Deconvolution and Compressed Sensing

Jean-Luc Starck, Florent Sureau
J. Bobin, N. Barbey, A. Woiselle
Outline of this Presentation

• Deconvolution and CS?
  – Inverse Problems
  – Sparsity as regularization?
• Deconvolution and sparsity
  – Classical approaches
  – Sparsity-based regularization
  – First PE on simulated Hubble images
• Compressed Sensing
  – Early experiments and theory
  – Radio-Interferometry
  – Herschel Example
  – Second PE on CS/radiointerferometry
Outline of this Presentation

• Deconvolution and CS?
  – Inverse Problems
  – Sparsity as regularization?

• Deconvolution and sparsity
  – Classical approaches
  – Sparsity-based regularization
  – First PE on simulated Hubble images

• Compressed Sensing
  – Early experiments and theory
  – Radio-Interferometry
  – Herschel Example
  – Second PE on CS/radiointerferometry
Why Deconvolution and CS?

• Two historically and theoretically quite different fields:
  – Deconvolution: extensively covered in the last 30 years. Problem arising in all observations. Inversion of a singular structured operator? Noise Amplification? Tradeoff Resolution/Noise? Choice of hyperparameters...
  – Compressed sensing: young field (about 10 years), with up to now only a few applications. Framework for new sensing approaches in case of sparse signals. Theoretical proof of unique solution? Guarantee to recover the solution, even in presence of noise? Finding applications.

• Two ill-posed inverse problems:
  – Spectrum of the convolution operator (e.g. Gaussian kernel)
  – CS: Underdetermined problem (less measurements than parameters)

• Both have benefited from multiresolution approaches and use essentially the same algorithms to recover a solution
  → Good illustration of what you can get from sparsity

• Radiointerferometry, sensing only part of Fourier space, can be seen both as a (badly conditioned) deconvolution problem and a soft CS problem.
Linear Inverse Problems

\[ Y = HX + N \]

\( Y \) is the measured data of size \( m \)

\( X \) is the vector of parameters to estimate of size \( n \)

\( N \) is a \( m \)-vector of additive noise (Gaussian) \( N \sim \mathcal{N}(0, \Sigma) \)

\( H \) is the \( m \times n \) system matrix

E.g. : \( H \) Toeplitz matrix or stationary convolution matrix, \( H \) incoherent with sparsity pattern for compressed sensing (with \( m << n \))

Weighted Least Square solution: \( \arg \min_X ||Y - HX||_\Sigma^2 \)

→ Ill-posed problem (no unique stable solution): needs to constrain more the solution

Weighted Least Square solution + regularization: \( \arg \min_X ||Y - HX||_\Sigma^2 + R(X) \)

→ Enforcing sparsity as regularization (among other possible constraints)

→ Find efficient representation for the signal (number coefficients)
What is a good representation for data?

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

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

**Dictionary (basis, frame)**

Ex: Haar wavelet

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

- Fast calculation of the coefficients
- Tight frame (no change of the $l_2$ norm)
- Analyze the signal through the statistical properties of the coefficients
- Approximation theory uses the sparsity of the coefficients

$f(x) = c_0 + \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} c_{jk} \psi_{jk}(x)$. 

![Graph of signal representation](image)
How to Measure and Enforce Sparsity?

with \( 0^0 = 0, \quad \| \mathbf{\alpha} \|_0 = \sum_k \alpha_k^0 = \# \{ \alpha_k \neq 0 \} \)

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

\[
\begin{align*}
(P0) \quad & \underset{\alpha}{\text{minimize}} \quad \| \mathbf{\alpha} \|_0 \\
& \text{subject to } \mathbf{s} = \Phi \mathbf{\alpha}
\end{align*}
\]

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

\[
\begin{align*}
(P1) \quad & \underset{\alpha}{\text{minimize}} \quad \| \mathbf{\alpha} \|_1 \\
& \text{subject to } \mathbf{s} = \Phi \mathbf{\alpha}
\end{align*}
\]

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 (P1), then it is identical to the solution of (P0).

**Our optimizations Problem to enforce sparsity:**

\[
\begin{align*}
& \min_{\mathbf{\alpha}} \| \mathbf{\alpha} \|_p^p \quad \text{subject to } \quad \mathbf{Y} = \mathbf{H} \Phi \mathbf{\alpha} \\
& \quad \quad \quad \quad \quad p < 2 \\
& \min_{\mathbf{\alpha}} \| \mathbf{\alpha} \|_p^p \quad \text{subject to } \quad \| \mathbf{Y} - \mathbf{H} \Phi \mathbf{\alpha} \|_2^2 \leq \epsilon
\end{align*}
\]
Outline of this Presentation

• Deconvolution and CS?
  – Inverse Problems
  – Sparsity as regularization?

• Deconvolution and sparsity
  – Classical approaches
  – Sparsity-based regularization
  – First PE on simulated Hubble images

• Compressed Sensing
  – Early experiments and theory
  – Radio-Interferometry
  – Herschel Example
Stationary Convolution

- System Matrix $H$ described by convolution integral (PSF):
  \[ Y(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} H(x - x_1, y - y_1) X(x_1, y_1) dx_1 dy_1 + N(x, y) \]

- Operator $H$ diagonal in Fourier
  (spectrum of the operator=Fourier spectrum):
  \[ \tilde{Y}(u, v) = \tilde{H}(u, v) \tilde{X}(u, v) + \tilde{N}(u, v) \]

- Operator $H$ badly conditioned:
  - Potential zero eigenvalues (zeros in frequency)?
  - Noise dominating regime in Fourier space
  - Slow iterative algorithms
Knowing the PSF

- Other potential pitfalls:
  - What is the PSF? Estimate it from data (noise?), or from optical model of the imaging telescope (accuracy?)
  - Is the PSF stationary or varying in the field?
  - In astronomy, stars are point-like objects which can be used to construct the PSF
  - Radially symmetric model proposed in several studies:

\[
H(r) \propto \left(1 + \frac{r^2}{R^2}\right)^{-\beta}
\]

with $\beta$ and $R$ fitted to the data
Example PSF: ISOCAM

Simulation: weak galax. near bright *, convolv. with ISOCAM Psf, noise

max on 17 11
First Attempt: Brute Force Inversion

- Ignore Problems:

\[ \hat{X}(u, v) = \frac{\tilde{H}^* \tilde{Y}(u, v)}{\|\tilde{H}(u, v)\|^2} = \hat{X}(u, v) + \frac{\tilde{H}^* \tilde{N}(u, v)}{\|\tilde{H}(u, v)\|^2} \]

- Advantages:
  - Very fast
  - Noise in the solution predictable

- Disadvantages:
  - Very noisy
  - Singularities in the spectrum? (Truncated SVD)

- Never consider using it in presence of noise/singular operator
More Sophisticated Techniques

- Deconvolution methods used in astronomy:
  - Wiener Solution
  - Richardson-Lucy → non-regularized so noise amplification
  - Maximum Entropy Method → Problem to restore stars
  - CLEAN Method → Problem to restore extended sources

- Signal Processing:
  - Markov Random Fields
Framework of Bayesian Analysis

- Using Bayes Law: \[ \Pr(X|Y) = \frac{\Pr(Y|X) \Pr(X)}{\Pr(Y)} \]

- Minimizing the antiloglikelihood:
  \[
  \text{minimize } - \log(\Pr(Y|X)) - \log(\Pr(X)) \rightarrow \text{data fidelity + regularization}
  \]

- Without any prior: maximum likelihood estimator

- With priors: maximizing the posterior (MAP approach)

- Examples:
  - Poisson Noise, maximum likelihood using steepest descent:
    \[
    \Pr(Y|X) = \prod_{(x,y)} \frac{(HX(x,y))^{Y(x,y)} \exp(-HX(x,y))}{Y(x,y)}
    \]
    \[
    X^{n+1} = X^n \odot H^* \left[ \frac{Y}{HX^n} \right]
    \]
    \[\rightarrow\] Richardson-Lucy deconvolution
Framework of Bayesian Analysis

- Other Examples:
  - Gaussian Noise, maximum likelihood using steepest descent:
    \[
    \Pr(Y|X) = \frac{1}{2\pi|\Sigma|} \exp((Y - HX)^T\Sigma^{-1}(Y - HX))
    \]
    \[
    \hat{X}^{n+1} = \hat{X}^n + \gamma H^*\Sigma^{-1}(Y - H\hat{X})
    \]
    Case i.i.d noise: \( \hat{X}(u, v) = \frac{\tilde{H}^*(u, v)\tilde{Y}(u, v)}{||\tilde{H}(u, v)||^2} \)
    
    → Direct inversion

  - \( \tilde{N}(u, v) \sim \mathcal{N}(0, \sigma_N^2(u, v)) \), \( \Pr(\tilde{X}(u, v)) \sim \mathcal{N}(0, \sigma_X^2(u, v)) \),
    MAP solution:
    \[
    \hat{X} = \frac{\tilde{H}^*(u, v)\tilde{Y}(u, v)}{||\tilde{H}(u, v)||^2 + \frac{\sigma_N^2(u, v)}{\sigma_X^2(u, v)}}
    \]
    
    → Wiener Filter
Generalized Linear Methods

- More General Linear inverse methods: 
  \[ \hat{X}(u,v) = \hat{W}(u,v) \frac{\tilde{Y}(u,v)}{\tilde{H}(u,v)} \]

  minimize \[ ||Y - HX||^2 + \lambda ||X||^2 \]

  \[ \hat{W}(u,v) = \frac{||\tilde{H}(u,v)||^2}{||\tilde{H}(u,v)||^2 + \lambda \tilde{H}(u,v)} \]

  \rightarrow \text{Tikhonov Regularization}

- Advantages of linear regularized methods:
  
  1. Very fast
  2. Noise predictable

- Disadvantages of linear regularized methods:
  
  1. Gibbs artefacts in the neighbourhood of discontinuities
  2. No firm constraints such as positivity (unphysical solutions)
  3. Degraded resolution (window function low pass filter). Trading resolution for noise
  4. Estimating noise with incorrect PSF: limited practical interest?
Adding other constraints

- Let’s call $\mathcal{P}_C(.)$ an operator enforcing a set of constraints (projection onto a convex set) e.g. Positivity: $\mathcal{C} = (\mathbb{R}^+)^n$, Spatial/frequency support:

$$
\mathcal{P}_C(X)_k = \begin{cases} 
X_k & \text{if } X_k \in D \\
0 & \text{otherwise}
\end{cases}
$$

- Iterative scheme to enforce these constraints e.g. Projected Landweber: $X^{n+1} = \mathcal{P}_C [X^n + \gamma H^*(Y - HX)]$

- Entropy based regularization:

$$
\text{minimize } ||Y - HX||^2 + \lambda \sum_{i,j} X_{i,j} \log X_{i,j}
$$

- Gradient based regularization:

$$
\text{minimize } ||Y - HX||^2 + \lambda \sum_{i,j} \phi\left( \frac{||\nabla X_{i,j}||_p}{\delta} \right)
$$

  e.g. TV: $p = 1, \phi(x) = |x|$

- Markov Random Fields etc.
Examples of $\phi$ functions:

1. $\phi_q(x) = x^2$: quadratic function.
2. $\phi_{TV}(x) = |x|$: Total Variation.
3. $\phi_2(x) = 2\sqrt{1 + x^2} - 2$: Hyper-Surface (Charbonnier et al, 1997).
4. $\phi_3(x) = x^2/(1 + x^2)$ (Geman and McClure, 1985) (Non convex).
5. $\phi_4(x) = 1 - e^{-x^2}$ (Perona and Malik, 1990).
6. $\phi_5(x) = \log(1 + x^2)$ (Herbert and Leahy, 1989).
Iterative Algorithms for Deconvolution

- Advantages of iteratives methods with constraints:
  1. Flexibility to add constraints
  2. Convergence proven to a unique minimum/minimum cost function (convex/strictly convex)
  3. Good regularization: better trade-offs noise/resolution

- Disadvantages of iteratives methods with constraints:
  1. Slow, or super slow (choose adequate algorithm)
  2. Non-linear methods: noise?
  3. Choice of hyperparameters in Lagrangian formulation
MultiResolution Approaches (1)

- Based on the ideas of Linear inversion and wavelet denoising: Wavelet-Vaguelette decomposition (Donoho 1995)

- Case of deconvolution: wavelets compactly supported in Fourier to perform inversion, then (hard/soft) thresholding of wavelet coefficients to regularize.

- Problem: classical wavelets not able to concentrate energy of colored noise after inversion: low pass filter.

- Solution: Mirror wavelet bases (Khalifa, Mallat 2000).
MultiResolution Approaches (2)

- Adding Sparsity constraints using Wavelets ($\mathcal{W}$):

$$||Y - HX||^2 + \lambda \sum_{j,k,l} \phi(||(\mathcal{W}X)_{j,k,l}||_p^p)$$

$$||Y - H\Phi\alpha||^2 + \lambda \sum_i^I \phi(||\alpha_i||_p^p)$$

- E.g. $\phi(x) = x \rightarrow \ell_p$ norm. When Haar wavelet with 1 decomposition level: very close to TV

- In a Bayesian framework, leptokurtic distributions (generalized gaussians)

$$\text{pdf}(\alpha_1, ..., \alpha_i) \propto \prod_{i=1}^I \exp(-\lambda||\alpha_i||_p^p)$$
Algorithms for Wavelet regularization

\[
\min_{\alpha} \|Y - H\Phi\alpha\|^2 + \lambda \sum_{i} |\alpha_i|
\]

- Forward Backward Algorithm (FBA)/ Iterative Soft Thresholding (ISTA) (Daubechies 04, Combettes 05):

\[
X^{n+1} = S\!T_{\lambda\tau}(X^n + 2\tau H^*(y - HX))
\]

where \(S\!T_{\lambda}(X)_i = (|x_i| - \lambda)_+ \text{sign}(x_i)\)

- Accelerated version exists (FISTA)

- Iterative Thresholding with varying threshold \(\lambda_n\) (Starck04, Elad 05) allows to accelerate the convergence.

\[
\begin{align*}
X^{n+1} &= S\!T_{\lambda_n\tau}(X^n + 2\tau H^*(y - HX)) \\
X^{n+1} &= H\!T_{\lambda_n\tau}(X^n + 2\tau H^*(y - HX))
\end{align*}
\]

where \(H\!T_{\lambda}(X)_i = X\delta(|X| > \lambda)\)

- More Refs: Vonesch et al, 2007; Elad et al 2008; Wright et al., 2008; Nesterov, 2008 and Beck-Teboulle, 2009; Blumensath, 2008; Maleki et Donoho, 2009 ; etc.
MultiResolution Support

- The idea is to detect the significant coefficients in the data prior to inversion: multiresolution mask $M$

- Use this information as regularization:
  - For Landweber iteration (can add positivity constraint):
    \[
    X^{n+1} = X^n + H^*(\mathcal{W}^{-1}M\mathcal{W}(Y - HX^n))
    \]
  - For Richardson-Lucy iteration:
    \[
    X^{n+1} = X^n \odot H^* \left[ \frac{HX^n + \mathcal{W}^{-1}M\mathcal{W}(Y - HX^n)}{HX^n} \right]
    \]
Practical Exercise no 1
Outline of this Presentation

- Deconvolution and CS?
  - Inverse Problems
  - Sparsity as regularization?
- Deconvolution and sparsity
  - Classical approaches
  - Sparsity-based regularization
  - First PE on simulated Hubble images
- Compressed Sensing
  - Early experiments and theory
  - Radio-Interferometry
  - Herschel Example
  - Second PE on CS/radiointerferometry
A Surprising Experiment

Randomly throw away 83% of samples

* E.J. Candes, J. Romberg and T. Tao.
A Surprising Result*

Minimum - norm conventional linear reconstruction

* E.J. Candes, J. Romberg and T. Tao.
A Surprising Result*

Minimum - norm conventional linear reconstruction

E.J. Candes
Compressed Sensing


A non linear sampling theorem

“Signals with exactly K components different from zero can be recovered perfectly from ~ K log N incoherent measurements”

Replace samples with few linear projections

\[ y = \Theta x \]

\[ M \times 1 \]
measurements
\[ \Theta \]
N \times 1
sparse signal
\[ x \]
M \times N
K
nonzero entries

\[ K < M \ll N \]

Reconstruction via non linear processing:

\[ \min_x \|x\|_1 \quad \text{s.t.} \quad y = \Theta x \]

Conditions on the sensing matrix to get stable and robust recovery (RIP)

⇒ Application: Compression, tomography, ill posed inverse problem.
In practice, $x$ is sparse in a given dictionary:

$$x = \Phi \alpha$$

and we need to solve:

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

The mutual incoherence is defined as

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

and the number of required measurements is:

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

\[ Y = \Theta X = \Theta \Phi \alpha \]

\[ \mu_{\Theta, \Phi} = \max_{i,k} |\langle \Theta_i, \Phi_k \rangle| \]

\( y \)
\( \Theta \)
\( \Phi \)
\( \alpha \)

Measurement System
Prior: Data Representation System
power-law
sorted index
\( K \)
\( N \)
Radio-Interferometry

CLEAN Algorithm: $\Phi = I_d$  
(Wagbom, 1974)

Wavelet Clean $\Phi = \text{Wavelet Transform}$  
(Wakker, and Schwarz, 1988; Starck et al. 1994)
Radio-Interferometry

\[ y = \Theta x \]

\( \Theta \) = Fourier transform
\( \Phi \) = Id (or Wavelet transform)

\[ \min_{\alpha} \|\alpha\|_1 \text{ s.t. } y = \Theta \Lambda \Phi \alpha \]


Wavelet - CLEAN minimizes well the \( l_0 \) norm

But recent \( l_0-l_1 \) minimization algorithms would be clearly much faster.

\[ \Rightarrow \] See (McEwen et al, 2011; Wenger et al, 2010; Wiaux et al, 2009; Cornwell et al, 2009; Suskimo, 2009; Feng et al, 2011).

The future Square Kilometre Array (SKA) radio-interferometer will certainly use such techniques.
The Applications of Compressive Sensing to Radio Astronomy: I Deconvolution

Australian Square Kilometer Array Pathfinder (ASKAP) radio telescope.
CS-Radio Astronomy

Hogbom CLEAN

MEM residual
Missing Data

- Period detection in temporal series
  COROT: HD170987

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

- Sandrine’s presentation on Wednesday
This space telescope has been designed to observe in the far-infrared and sub-millimeter wavelength range:

Goals:
- Understand the beginning of stars formation (molecular clouds).
- Observe galaxies at their formation epoch.

On bord: three instruments:
- HIFI: spectromètre hétérodyne (SRON-NL)
- PACS: Spectromètre et Photomètre 57-205 µm (MPE-D)
- SPIRE: Spectromètre et Photomètre - 200-607 µm

The shortest wavelength band, 57-210 microns, is covered by PACS (Photodetector Array Camera and Spectrometer).

PACS: 8 matrices of 16x16 pixels, cooled down to 300 mK.
Herschel Status

- Herschel produces already beautiful images:

Herschel/PACS Images of M51 ("Whirlpool Galaxy")

160 μm  100 μm  70 μm

© ESA & The PACS Consortium
Herschel data transfer problem:

- no time to do sophisticated data compression on board.
- a compression ratio of 8 must be achieved.

==> solution: averaging of eight successive images on board

CS may offer another alternative.


Compressed Sensing presents several interesting properties for data compress:

- Compression is very fast ==> good for on-board applications.
- Very robust to bit loss during the transfer.
- Decoupling between compression/decompression.
- Data protection.
- Linear Compression.

But clearly not as competitive to JPEG or JPEG2000 to compress an image.
Transfering Spatial Data to the Earth

Good measurements must be incoherent with the basis in which the data are assumed to be sparse.

Noiselets (Coifman, Geshwind and Meyer, 2001) are an orthogonal basis that is shown to be highly incoherent with a wide range of practical sparse representations (wavelets, Fourier, etc).

Advantages:
Low computational cost \( (O(n)) \)
Most astronomical data are sparsely represented in a wide range of wavelet bases
Herschel image packets decompression

\[
\begin{align*}
\min_{\alpha} & \|\alpha\|_{\ell_1} \text{ s.t. } \|y - \Theta \Phi \alpha\|_{\ell_2} \leq \epsilon \\
\end{align*}
\]

\[
x = \Phi \alpha
\]

Physical priors

Eight consecutive observations of the same field can be decompressed together (forward-backward splitting algorithm)

\[
\alpha^{(t+1)} = \text{SoftThresh}_{\mu_t \lambda(t)} \left( \alpha^{(t)} + \mu_t \frac{1}{P} \sum_{i=1}^{P} \Phi^* S_{-d_i} \left( \Theta_{\Lambda_i}^* \left( y_i - \Theta_{\Lambda_i} S_{d_i} \left( \Phi \alpha^{(t)} \right) \right) \right) \right)
\]

where \( \mu_t \in (0, 2P / \sum_i \Theta_{\Lambda_i}^2 \Phi^2) \).

At each iteration, the sought after image is reconstructed from the coefficients \( \alpha^{(t)} \) as \( x^{(t)} = \Phi \alpha^{(t)} \).
Sensitivity: CS versus mean of 8 images
Resolution: CS versus Mean

Simulated image

Simulated noisy image with flat and dark

Mean of six images

Compressed sensing reconstructed images

Resolution limit versus SNR

<table>
<thead>
<tr>
<th>SNR</th>
<th>-17.3</th>
<th>-9.35</th>
<th>-3.3</th>
<th>0.21</th>
<th>2.7</th>
<th>4.7</th>
<th>6.2</th>
<th>7.6</th>
<th>8.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intensity</td>
<td>900</td>
<td>2250</td>
<td>4500</td>
<td>6750</td>
<td>9000</td>
<td>11250</td>
<td>13500</td>
<td>15750</td>
<td>18000</td>
</tr>
<tr>
<td>MO6</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>CS</td>
<td>2.33</td>
<td>2.33</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

The CS-based compression entails a resolution gain equal to a 30% of the spatial resolution provided by MO6.
Data Fusion: JPEG versus Compressed Sensing

Simulated source

One of the 10 observations

Averaged of the 10 JPEG compressed images (CR=4)

Reconstruction from the 10 compressed sensing images (CR=4)
JPEG-2000 Versus Compressed Sensing

Compression Rate: 25

One observation | 10 observations | 20 observations | 100 observations

![Images showing compression results](image1.png)

![Graph showing PSNR vs Number of Observations](image2.png)
ESA wants to test CS

- CS compression has been implemented in the Herschel on-board software (as an option).

- Astronet Grant: 1 postdoc for 3 years, from 2009 to 2011.

- Uncompressed data observed in November/December 2009 in order to evaluate CS.

- The CS decompression needs to be fully integrated in the data processing pipeline.

- Software developments required for an efficient decompression which takes into calibration problems.
PACS Calibration

The following is applied to the data before the inversion step (back-projection):

– Removal of the offset by removing the temporal mean along each timeline
– Deglitching
– Pink noise filtering using high pass median filtering (removes only the slowly drifting part)

\[ Y(i, j, t) = s(i, j) + B(i, j, t) + N(i, j) \]

We need to recover \( s \) and \( B \)
\[ Y(i, j, t) = s(i, j) + B(i, j, t) + N(i, j); \]
Transfering Data to the Earth

Observed Herschel Data During the Calibration Phase, November 2010, without any compression.

A scan (16 x 16 pixels at 40 Hz during 25 min each, we obtained 60000 images.

**Compressed Sensing Reconstruction**

**Official Pipeline Reconstruction: Averaging**
Map from Uncompressed Data

Compressed Sensing Reconstruction

Official Pipeline

Averaging + Deblurring
Conclusions on CS for Herschel

- CS works.
- CS is clearly better than Averaging, as predicted from the toy model simulations.
- But only slightly better than Averaging + Deblurring
  ==> at this point, the improvement does not justify to use the CS mode as the standard mode.
- The Averaging-Deblurring solution has been developed thanks to the CS spirit.

- It can however be very useful for some scientific programs, where the resolution is the key of the success.

Maybe, it is not the end of the story. Possible improvement:
- Better drift removal.
- Matrix choice (Hadamard, noiselet, etc)
- Dictionary choice.
- Deconvolution would be possible with CS, and further improve the resolution.
Sparsity in Astronomy: conclusions

Sparsity is very efficient for

- Inverse problems: denoising, deconvolution, etc (ISO, XMM, Chandra, FERMI).
- Inpainting, Component Separation (PLANCK).

Compressed Sensing Perspectives

- Radio-Astronomy (LOFAR, SKA).
- Tomographic 3D Weak Lensing (Euclid, DES, LSST).
Practical Exercise no 2