A sparsity-based method to restore MUSE hyperspectral data

Dedicated Algorithms for HyperspectraL Imaging in Astronomy

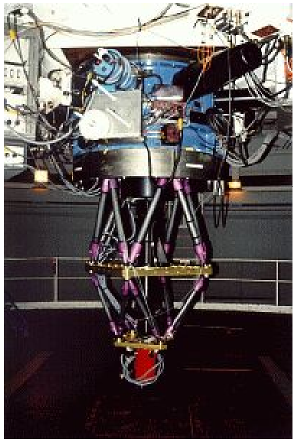
url : dahlia.oca.eu

E.Slezak – S.Bourguignon – D.Mary
(UNS / CNRS / Observatoire de la Côte d'Azur)

- References -

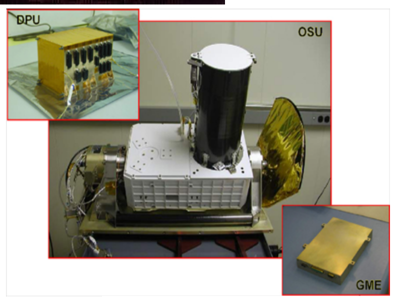
Bourguignon,S.,Mary,D.,Slezak,E.: 2011, Statistical Methodology 9(1)
Bourguignon,S.,Mary,D.,Slezak.E,: 2011, IEEE Selected topics on Signal Proc. 5(5),1002-1013

The era of integral-field spectroscopy

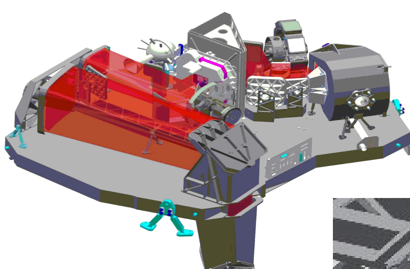


sauron@wht

- giraffe/flames@vlt
- sinfoni@vlt
- kmos@vlt2
- osiris@keck
- miri@jwst

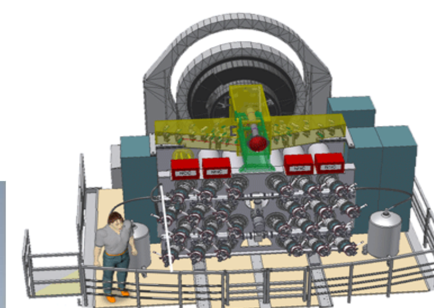


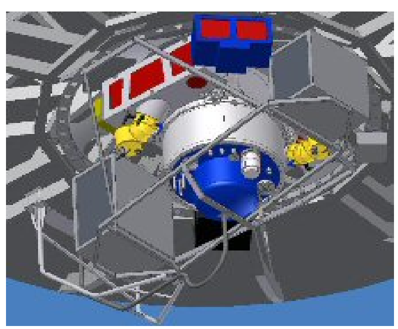
crism@mro



nirspec@jwst

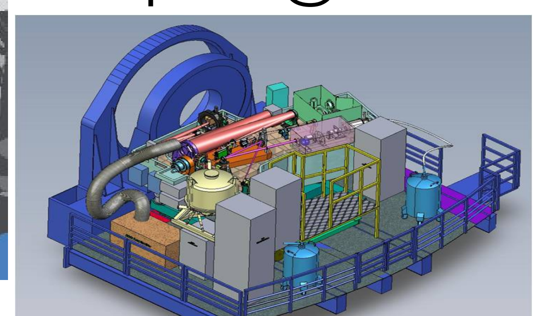
muse@vlt2





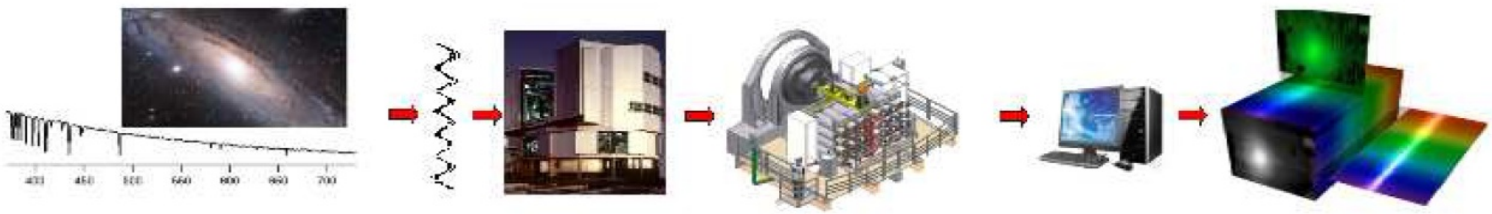
XShooter@vlt2

sphere@vlt2



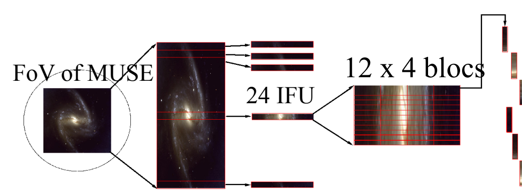
Astronomical hyperspectral data : specificities

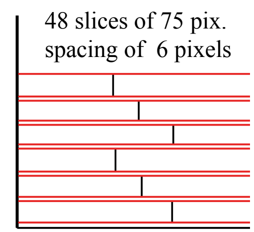
- **Low signal to noise ratio** (S/N < 1 dB) :
  - → the high noise level and its detailed statistics have to be taken into account
  - → fusion/combination of exposures (deep field, large field : mosaicing)
- **Ground-based observations** mostly (cf. MUSE @ VLT2) :



- – perturbations from : atmosphere, AO, telescope, instrument, ...
  - → PSF determination (spatial and wavelength dependencies) + deconvolution
- – objects are superimposed onto a (varying) background
  - → background estimation and subtraction
- **Astrophysical sources are diverse** :
  - – spatial and spectral shapes may be very different → **segmentation / detection**
  - – overlaps and crowding → source separation
- Massive and complex data :

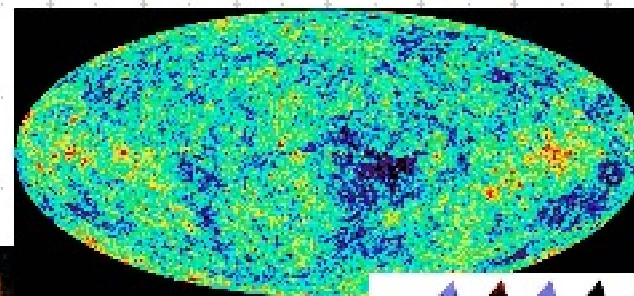
→ data visualization and simulation
300x300 x 4000 elements



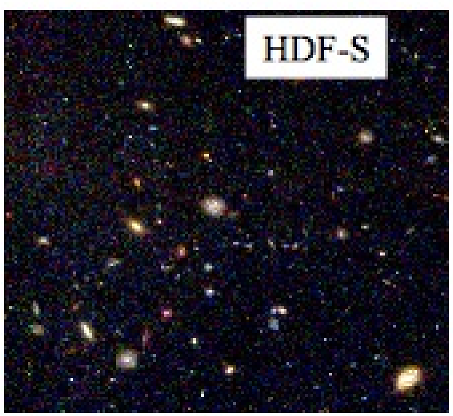


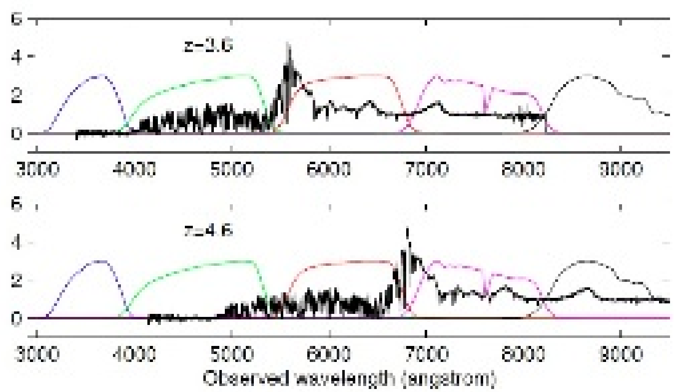
Interest of hyperspectral data in astrophysics

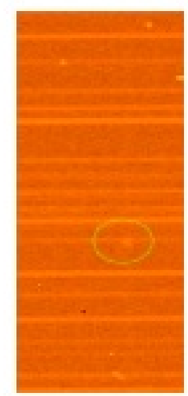
- various physical processes with different spectral behaviours



- discovery of distant (emission-line) objects whatever their redshifts are



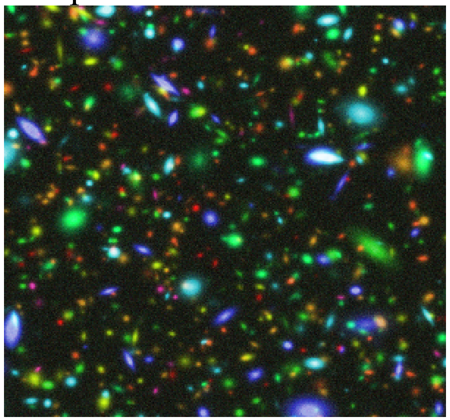




MUSE hyperspectral data Segmentation

astronomical objects vs. astrophysical sources vs. *homogeneous zones*

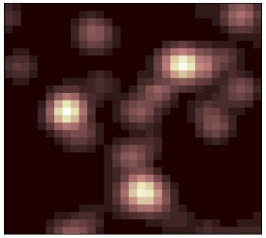
deep field v1 without noise

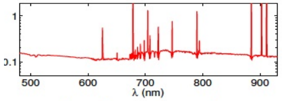


- **reminder :**
  - PSF variations + expected DRS residuals
  - very low S/N ; a spectrally varying noise
  - lack of spectral continuum may occur → LAEs
  - dimension heterogeneity : *images / spectra*

- **Two strategies :**
  - a 2+1 D cube : a set of shapes with a spectral information (Marked Point Process)
  - a 1+2 D cube : a set of spectra (cf. MUSE instrumental characteristics) to be aggregated

noise-free image (part)





noise-free spectrum

1) use and restore the spectral information

(noise reduction, deconvolution, characterization)

2) segmentation

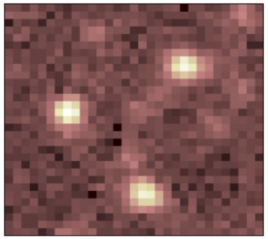
chosen approach

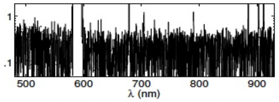
huge data set → dimension reduction is welcomed

very low S/N → a robust method is mandatory

an inverse problem with parcimony constraints

noisy image (part)



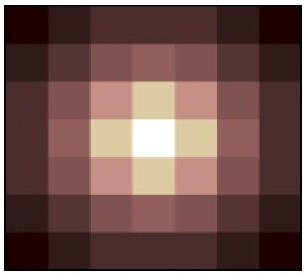


noisy spectrum

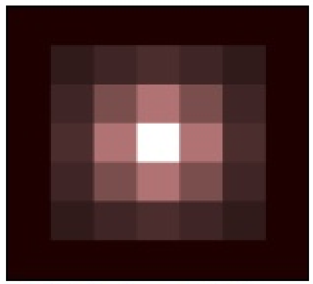
MUSE instrument and noise specificities

• MUSE response $\rightarrow$ spatial and spectral spreading variable with the wavelength

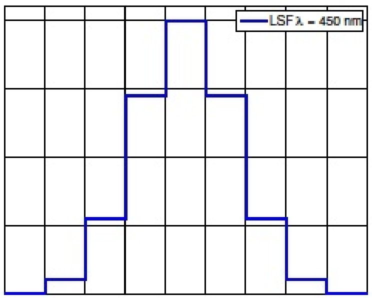
PSF : *Field Spread Function* (FSF) + *Line Spread Function* (LSF)

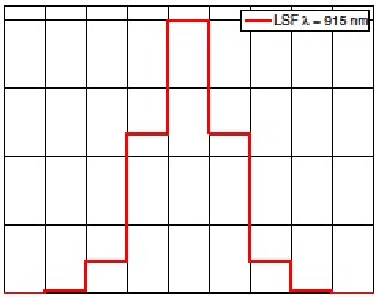


FSF @ $\lambda = 450$ nm

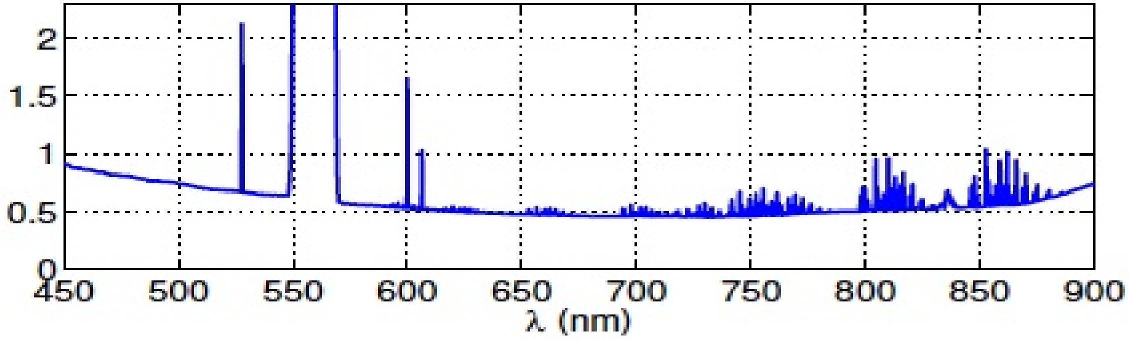


FSF @ $\lambda = 915$ nm





• a spectrally variable noise : (strong) atmospheric emission lines, q. efficiency, laser star (AO)



Restoration of MUSE-like data

Inverse problem : observations $\mathbf{Y} = \mathbf{H}\ \mathbf{X} + \mathbf{N}$ $\rightarrow$ estimate $\mathbf{X}$?

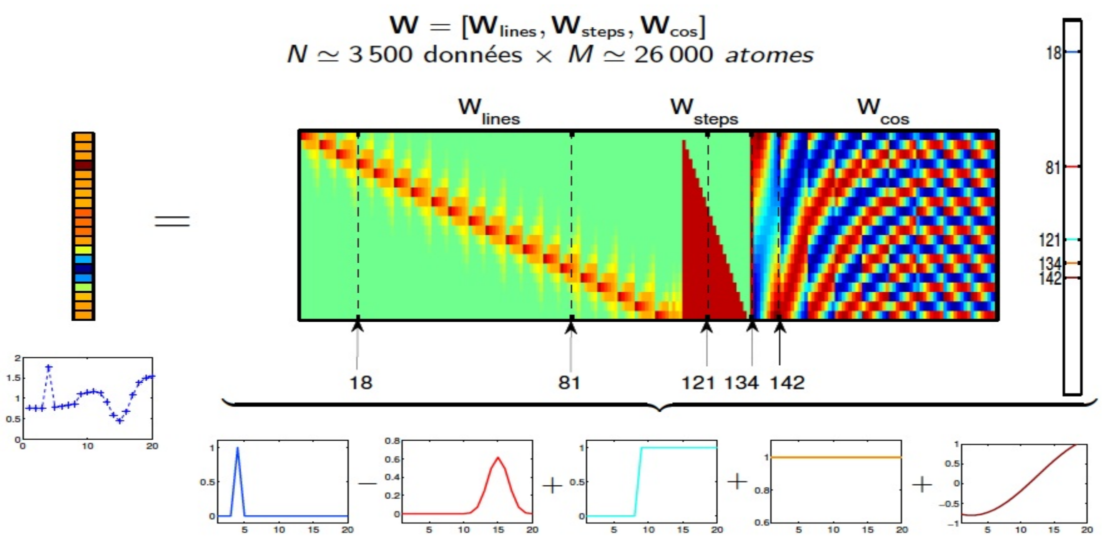
Direct inversion : $\mathbf{H}^{-1}\ \mathbf{X} + \mathbf{H}^{-1}\ \mathbf{N}$ $\rightarrow$ noise amplification

> constrain restoration with additional prior assumptions <

Prior information : sparsity in the spectral domain where the information is located

> galaxy spectrum : a set of elementary features (continuum, lines, discontinuities, etc.) <

$\rightarrow$ dictionary $\mathbf{W}$ of possible spectral features $\rightarrow$ $x = \mathbf{W}\ u$ with u sparse



Estimation setting

Considering $\mathbf{Y} = \mathbf{H}\,\mathbf{X} + \mathbf{N}$ the estimate of $\mathbf{X}$ is given by $\min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{H}\,\mathbf{X}\|^2$

so that, for each pixel k (1,...,K), $\mathrm{x}_{\mathrm{k}} = \mathbf{W}\,\mathrm{x}_k$ with sparse x_k

s.t. $\mathbf{X} = \mathbf{W}^{(K)}\,\mathbf{U}$ with $\mathbf{W}^{(K)}$ made of K blocks $\mathbf{W}$

$$\hat{\mathbf{U}} = \arg\min_{\mathbf{U}} \left\|\mathbf{Y} - \mathbf{H}\mathbf{W}^{(K)}\mathbf{U}\right\|^2 + \alpha\,\|\mathbf{U}\|_1 \,, \alpha > 0$$

→ estimation of active coefficients in each spectrum : $\hat{\mathbf{U}}_\star = \{\hat{\mathbf{U}}_j \neq 0\}$

i.e. *detection of significant spectral features*

$\|.\|1 \rightarrow$ amplitudes are biaised $\rightarrow$ $\hat{\mathbf{U}}'_\star = \arg\min_{\mathbf{U}_\star} \left\|\mathbf{Y} - \mathbf{H}\mathbf{W}_\star^{(K)}\mathbf{U}_\star\right\|^2$

→ spectra restoration : $\hat{\mathbf{X}} = \mathbf{W}_\star^{(K)}\hat{\mathbf{U}}'_\star$

- spatial and spectral deconvolution with *only spectral prior* information
- 15 x 15 pixels, 4 000 wavelengths → 9 10^5 data points, 7 10^6 unknowns

A two-step (sub-optimal) procedure

$$\hat{\mathbf{U}} = \arg\min_{\mathbf{U}} \left\| \mathbf{Y} - \mathbf{H}\mathbf{W}^{(K)}\mathbf{U} \right\|^2 + \alpha \|\mathbf{U}\|_1 , \alpha > 0 \qquad \text{(eq.1)}$$

i) Decompose all spectra $\mathbf{y}_k$ in $\mathbf{Y}$ *independently* :

$\rightarrow$ sparse approximation of a spatially spread spectral information

$$\hat{\mathbf{u}}_k = \arg\min_{\mathbf{u}} \|\mathbf{y}_k - \mathbf{L}_k\mathbf{W}\mathbf{u}_k\|^2 + \beta \|\mathbf{u}_k\|_1 , \quad \beta > 0 \qquad (\mathbf{L}_k : \text{LSF at pixel } k)$$

$\rightarrow$ detection of significant atoms in y_k : $\mathbf{W}_{\Omega_k}$, $\Omega_k = \{ j \mid \mathbf{u}_{kj} \neq 0 \}$

in pratice : low β values $\rightarrow$ detection of faint features, but more false alarms

ii) Consider eq.1 with **W** being restricted to the atoms selected at step i)

- $\mathbf{x}_k = \mathbf{W}_{\Omega_k} \mathbf{u}_{\Omega_k}$
- $\mathbf{x}_k = \mathbf{W}_{\Omega'_k} \mathbf{u}_{\Omega'_k}$ with $\Omega'_k = \cup_{j \in V(k)} \Omega_j$ and V(k) a local neighbourhood < FSF

$\rightarrow$ optimization in a parameter space Ω of lower dimensionality

$$\hat{\mathbf{U}} = \arg\min \|\mathbf{Y} - \mathbf{H}\mathbf{W}_\Omega\mathbf{U}_\Omega\|^2 + \gamma \|\mathbf{U}_\Omega\|_1 , \quad \gamma > 0$$

in practice : the 2nd sparsity constraint allows one to remove false alarms from i)

Implementations details

i) Dictionary normalization :

Let us consider a Gaussian noise with covariance matrix $\boldsymbol{\Sigma}$. The data misfit term is :

$$\|\mathbf{Y}-\mathbf{HWU}\|_{\boldsymbol{\Sigma}}^2 = (\mathbf{Y}-\mathbf{HWU})^T \boldsymbol{\Sigma}^{-1} (\mathbf{Y}-\mathbf{HWU}) = \left\| \boldsymbol{\Sigma}^{-1/2}\mathbf{Y} - \boldsymbol{\Sigma}^{-1/2}\mathbf{HWU} \right\|^2$$

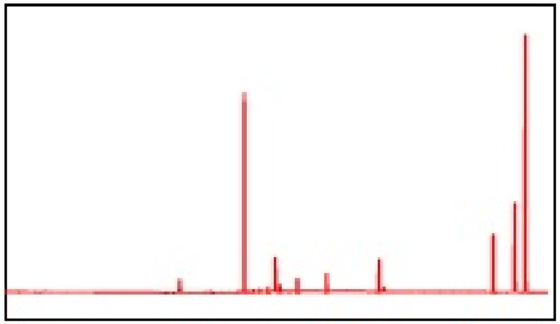
- the equivalent dictionary $\boldsymbol{\Sigma}^{-1/2}\,\mathbf{H}\,\mathbf{W}$ is *not normalized*
- column normalization is necessary for coherent detection statistics

ii) Optimization specificities :

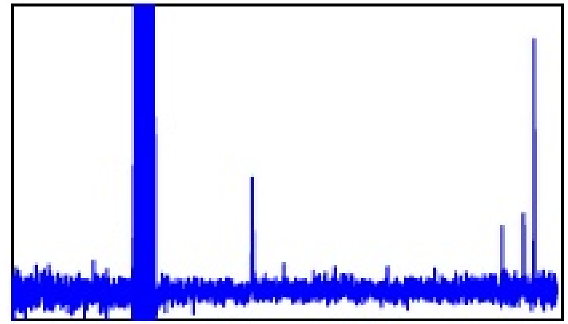
- W structure : no fast transform is available
- involv. sizes: matrix storage is impossible

→ Iterative Coordinate Descent algorithm with *specific accelerations*

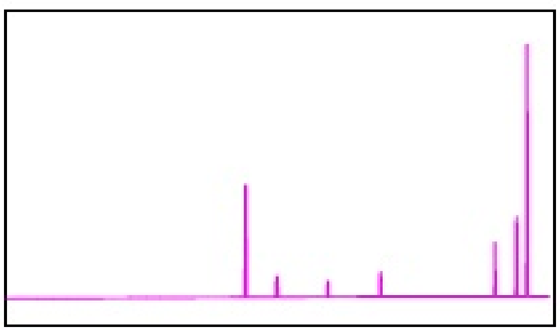
Experimental results : a single spectrum



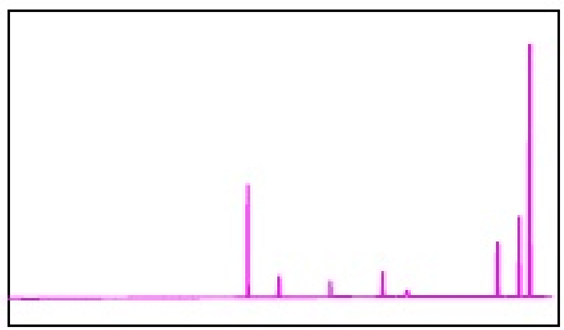
initial spectrum



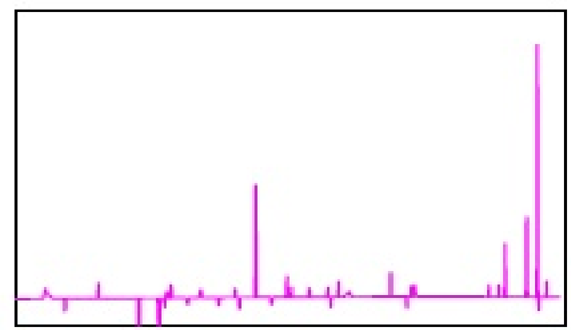
convolved and noisy spectrum



restored : β = 4.2

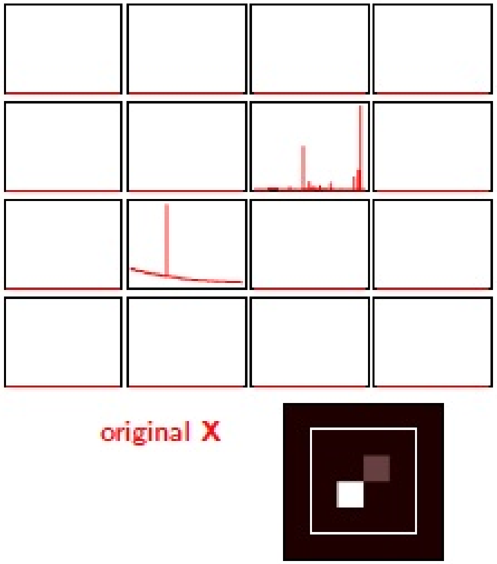


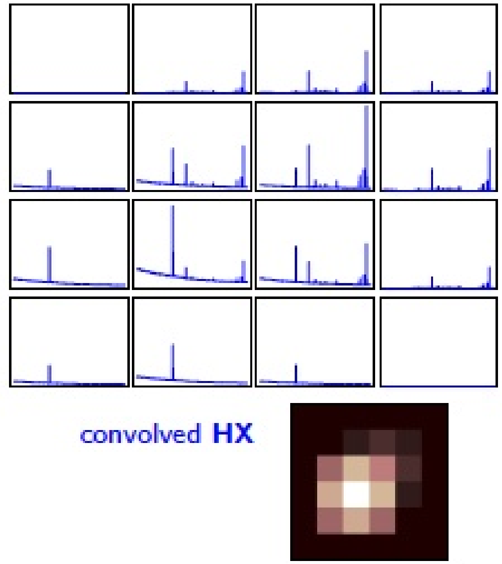
restored : β = 3.5

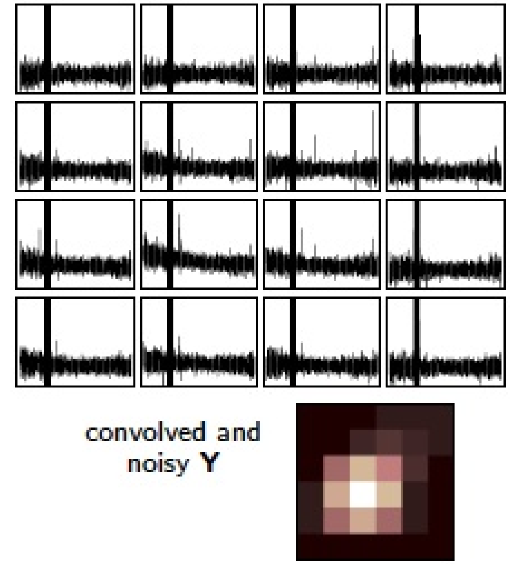


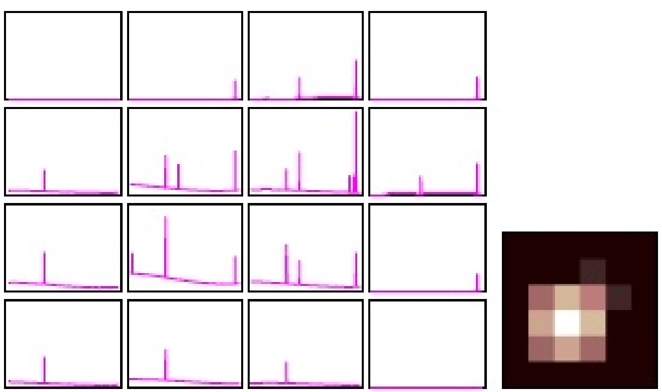
restored : β = 2.5

Experimental results : two point sources

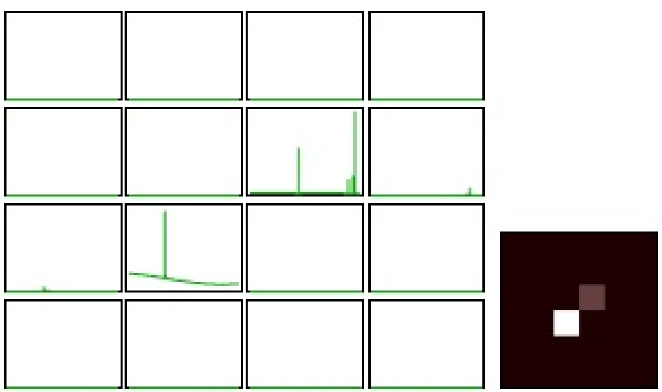








"individual" spectral restoration



spatial-spectral restoration

Computational aspects

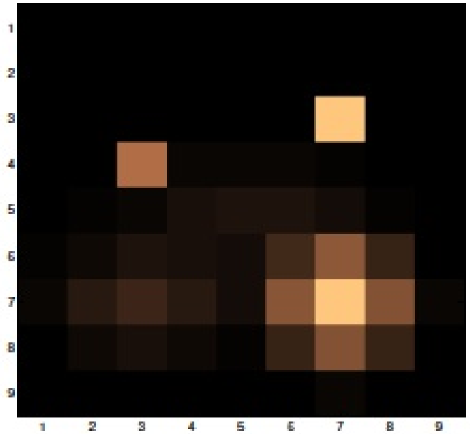
Image 9×9 ; FSF 5×5 $\longrightarrow$ restored image : 13×13 pixels

3 463 wavelengths $\longrightarrow$ 280 000 data

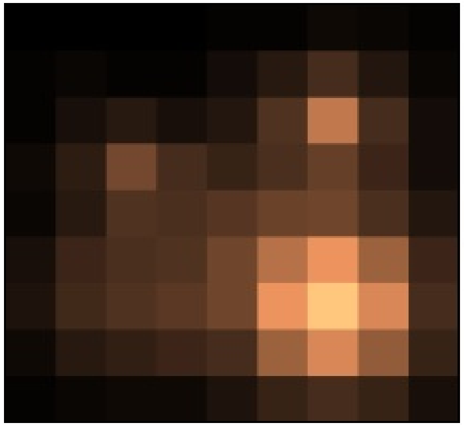
- whole dictionary : $\sim 4\,10^6$ unknowns, computation time = ???
- restoration of x_k using only atoms selected on pixel k

  $\sim 10^3$ unknowns, 82 active coefficients, execution time = 6 mn
- restoration of x_k using selected atoms within a 5×5 neighbourhood

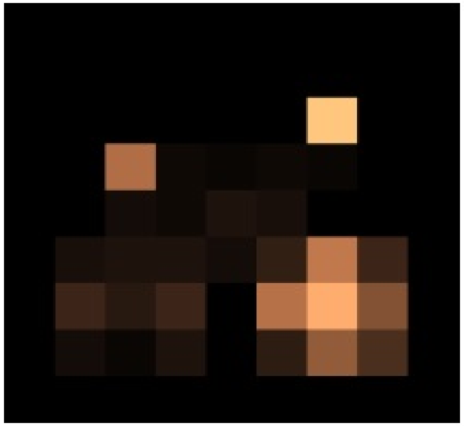
  $\sim 10^4$ unknowns, 89 active coefficients, execution time = 45 mn



initial data



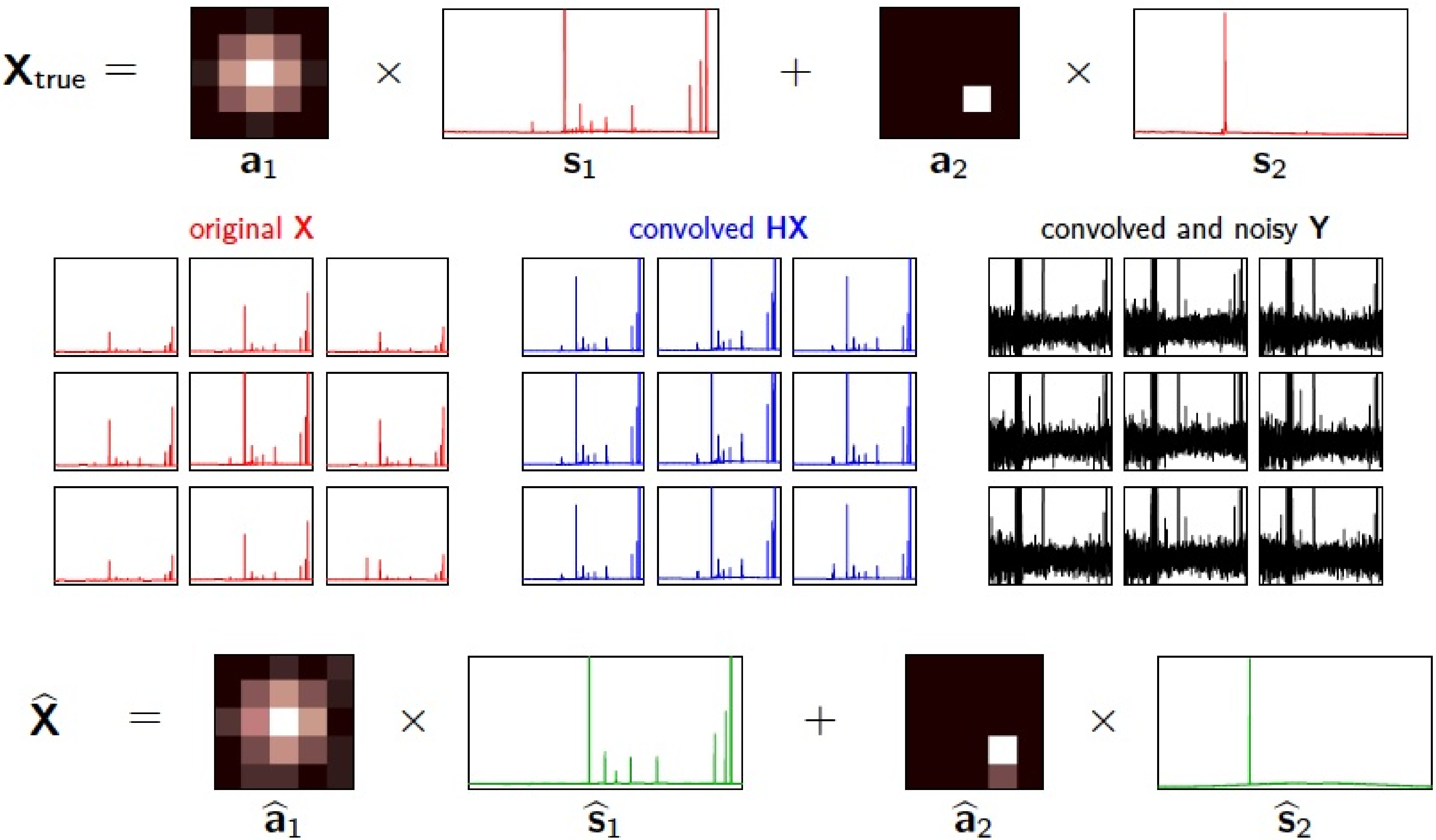
blurred and noisy data



restored data

averaged images

Experimental results : spectral unmixing



Summary and Perspectives

- a restoration scheme which accounts for spectrally varying PSF and noise
- *spatial* and *spectral* deconvolution with *spectral* prior information
- sparsity : physically motivated (vs. generic) set of atoms → robustness
- efficiency in separating point sources and in “simple” unmixing cases

> topics in progress <

→ dictionaries

→ automatic tuning of hyperparameters

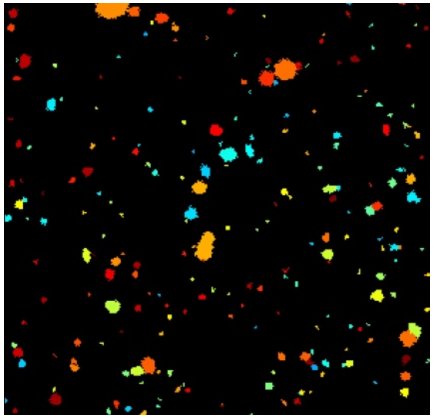
→ more complex priors for specific problems (cf. extended sources)

→ source detection and characterization :

– group pixels with the *same* spectral *signature*

(distance measure : KL ; statistical decision)

a first attempt with a basic criterion ... →



– no constraint on morphologies

– vs. input for a MPP approach ?

→ very faint but extended structures