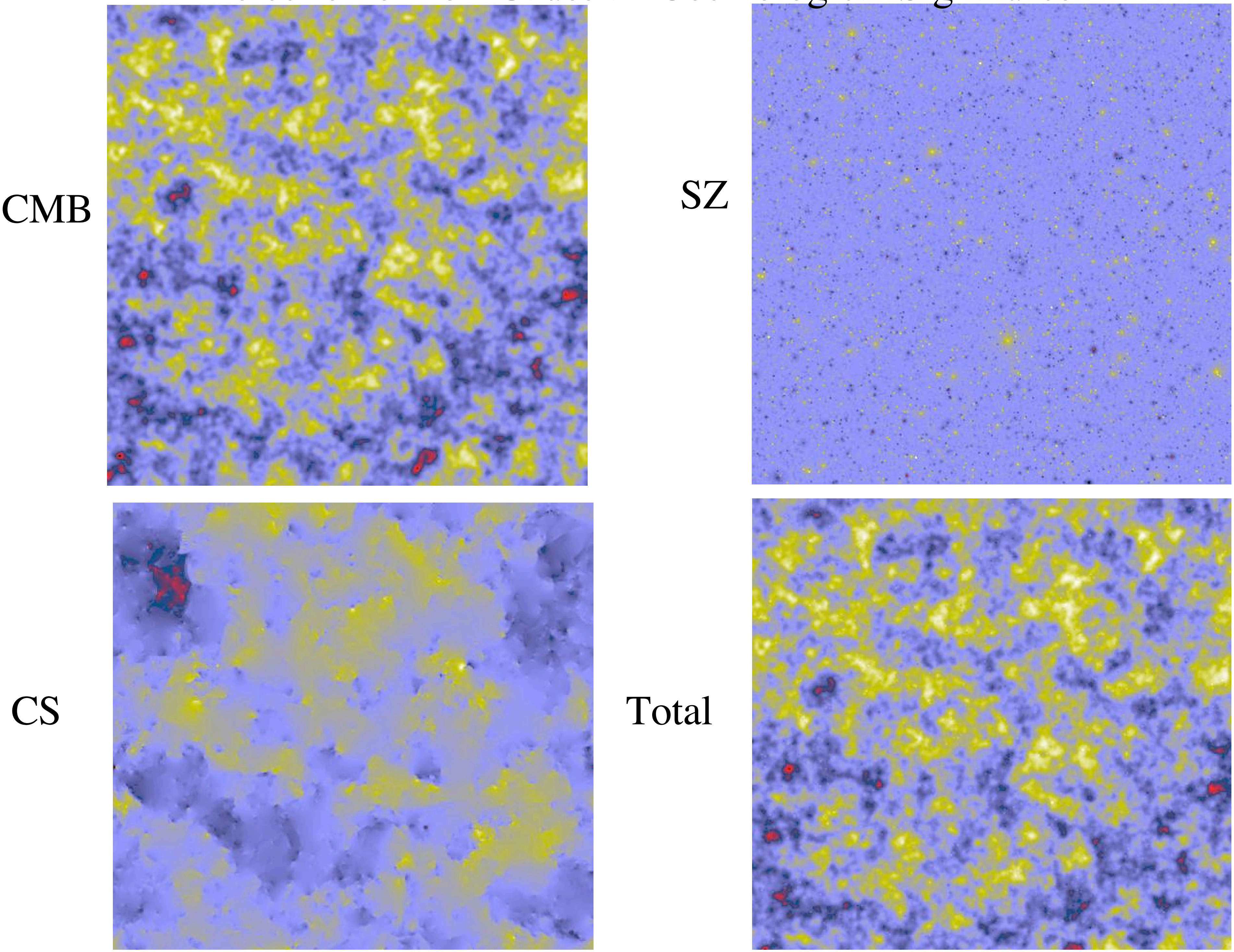


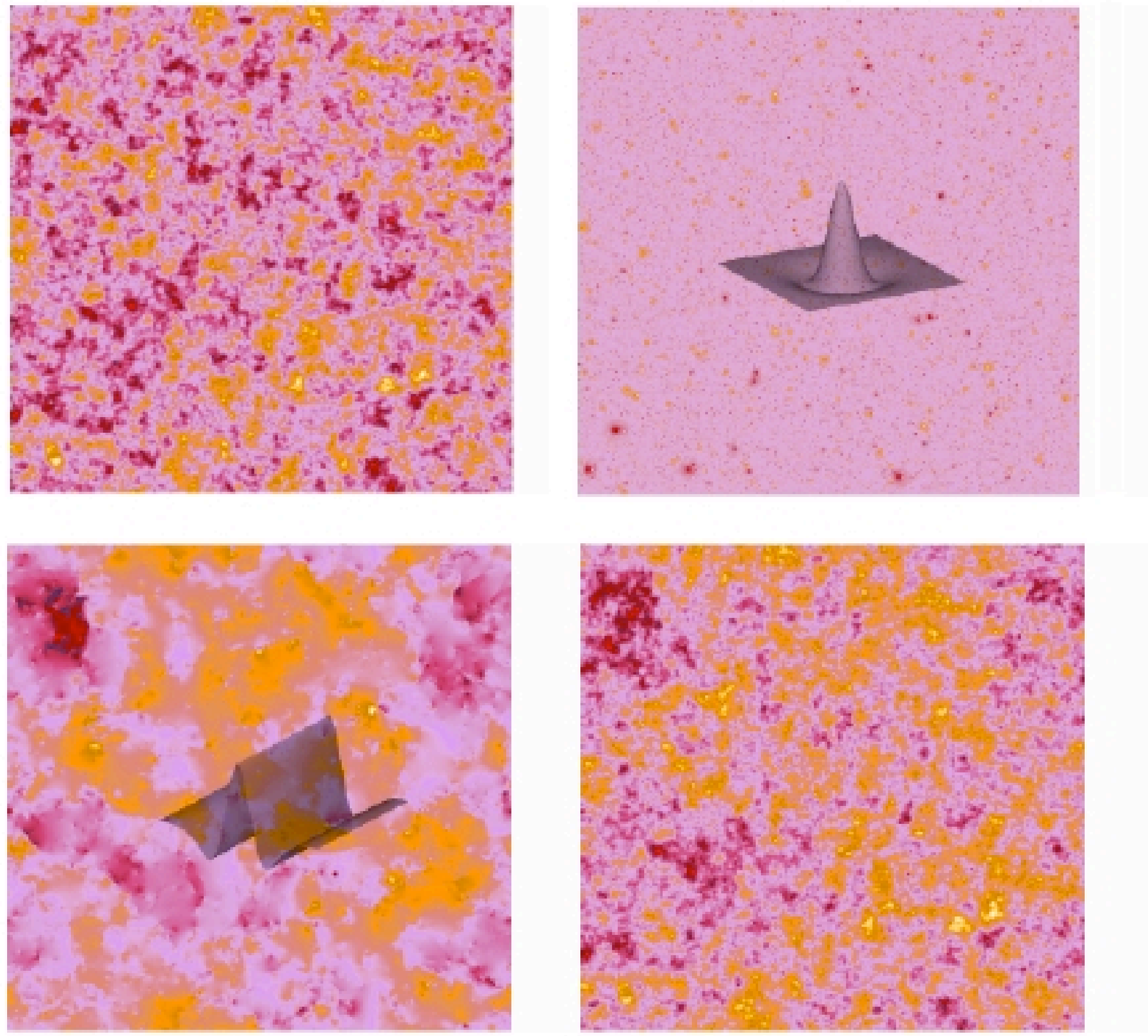
The Cosmic Microwave Background (CMB) is a relic radiation (with a temperature equals to 2.726 Kelvin) emitted 13 billion years ago when the Universe was about 370000 years old.





Detection of non-Gaussian Cosmological Signatures





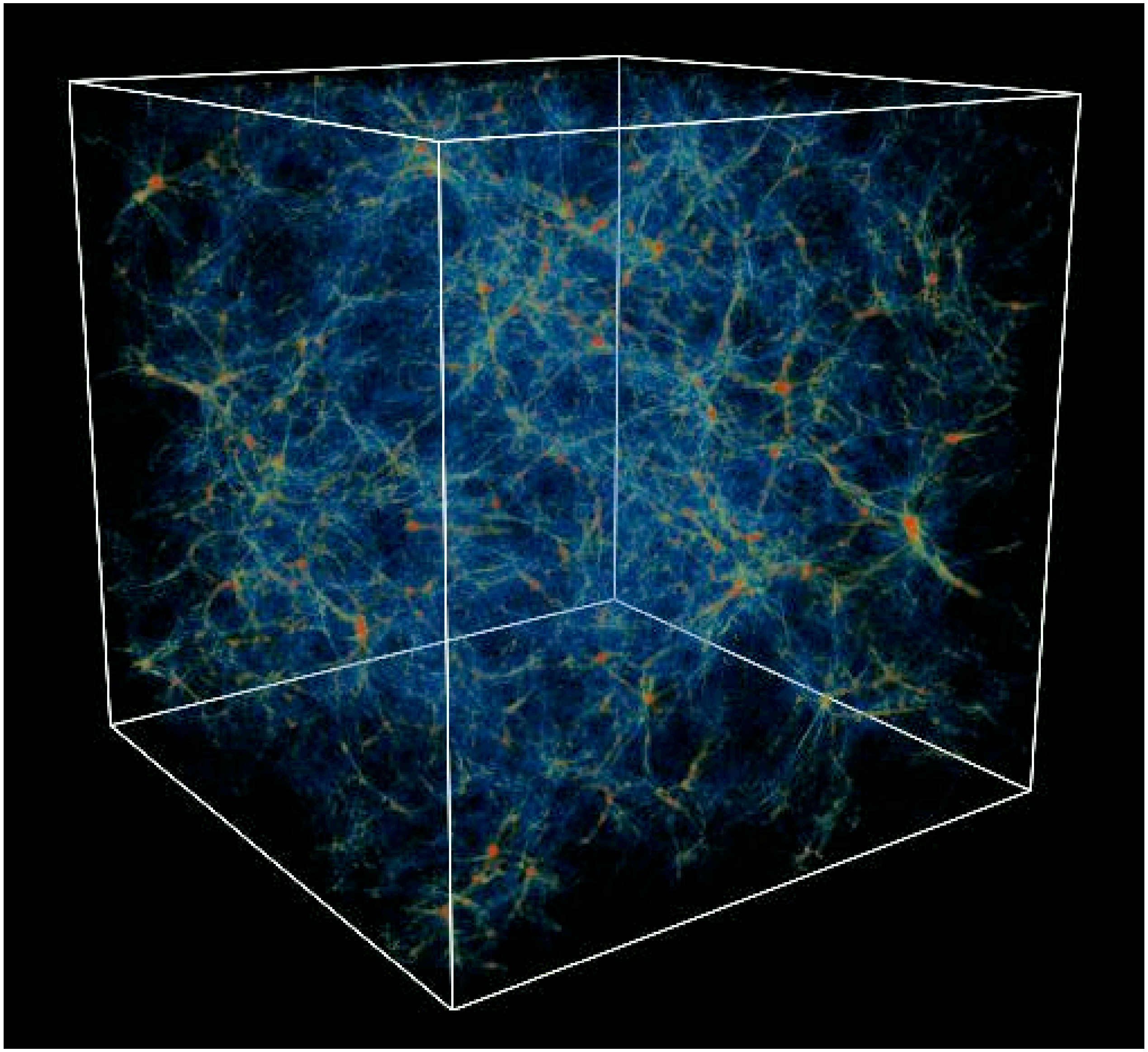
Results

- **Curvelets are NOT sensible to KSZ and sensitive to cosmic string**

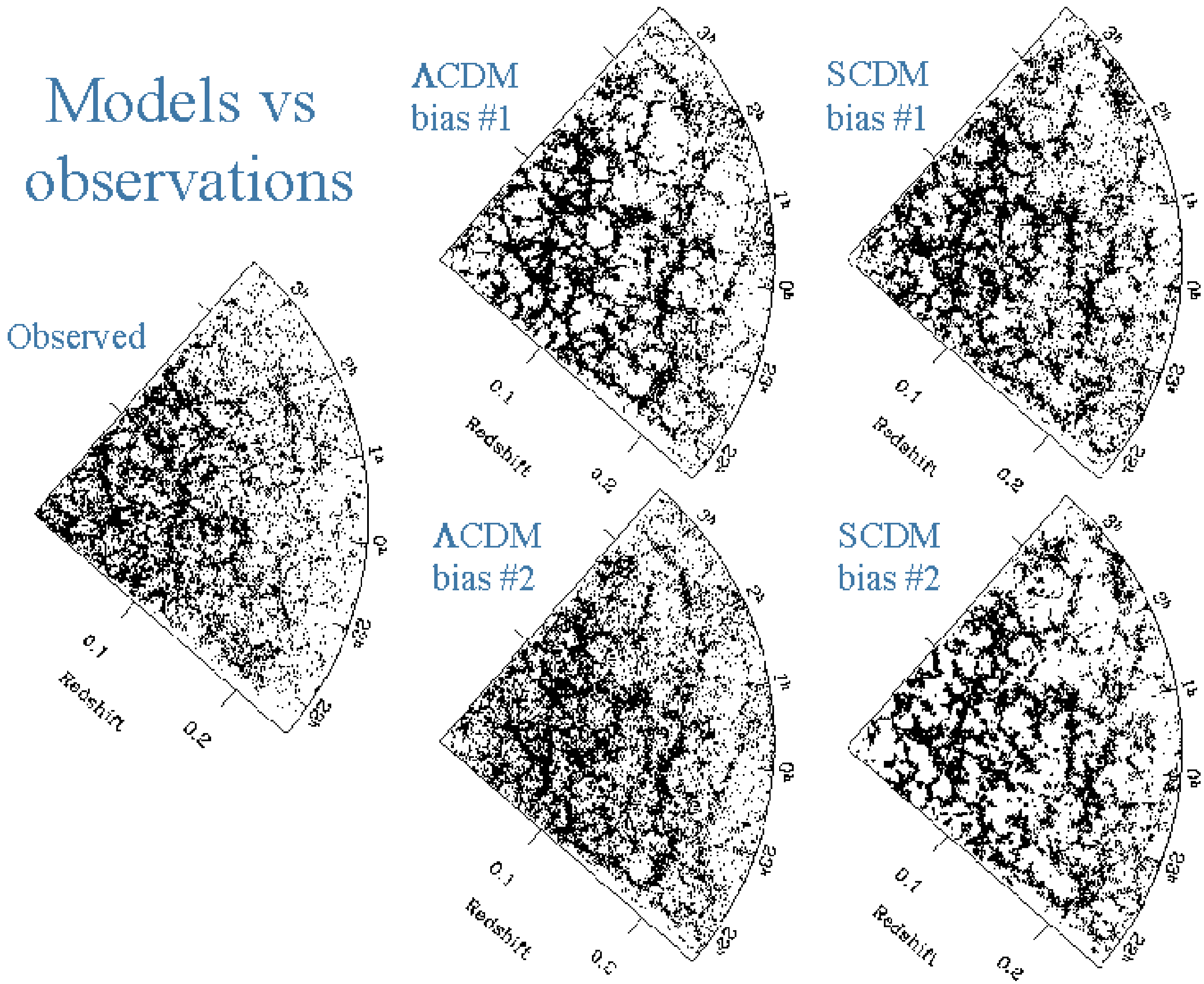
	Bi-orthogonal WT	Ridgelet	Curvelet
CMB+KSZ	1106.	0.1	10.12
CMB+CS	1813.	5.7	198.
CMB+CS+KSZ	1040.	5.9	165.

Detecting cosmological non-Gaussian signatures by multi-scale methods, *Astron. and Astrophys., 416, 9--17, 2004* .

Spatial distribution of the galaxies



Models vs observations

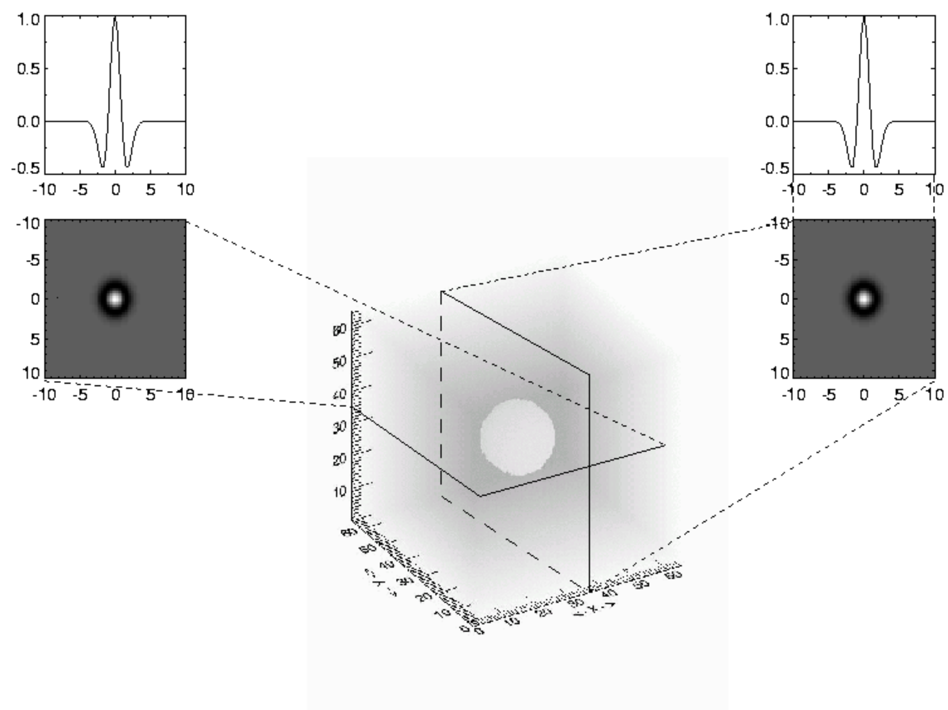


3D MULTISCALE TRANSFORMS

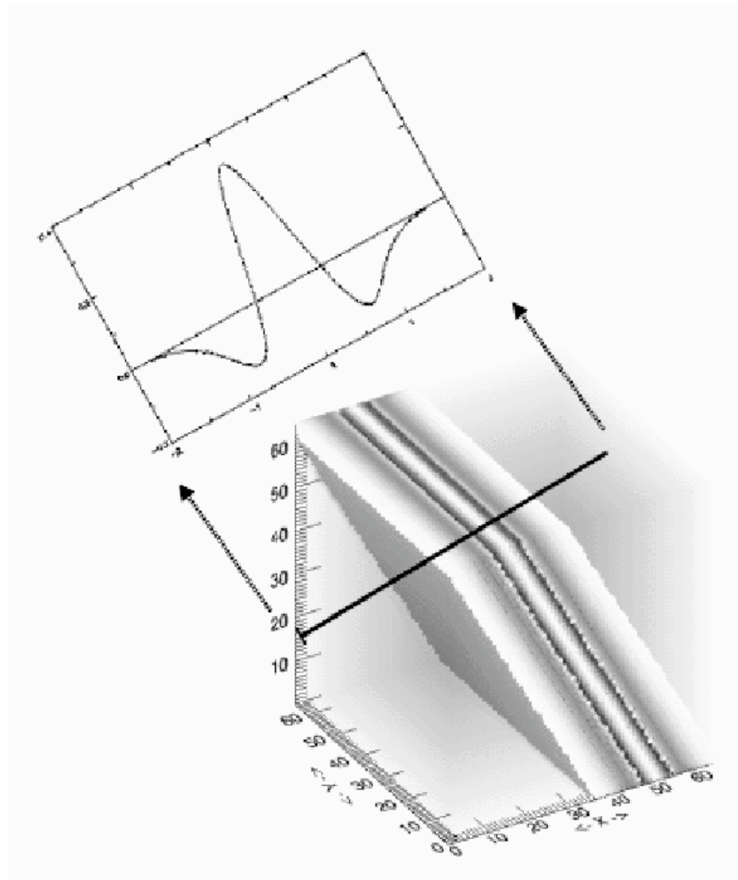
1) **3D WAVELET TRANSFORM: Isotropic Structures**
2) **3D RIDGELET TRANSFORM: Sheet like Structures**
3) **3D BEAMLET TRANSFORM: Filaments**

=> Statistical information extraction.

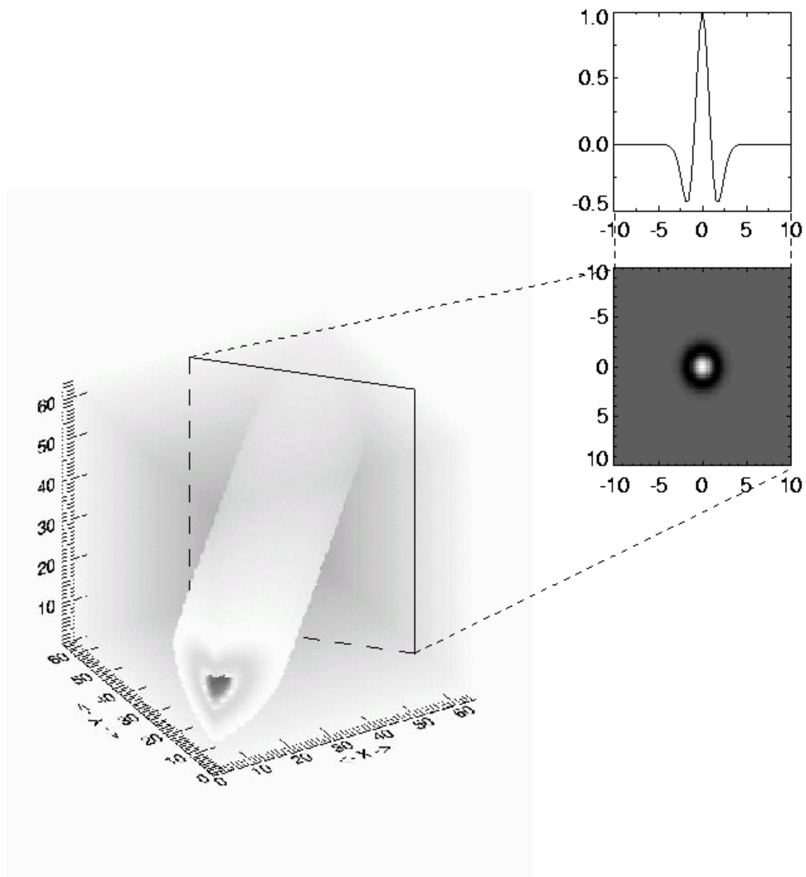
3D Wavelet Function

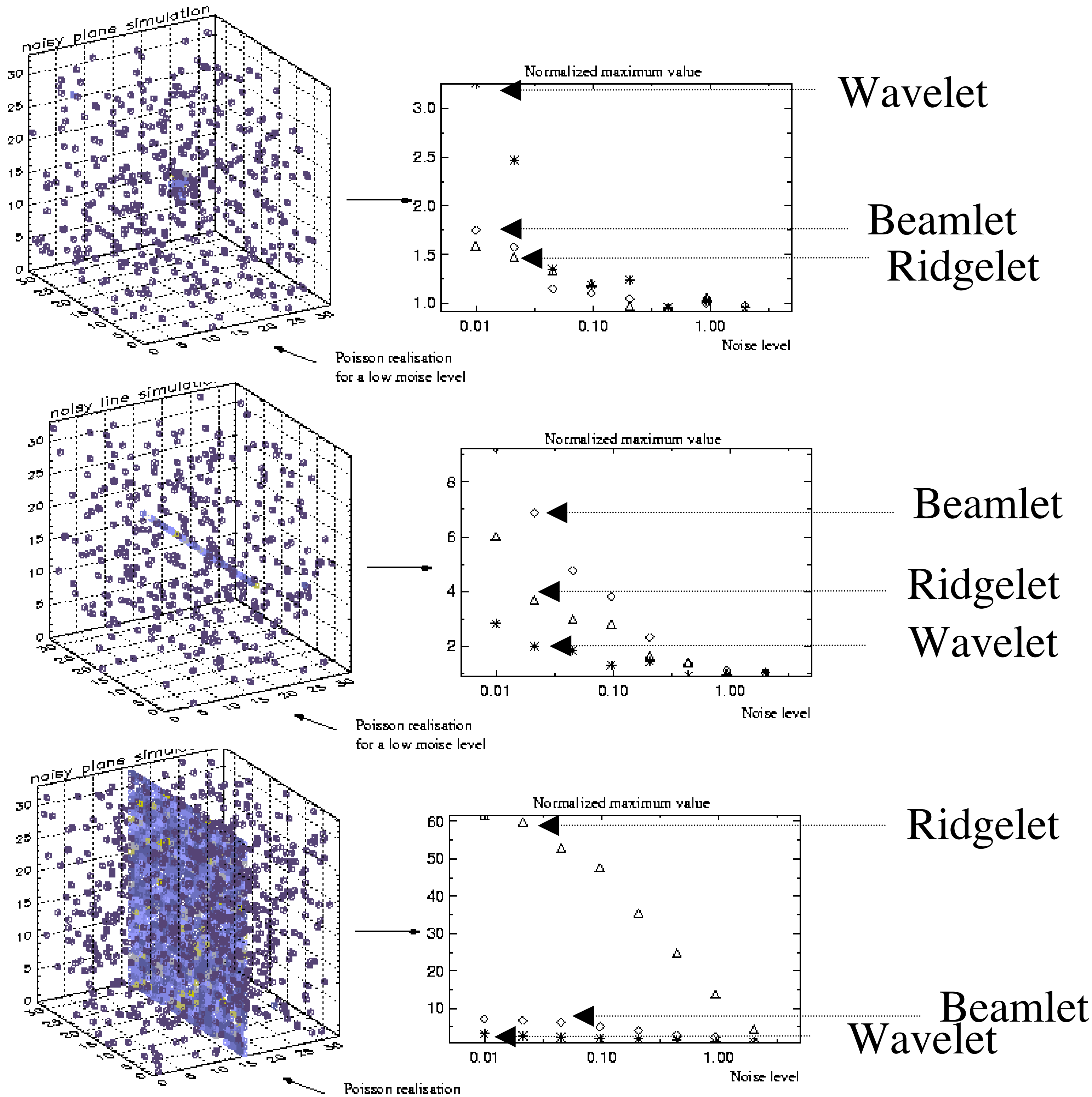


3D Ridgelet Function

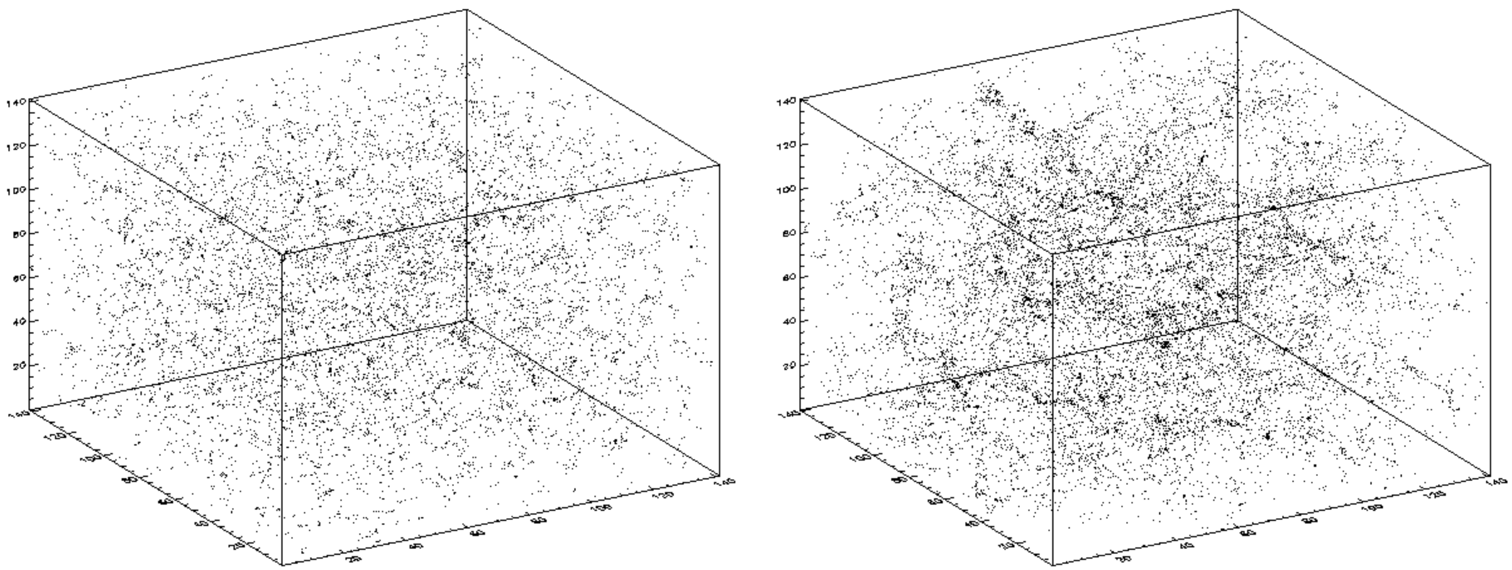


3D Beamlet Function




noisy plane simulation
Normalized maximum value
Noise level
Poisson realisation
for a low moise level
Wavelet
Beamlet
Ridgelet
noisy line simulation
Normalized maximum value
Noise level
Poisson realisation
for a low moise level
Beamlet
Ridgelet
Wavelet
noisy plane simul
Normalized maximum value
Noise level
Poisson realisation
Ridgelet
Beamlet
Wavelet

Simulations



Skewness and Kurtosis