Inverse problem approaches for hyperspectral data:

toward a better exploitation of integral field spectrographs and polychromatic interferometers.

Ferréol Soulez

Centre de Recherche Astrophysique de Lyon

30 Janvier 2014

Outlines

1. Prologue

   Forward model, inverse problems, deconvolution...
2. Deconvolution in Integral Field Spectrography
3. Demixing in Integral Field Spectrography
4. Image reconstruction in interferometry

Prologue

Astrophysical context

Some key questions:

- What is the origin and evolution of stars and planets?
- How do galaxies form and evolve?
- How black holes interact with their environment?

Needs of high angular and spectral resolution.

Integral fields spectrography

- already high throughput,
- good detectors,
- low SNR for distant galaxies,
- few ELTs.

Optical interferometry

- only bright objects,
- few telescopes,
- model fitting Vs image reconstruction

Astrophysical context

Some key questions:

- What is the origin and evolution of stars and planets?
- How do galaxies form and evolve?
- How black holes interact with their environment?

Needs of high angular and spectral resolution.

Integral fields spectrography

- already high throughput,
- good detectors,
- low SNR for distant galaxies,
- few ELTs.

Optical interferometry

- only bright objects,
- few telescopes,
- model fitting Vs image reconstruction

Astrophysical context

Some key questions:

- What is the origin and evolution of stars and planets?
- How do galaxies form and evolve?
- How black holes interact with their environment?

Needs of high angular and spectral resolution.

Integral fields spectrography

- already high throughput,
- good detectors,
- low SNR for distant galaxies,
- few ELTs.

Optical interferometry

- only bright objects,
- few telescopes,
- model fitting Vs image reconstruction

Astrophysical context

Some key questions:

- What is the origin and evolution of stars and planets?
- How do galaxies form and evolve?
- How black holes interact with their environment?

Needs of high angular and spectral resolution.

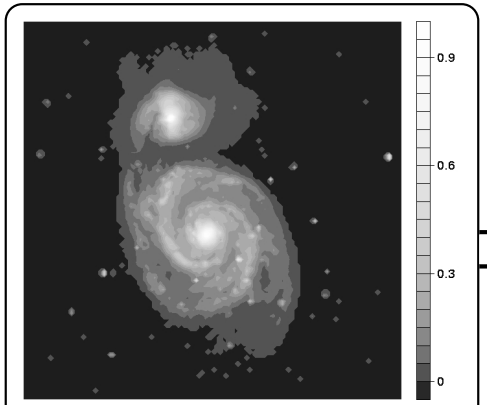
Integral fields spectrography

- already high throughput,
- good detectors,
- low SNR for distant galaxies,
- few ELTs.

Optical interferometry

- only bright objects,
- few telescopes,
- model fitting Vs image reconstruction

Observation model



Object x

Observation medium

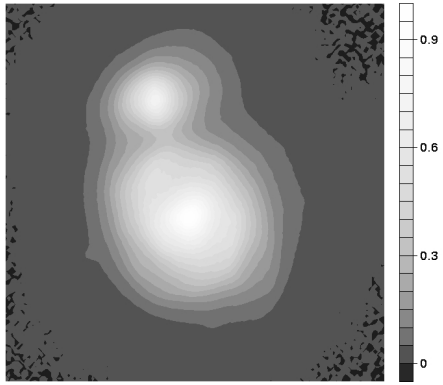
- scattering,
- turbulence,
- ...

Instrumental set up

- motion,
- defocus,
- aberration,
- ...

Sensor

- quantization,
- integration,
- noise ϵ,
- ...



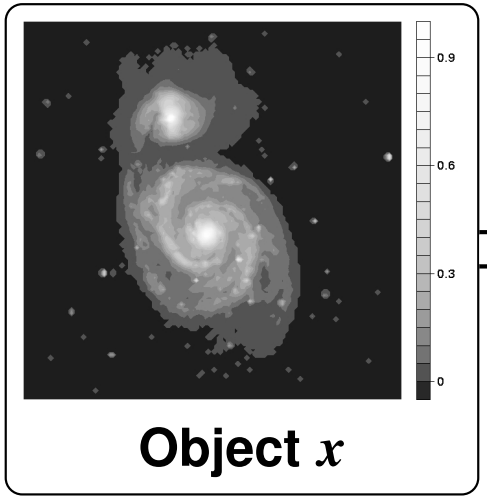
Data y

Direct model of data formation:

$$y = \mathcal{H}(x) + \epsilon$$

$\mathcal{H}()$: observation operator

Observation model



Object x

Observation medium

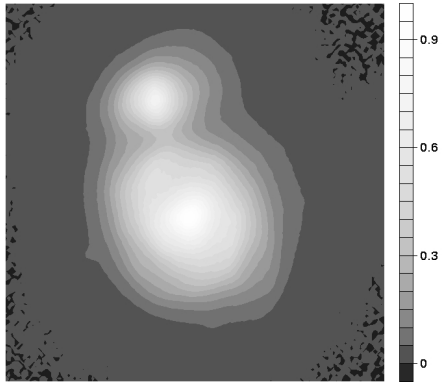
- scattering,
- turbulence,
- ...

Instrumental set up

- motion,
- defocus,
- aberration,
- ...

Sensor

- quantization,
- integration,
- noise ϵ,
- ...



Data y

Direct model of data formation:

$$y = \mathcal{H}(x) + \epsilon$$

$\mathcal{H}()$: observation operator

What is an inverse problem?

We have a model $\mathcal{H}$ linking parameters $\boldsymbol{x}$ and data $\boldsymbol{y}$ with some errors $\boldsymbol{e}$:

$$\boxed{\boldsymbol{y} = \mathcal{H}(\boldsymbol{x}) + \boldsymbol{e}}$$

Inverse problem

What are the best parameters $\boldsymbol{x}$ according to data $\boldsymbol{y}$, the model $\mathcal{H}$ and eventually some priors?

Many parameters (*e.g.* $\gtrsim 10^6$), possibly more parameters than measures: $N_{\boldsymbol{x}} \gtrsim N_{\boldsymbol{y}}$.

In general, inverses problems are ill-conditioned (unstable solution) or ill-posed (several or no solutions).

What is an inverse problem?

We have a model $\mathcal{H}$ linking parameters $\boldsymbol{x}$ and data $\boldsymbol{y}$ with some errors $\boldsymbol{e}$:

$$\boxed{\boldsymbol{y} = \mathcal{H}(\boldsymbol{x}) + \boldsymbol{e}}$$

Inverse problem

What are the best parameters $\boldsymbol{x}$ according to data $\boldsymbol{y}$, the model $\mathcal{H}$ and eventually some priors?

Many parameters (*e.g.* $\gtrsim 10^6$), possibly more parameters than measures: $N_{\boldsymbol{x}} \gtrsim N_{\boldsymbol{y}}$.

In general, inverses problems are ill-conditioned (unstable solution) or ill-posed (several or no solutions).

What is an inverse problem?

We have a model $\mathcal{H}$ linking parameters $\boldsymbol{x}$ and data $\boldsymbol{y}$ with some errors $\boldsymbol{e}$:

$$\boxed{\boldsymbol{y} = \mathcal{H}(\boldsymbol{x}) + \boldsymbol{e}}$$

Inverse problem

What are the best parameters $\boldsymbol{x}$ according to data $\boldsymbol{y}$, the model $\mathcal{H}$ and eventually some priors?

Many parameters (*e.g.* $\gtrsim 10^6$), possibly more parameters than measures: $N_{\boldsymbol{x}} \gtrsim N_{\boldsymbol{y}}$.

In general, inverses problems are ill-conditioned (unstable solution) or ill-posed (several or no solutions).

What is an inverse problem?

We have a model $\mathcal{H}$ linking parameters $\boldsymbol{x}$ and data $\boldsymbol{y}$ with some errors $\boldsymbol{e}$:

$$\boxed{\boldsymbol{y} = \mathcal{H}(\boldsymbol{x}) + \boldsymbol{e}}$$

Inverse problem

What are the best parameters $\boldsymbol{x}$ according to data $\boldsymbol{y}$, the model $\mathcal{H}$ and eventually some priors?

Many parameters (*e.g.* $\gtrsim 10^6$), possibly more parameters than measures: $N_{\boldsymbol{x}} \gtrsim N_{\boldsymbol{y}}$.

In general, inverses problems are ill-conditioned (unstable solution) or ill-posed (several or no solutions).

What is an inverse problem?

We have a model $\mathcal{H}$ linking parameters $\boldsymbol{x}$ and data $\boldsymbol{y}$ with some errors $\boldsymbol{e}$:

$$\boxed{\boldsymbol{y} = \mathcal{H}(\boldsymbol{x}) + \boldsymbol{e}}$$

Inverse problem

What are the best parameters $\boldsymbol{x}$ according to data $\boldsymbol{y}$, the model $\mathcal{H}$ and eventually some priors?

Many parameters (*e.g.* $\gtrsim 10^6$), possibly more parameters than measures: $N_{\boldsymbol{x}} \gtrsim N_{\boldsymbol{y}}$.

In general, inverses problems are ill-conditioned (unstable solution) or ill-posed (several or no solutions).

What is an inverse problem?

We have a model $\mathcal{H}$ linking parameters $\boldsymbol{x}$ and data $\boldsymbol{y}$ with some errors $\boldsymbol{e}$:

$$\boxed{\boldsymbol{y} = \mathcal{H}(\boldsymbol{x}) + \boldsymbol{e}}$$

Inverse problem

What are the best parameters $\boldsymbol{x}$ according to data $\boldsymbol{y}$, the model $\mathcal{H}$ and eventually some priors?

Many parameters (*e.g.* $\gtrsim 10^6$), possibly more parameters than measures: $N_{\boldsymbol{x}} \gtrsim N_{\boldsymbol{y}}$.

In general, inverses problems are ill-conditioned (unstable solution) or ill-posed (several or no solutions).

What is an inverse problem?

We have a model $\mathcal{H}$ linking parameters $\boldsymbol{x}$ and data $\boldsymbol{y}$ with some errors $\boldsymbol{e}$:

$$\boxed{\boldsymbol{y} = \mathcal{H}(\boldsymbol{x}) + \boldsymbol{e}}$$

Inverse problem

What are the best parameters $\boldsymbol{x}$ according to data $\boldsymbol{y}$, the model $\mathcal{H}$ and eventually some priors?

Many parameters (*e.g.* $\gtrsim 10^6$), possibly more parameters than measures: $N_{\boldsymbol{x}} \gtrsim N_{\boldsymbol{y}}$.

In general, inverses problems are ill-conditioned (unstable solution) or ill-posed (several or no solutions).

What is an inverse problem?

We have a model $\mathcal{H}$ linking parameters $\boldsymbol{x}$ and data $\boldsymbol{y}$ with some errors $\boldsymbol{e}$:

$$\boxed{\boldsymbol{y} = \mathcal{H}(\boldsymbol{x}) + \boldsymbol{e}}$$

Inverse problem

What are the best parameters $\boldsymbol{x}$ according to data $\boldsymbol{y}$, the model $\mathcal{H}$ and eventually some priors?

Many parameters (*e.g.* $\gtrsim 10^6$), possibly more parameters than measures: $N_{\boldsymbol{x}} \gtrsim N_{\boldsymbol{y}}$.

In general, inverses problems are ill-conditioned (unstable solution) or ill-posed (several or no solutions).

The deconvolution problem

$$\boldsymbol{y} = \mathbf{H} \cdot \boldsymbol{x} + \boldsymbol{e} \, .$$

The deconvolution problem

$$\boldsymbol{y} = \mathbf{H} \cdot \boldsymbol{x} + \boldsymbol{e} \, .$$

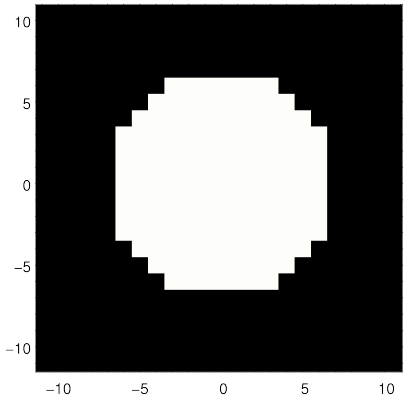
Object

The deconvolution problem

$$\boldsymbol{y} = \mathbf{H} \cdot \boldsymbol{x} + \boldsymbol{e} \, .$$

Object

$*$



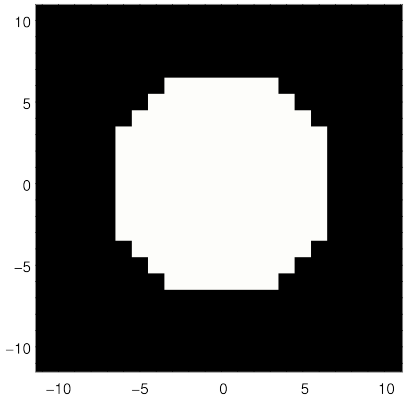
PSF

The deconvolution problem

$$\boldsymbol{y} = \mathbf{H} \cdot \boldsymbol{x} + \boldsymbol{e} \,.$$

Object

$*$



PSF

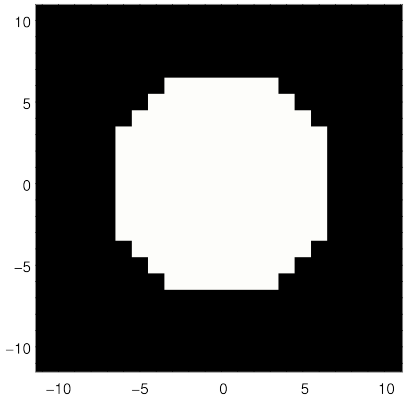
+ Noise

The deconvolution problem

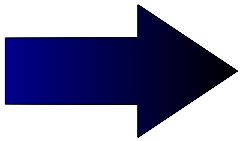
$$y = \mathbf{H} \cdot x + e \, .$$

Object

$*$



PSF

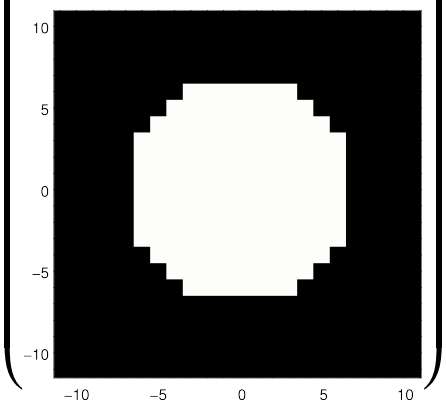


+ Noise

Data

Inversion

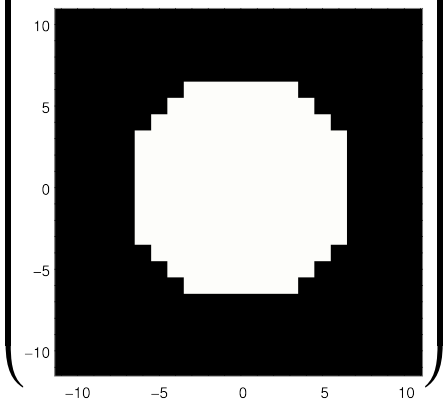
$$\boldsymbol{x} = \mathrm{FFT}^{-1}\left(\frac{\mathrm{FFT}\left(\right.\right.}{\mathrm{FFT}\left(\right.}\right) =$$

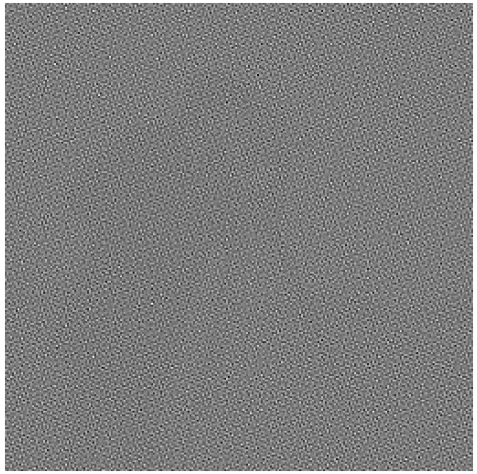


$$\widehat{x}(\boldsymbol{\nu}) = \frac{\widehat{y}(\boldsymbol{\nu})}{\widehat{h}(\boldsymbol{\nu})}$$

Inversion

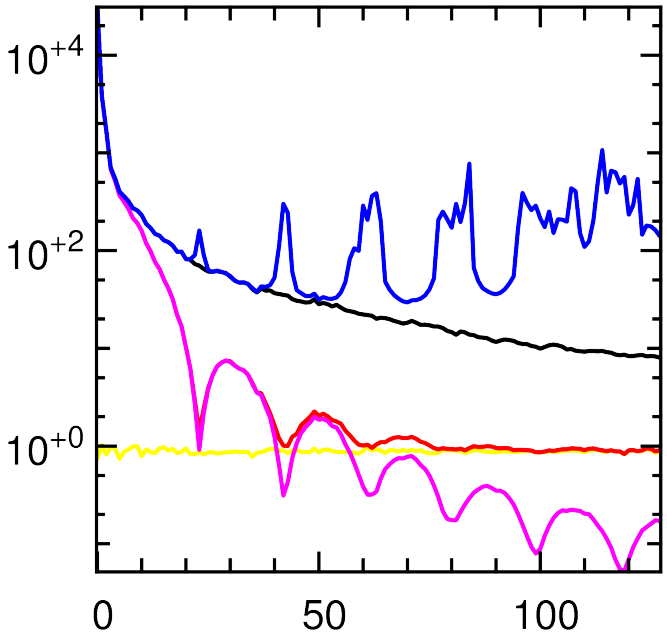
$$\boldsymbol{x} = \mathrm{FFT}^{-1}\left(\frac{\mathrm{FFT}\left(\right.\right.}{\mathrm{FFT}\left(\right.}\right) =$$





$$\widehat{x}(\boldsymbol{\nu}) = \frac{\widehat{y}(\boldsymbol{\nu})}{\widehat{h}(\boldsymbol{\nu})}$$

What is the problem?



power spectra profiles

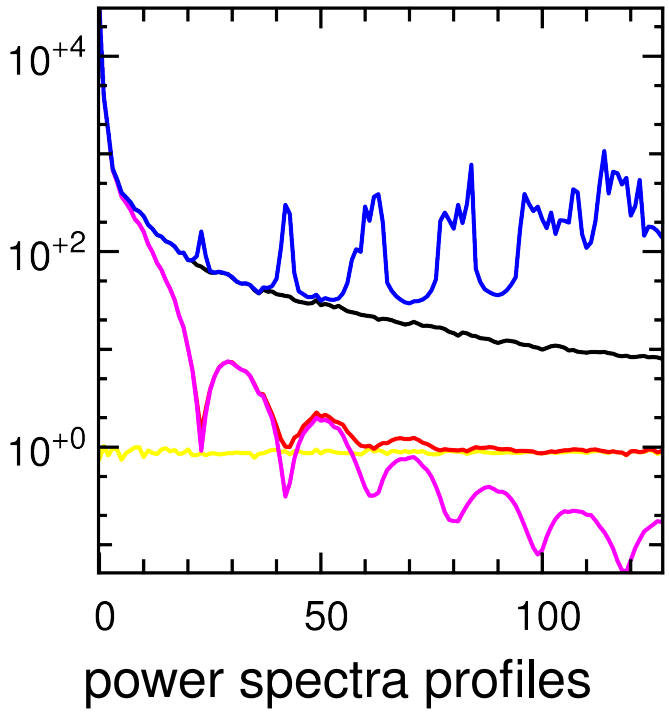
- — object,
- — blurred object,
- — noise,
- — blurred object + noise,
- — direct inversion.

- direct inversion :

$$\widehat{x}_{\mathrm{ML}}(\nu) = \frac{\widehat{y}(\nu)}{\widehat{h}(\nu)} = \widehat{x}(\nu) + \frac{\widehat{e}(\nu)}{\widehat{h}(\nu)}$$

- Noise amplification. It is an ill conditioned problem.

What is the problem?



power spectra profiles

- — object,
- — blurred object,
- — noise,
- — blurred object + noise,
- — direct inversion.

- direct inversion :

$$\widehat{x}_{\mathrm{ML}}(\nu) = \frac{\widehat{y}(\nu)}{\widehat{h}(\nu)} = \widehat{x}(\nu) + \frac{\widehat{e}(\nu)}{\widehat{h}(\nu)}$$

- Noise amplification. It is an ill conditioned problem.

Maximum Likelihood

$$\boxed{\boldsymbol{x}_{\mathrm{ML}} = \arg\max_{\boldsymbol{x}} \Pr(\boldsymbol{y}|\boldsymbol{x})}$$

- Linear model, Gaussian noise:

$$\begin{aligned} \boldsymbol{x}_{\mathrm{ML}} &= \arg\min_{\boldsymbol{x}} f_{\mathrm{data}}(\boldsymbol{x}) \\ f_{\mathrm{data}}(\boldsymbol{x}) &= (\boldsymbol{y} - \mathbf{H}\cdot\boldsymbol{x})^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot (\boldsymbol{y} - \mathbf{H}\cdot\boldsymbol{x}) \end{aligned}$$

- Optimality conditions:

$$\begin{aligned} & \left.\frac{\partial f_{\mathrm{data}}(\boldsymbol{x})}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=\boldsymbol{x}_{\mathrm{ML}}} = \mathbf{0} \\ \Longleftrightarrow \quad & \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H} \cdot \boldsymbol{x}_{\mathrm{ML}} = \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y} \\ \Longleftrightarrow \quad & \boxed{\boldsymbol{x}_{\mathrm{ML}} = \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H}\right)^{-1} \cdot \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y}} \end{aligned}$$

Maximum likelihood

- Circulant approximation: $\mathbf{F} \cdot \mathbf{H} = \mathrm{diag}(\hat{\boldsymbol{h}})$.
- Gaussian white noise :$\mathbf{C}_{\boldsymbol{e}} = \sigma_e \mathbf{I}$.
-

$$x^{\mathrm{MV}} = \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{H}\right)^{-1} \cdot \mathbf{H}^{\mathrm{T}} \cdot \boldsymbol{y} \,,$$

$$\hat{x}_u^{(\mathrm{MV})} = \frac{\hat{h}_u^* \hat{y}_u}{|\hat{h}_u|^2} = \frac{\hat{y}_u}{\hat{h}_u} \,,$$

- It is direct inversion
- The best (minimum variance) solution without bias.

Maximum likelihood

- Circulant approximation: $\mathbf{F} \cdot \mathbf{H} = \mathrm{diag}(\hat{\boldsymbol{h}})$.
- Gaussian white noise :$\mathbf{C}_{\boldsymbol{e}} = \sigma_e \mathbf{I}$.
-

$$\begin{aligned} x^{\mathrm{MV}} &= \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{H}\right)^{-1} \cdot \mathbf{H}^{\mathrm{T}} \cdot \boldsymbol{y}\,, \\ \hat{x}_u^{(\mathrm{MV})} &= \frac{\hat{h}_u^* \hat{y}_u}{|\hat{h}_u|^2} = \frac{\hat{y}_u}{\hat{h}_u}\,, \end{aligned}$$

- It is direct inversion
- The best (minimum variance) solution without bias.

Maximum likelihood

- Circulant approximation: $\mathbf{F} \cdot \mathbf{H} = \mathrm{diag}(\hat{\boldsymbol{h}})$.
- Gaussian white noise :$\mathbf{C}_{\boldsymbol{e}} = \sigma_e \mathbf{I}$.
-

$$\begin{aligned} x^{\mathrm{MV}} &= \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{H}\right)^{-1} \cdot \mathbf{H}^{\mathrm{T}} \cdot \boldsymbol{y}\,, \\ \hat{x}_u^{(\mathrm{MV})} &= \frac{\hat{h}_u^* \hat{y}_u}{|\hat{h}_u|^2} = \frac{\hat{y}_u}{\hat{h}_u}\,, \end{aligned}$$

- It is direct inversion
- The best (minimum variance) solution without bias.

Restricting the parameters space

Cutting high frequency ($u_c = 40$ frequels)

$\mathrm{MSE} = 27\,\mathrm{dB}$

Restricting the parameters space

Cutting high frequency ($u_c = 40$ frequels)
$\mathrm{MSE} = 27\,\mathrm{dB}$

Positivity constraints (ISRA Daube-Witherspoon and Muehllehner (1986)) $\mathrm{MSE} = 38.8\,\mathrm{dB}$
Iterative method.

Richardson-Lucy and Expectation Maximization algorithm

Maximum likelihood method based on a Poisson noise model

$$\boldsymbol{y} = \mathcal{P}(\mathbf{H} \cdot \boldsymbol{x})$$

Poisson likelihood:

$$\Pr(\boldsymbol{y}\,|\,\boldsymbol{x}) = \prod_k \frac{(\mathbf{H} \cdot \boldsymbol{x})_k^{y_k} \, \exp\left(-(\boldsymbol{H} \cdot \boldsymbol{x})_k\right)}{y_k\,!} \,.$$

Maximum likelihood:

$$\boldsymbol{x}^+ = \arg\max_{\boldsymbol{x}} \log\left(\Pr(\boldsymbol{y}\,|\,\boldsymbol{x})\right)$$

Derivatives vanish at maximum :

$$\nabla_x \log\left(\Pr(\boldsymbol{y}\,|\,\boldsymbol{x})\right) = \mathbf{H}^{\mathrm{T}} \operatorname{diag}(\mathbf{H} \cdot \boldsymbol{x})^{-1} \boldsymbol{y} - \mathbf{H}^{\mathrm{T}} \mathbf{1} = 0$$

Fixed point gives the following multiplicative update:

$$\boldsymbol{x}^{(k+1)} = \frac{\boldsymbol{x}^{(k)}}{\mathbf{H}^{\mathrm{T}} \cdot \mathbf{1}} \left[\mathbf{H}^{\mathrm{T}} \cdot \left(\frac{\boldsymbol{y}}{\mathbf{H} \cdot \boldsymbol{x}^{(k)}} \right) \right]$$

Richardson-Lucy and Expectation Maximization algorithm

Converge toward the maximum likelihood solution!

Three possibilities:

1. stop iterations early (as RLA have a slow convergence rate),
2. apply extra-constraints using projection $\mathcal{P}_c$:

$$\boldsymbol{x}^{(k+1)} = \frac{\boldsymbol{x}^{(k)}}{\mathbf{H}^{\mathrm{T}} \cdot \mathbf{1}} \mathcal{P}_c \left[\mathbf{H} \cdot \left(\frac{\boldsymbol{y}}{\mathbf{H} \cdot \boldsymbol{x}^{(k)}} \right) \right],$$

3. non-linear residual denoising [Starck *et al.*, 1994, 2002]:
   RLA rewritten introducing model $\boldsymbol{m}^{(k)} = \mathbf{H} \cdot \boldsymbol{x}^{(k)}$ and residuals $\boldsymbol{r}^{(k)} = \boldsymbol{m}^{(k)} - \boldsymbol{y}$:

$$\boldsymbol{x}^{(k+1)} = \frac{\boldsymbol{x}^{(k)}}{\mathbf{H}^{\mathrm{T}} \cdot \mathbf{1}} \left[\mathbf{H} \cdot \left(\frac{\boldsymbol{m}^{(k)} - \boldsymbol{r}^{(k)}}{\boldsymbol{m}^{(k)}} \right) \right],$$

   With non-linear denoising operator $\mathcal{D}$:

$$\boldsymbol{x}^{(k+1)} = \frac{\boldsymbol{x}^{(k)}}{\mathbf{H}^{\mathrm{T}} \cdot \mathbf{1}} \left[\mathbf{H} \cdot \left(\frac{\boldsymbol{m}^{(k)} - \mathcal{D}(\boldsymbol{r}^{(k)})}{\boldsymbol{m}^{(k)}} \right) \right],$$

Richardson-Lucy and Expectation Maximization algorithm

Converge toward the maximum likelihood solution!

Three possibilities:

1. stop iterations early (as RLA have a slow convergence rate),
2. apply extra-constraints using projection $\mathcal{P}_c$:

$$\boldsymbol{x}^{(k+1)} = \frac{\boldsymbol{x}^{(k)}}{\mathbf{H}^{\mathrm{T}} \cdot \mathbf{1}} \mathcal{P}_c \left[\mathbf{H} \cdot \left(\frac{\boldsymbol{y}}{\mathbf{H} \cdot \boldsymbol{x}^{(k)}} \right) \right],$$

3. non-linear residual denoising [Starck *et al.*, 1994, 2002]:
   RLA rewritten introducing model $\boldsymbol{m}^{(k)} = \mathbf{H} \cdot \boldsymbol{x}^{(k)}$ and residuals $\boldsymbol{r}^{(k)} = \boldsymbol{m}^{(k)} - \boldsymbol{y}$:

$$\boldsymbol{x}^{(k+1)} = \frac{\boldsymbol{x}^{(k)}}{\mathbf{H}^{\mathrm{T}} \cdot \mathbf{1}} \left[\mathbf{H} \cdot \left(\frac{\boldsymbol{m}^{(k)} - \boldsymbol{r}^{(k)}}{\boldsymbol{m}^{(k)}} \right) \right],$$

   With non-linear denoising operator $\mathcal{D}$:

$$\boldsymbol{x}^{(k+1)} = \frac{\boldsymbol{x}^{(k)}}{\mathbf{H}^{\mathrm{T}} \cdot \mathbf{1}} \left[\mathbf{H} \cdot \left(\frac{\boldsymbol{m}^{(k)} - \mathcal{D}(\boldsymbol{r}^{(k)})}{\boldsymbol{m}^{(k)}} \right) \right],$$

Richardson-Lucy and Expectation Maximization algorithm

Converge toward the maximum likelihood solution!

Three possibilities:

1. stop iterations early (as RLA have a slow convergence rate),
2. apply extra-constraints using projection $\mathcal{P}_c$:

$$\boldsymbol{x}^{(k+1)} = \frac{\boldsymbol{x}^{(k)}}{\mathbf{H}^{\mathrm{T}} \cdot \mathbf{1}} \mathcal{P}_c \left[\mathbf{H} \cdot \left(\frac{\boldsymbol{y}}{\mathbf{H} \cdot \boldsymbol{x}^{(k)}} \right) \right],$$

3. non-linear residual denoising [Starck *et al.*, 1994, 2002]:
   RLA rewritten introducing model $\boldsymbol{m}^{(k)} = \mathbf{H} \cdot \boldsymbol{x}^{(k)}$ and residuals $\boldsymbol{r}^{(k)} = \boldsymbol{m}^{(k)} - \boldsymbol{y}$:

$$\boldsymbol{x}^{(k+1)} = \frac{\boldsymbol{x}^{(k)}}{\mathbf{H}^{\mathrm{T}} \cdot \mathbf{1}} \left[\mathbf{H} \cdot \left(\frac{\boldsymbol{m}^{(k)} - \boldsymbol{r}^{(k)}}{\boldsymbol{m}^{(k)}} \right) \right],$$

   With non-linear denoising operator $\mathcal{D}$:

$$\boldsymbol{x}^{(k+1)} = \frac{\boldsymbol{x}^{(k)}}{\mathbf{H}^{\mathrm{T}} \cdot \mathbf{1}} \left[\mathbf{H} \cdot \left(\frac{\boldsymbol{m}^{(k)} - \mathcal{D}(\boldsymbol{r}^{(k)})}{\boldsymbol{m}^{(k)}} \right) \right],$$

Richardson-Lucy and Expectation Maximization algorithm

Converge toward the maximum likelihood solution!

Three possibilities:

1. stop iterations early (as RLA have a slow convergence rate),
2. apply extra-constraints using projection $\mathcal{P}_c$:

$$\boldsymbol{x}^{(k+1)} = \frac{\boldsymbol{x}^{(k)}}{\mathbf{H}^{\mathrm{T}} \cdot \mathbf{1}} \mathcal{P}_c \left[\mathbf{H} \cdot \left(\frac{\boldsymbol{y}}{\mathbf{H} \cdot \boldsymbol{x}^{(k)}} \right) \right],$$

3. non-linear residual denoising [Starck *et al.*, 1994, 2002]:
   RLA rewritten introducing model $\boldsymbol{m}^{(k)} = \mathbf{H} \cdot \boldsymbol{x}^{(k)}$ and residuals $\boldsymbol{r}^{(k)} = \boldsymbol{m}^{(k)} - \boldsymbol{y}$:

$$\boldsymbol{x}^{(k+1)} = \frac{\boldsymbol{x}^{(k)}}{\mathbf{H}^{\mathrm{T}} \cdot \mathbf{1}} \left[\mathbf{H} \cdot \left(\frac{\boldsymbol{m}^{(k)} - \boldsymbol{r}^{(k)}}{\boldsymbol{m}^{(k)}} \right) \right],$$

   With non-linear denoising operator $\mathcal{D}$:

$$\boldsymbol{x}^{(k+1)} = \frac{\boldsymbol{x}^{(k)}}{\mathbf{H}^{\mathrm{T}} \cdot \mathbf{1}} \left[\mathbf{H} \cdot \left(\frac{\boldsymbol{m}^{(k)} - \mathcal{D}(\boldsymbol{r}^{(k)})}{\boldsymbol{m}^{(k)}} \right) \right],$$

Maximum A Posteriori

$$\boldsymbol{x}^{+} = \arg\min_{\boldsymbol{x}} f_{\text{post}}(\boldsymbol{x}) \,,$$

- $f_{\text{post}}(\boldsymbol{x})$ = ***a posteriori*** cost function:

$$f_{\text{post}}(\boldsymbol{x}) = f_{\text{data}}(\boldsymbol{x}) + f_{\text{prior}}(\boldsymbol{x})$$

- $f_{\text{data}}(\boldsymbol{x})$ = ***likelihood*** term (Gaussian case):

$$f_{\text{data}}(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{y} - \mathbf{H} \cdot \boldsymbol{x})^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{e}}^{-1} \cdot (\boldsymbol{y} - \mathbf{H} \cdot \boldsymbol{x}) \,.$$

- $f_{\text{prior}}(\boldsymbol{x})$ = ***regularization*** term (a priori):

$$f_{\text{prior}}(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^{\mathrm{T}} \cdot \mathbf{C}_{\text{prior}}^{-1} \cdot \boldsymbol{x} \,.$$

Maximum A Posteriori

$$\boldsymbol{x}^{+} = \arg\min_{\boldsymbol{x}} f_{\text{post}}(\boldsymbol{x}) \,,$$

- $f_{\text{post}}(\boldsymbol{x})$ = ***a posteriori*** cost function:

$$f_{\text{post}}(\boldsymbol{x}) = f_{\text{data}}(\boldsymbol{x}) + f_{\text{prior}}(\boldsymbol{x})$$

- $f_{\text{data}}(\boldsymbol{x})$ = ***likelihood*** term (Gaussian case):

$$f_{\text{data}}(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{y} - \mathbf{H} \cdot \boldsymbol{x})^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{e}}^{-1} \cdot (\boldsymbol{y} - \mathbf{H} \cdot \boldsymbol{x}) \,.$$

- $f_{\text{prior}}(\boldsymbol{x})$ = ***regularization*** term (a priori):

$$f_{\text{prior}}(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^{\mathrm{T}} \cdot \mathbf{C}_{\text{prior}}^{-1} \cdot \boldsymbol{x} \,.$$

Maximum A Posteriori (Gaussian case)

$$\boldsymbol{x}_{\mathrm{MAP}} = \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{e}}^{-1} \cdot \mathbf{H} + \mathbf{C}_{\mathrm{prior}}^{-1}\right)^{-1} \cdot \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{e}}^{-1} \cdot \boldsymbol{y}\,.$$

$$\mathbf{C}_{\mathrm{post}} = \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{e}}^{-1} \cdot \mathbf{H} + \mathbf{C}_{\mathrm{prior}}^{-1}\right)^{-1}\,.$$

$$\mathbf{C}_{\mathrm{post}} \leq \mathbf{C}_{\mathrm{prior}}$$ Better knowledge than before measurements

- If $\mathbf{C}_{\mathrm{prior}}$, $\mathbf{C}_{\mathrm{e}}$ and $\mathbf{H}$ are circulant (BTTB) and $\widehat{\mathbf{C}_{\mathrm{prior}}} = \mathrm{diag}\left|\widehat{x}\right|^2$, $\widehat{\mathbf{C}_{\mathrm{e}}} = \mathrm{diag}\left|\widehat{e}\right|^2$

$$\widehat{x}_u^{(\mathrm{MAP})} = \frac{\widehat{h}_u^* \widehat{y}_u}{|\widehat{h}_u|^2 + \dfrac{\left|\widehat{x}\right|_u^2}{\left|\widehat{e}\right|^2}}\,.$$

- This is the Wiener filter, in Gaussian case MAP = MMSE.

Maximum A Posteriori (Gaussian case)

$$\boldsymbol{x}_{\mathrm{MAP}} = \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{e}}^{-1} \cdot \mathbf{H} + \mathbf{C}_{\mathrm{prior}}^{-1}\right)^{-1} \cdot \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{e}}^{-1} \cdot \boldsymbol{y}\,.$$

$$\mathbf{C}_{\mathrm{post}} = \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{e}}^{-1} \cdot \mathbf{H} + \mathbf{C}_{\mathrm{prior}}^{-1}\right)^{-1}.$$

$$\mathbf{C}_{\mathrm{post}} \leq \mathbf{C}_{\mathrm{prior}}$$ Better knowledge than before measurements

- If $\mathbf{C}_{\mathrm{prior}}$, $\mathbf{C}_{\mathrm{e}}$ and $\mathbf{H}$ are circulant (BTTB) and $\widehat{\mathbf{C}_{\mathrm{prior}}} = \mathrm{diag}\left|\widehat{x}\right|^2$, $\widehat{\mathbf{C}_{\mathrm{e}}} = \mathrm{diag}\left|\widehat{e}\right|^2$

$$\widehat{x}_u^{(\mathrm{MAP})} = \frac{\widehat{h}_u^* \widehat{y}_u}{|\widehat{h}_u|^2 + \dfrac{\left|\widehat{x}\right|_u^2}{\left|\widehat{e}\right|^2}}\,.$$

- This is the Wiener filter, in Gaussian case MAP = MMSE.

Maximum A Posteriori (Gaussian case)

$$
\begin{aligned}
\boldsymbol{x}_{\mathrm{MAP}} &= \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{e}}^{-1} \cdot \mathbf{H} + \mathbf{C}_{\mathrm{prior}}^{-1}\right)^{-1} \cdot \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{e}}^{-1} \cdot \boldsymbol{y}\,. \\
\mathbf{C}_{\mathrm{post}} &= \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{e}}^{-1} \cdot \mathbf{H} + \mathbf{C}_{\mathrm{prior}}^{-1}\right)^{-1}\,. \\
\mathbf{C}_{\mathrm{post}} &\leq \mathbf{C}_{\mathrm{prior}}
\end{aligned}
$$

Better knowledge than before measurements

- If $\mathbf{C}_{\mathrm{prior}}$, $\mathbf{C}_{\mathrm{e}}$ and $\mathbf{H}$ are circulant (BTTB) and $\widehat{\mathbf{C}_{\mathrm{prior}}} = \mathrm{diag}\left|\widehat{x}\right|^2$, $\widehat{\mathbf{C}_{\mathrm{e}}} = \mathrm{diag}\left|\widehat{e}\right|^2$

$$
\widehat{x}_u^{(\mathrm{MAP})} = \frac{\widehat{h}_u^* \widehat{y}_u}{|\widehat{h}_u|^2 + \dfrac{\left|\widehat{x}\right|_u^2}{\left|\widehat{e}\right|^2}}\,.
$$

- This is the Wiener filter, in Gaussian case MAP = MMSE.

Maximum A Posteriori (Gaussian case)

$$
\begin{aligned}
\boldsymbol{x}_{\mathrm{MAP}} &= \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{e}}^{-1} \cdot \mathbf{H} + \mathbf{C}_{\mathrm{prior}}^{-1}\right)^{-1} \cdot \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{e}}^{-1} \cdot \boldsymbol{y}\,. \\
\mathbf{C}_{\mathrm{post}} &= \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{e}}^{-1} \cdot \mathbf{H} + \mathbf{C}_{\mathrm{prior}}^{-1}\right)^{-1}\,. \\
\mathbf{C}_{\mathrm{post}} &\leq \mathbf{C}_{\mathrm{prior}} \text{Better knowledge than before measurements}
\end{aligned}
$$

- If $\mathbf{C}_{\mathrm{prior}}$, $\mathbf{C}_{\mathrm{e}}$ and $\mathbf{H}$ are circulant (BTTB) and $\widehat{\mathbf{C}_{\mathrm{prior}}} = \mathrm{diag}\left|\widehat{x}\right|^2$, $\widehat{\mathbf{C}_{\mathrm{e}}} = \mathrm{diag}\left|\widehat{e}\right|^2$

$$
\widehat{x}_u^{(\mathrm{MAP})} = \frac{\widehat{h}_u^* \widehat{y}_u}{|\widehat{h}_u|^2 + \dfrac{\left|\widehat{x}\right|_u^2}{\left|\widehat{e}\right|^2}}\,.
$$

- This is the Wiener filter, in Gaussian case MAP = MMSE.

Maximum A Posteriori (Gaussian case)

$$
\begin{aligned}
\boldsymbol{x}_{\mathrm{MAP}} &= \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{e}}^{-1} \cdot \mathbf{H} + \mathbf{C}_{\mathrm{prior}}^{-1}\right)^{-1} \cdot \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{e}}^{-1} \cdot \boldsymbol{y}\,. \\
\mathbf{C}_{\mathrm{post}} &= \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\mathrm{e}}^{-1} \cdot \mathbf{H} + \mathbf{C}_{\mathrm{prior}}^{-1}\right)^{-1}\,. \\
\mathbf{C}_{\mathrm{post}} &\leq \mathbf{C}_{\mathrm{prior}} \text{Better knowledge than before measurements}
\end{aligned}
$$

- If $\mathbf{C}_{\mathrm{prior}}$, $\mathbf{C}_{\mathrm{e}}$ and $\mathbf{H}$ are circulant (BTTB) and $\widehat{\mathbf{C}_{\mathrm{prior}}} = \mathrm{diag}\left|\widehat{x}\right|^2$, $\widehat{\mathbf{C}_{\mathrm{e}}} = \mathrm{diag}\left|\widehat{e}\right|^2$

$$
\widehat{x}_u^{(\mathrm{MAP})} = \frac{\widehat{h}_u^* \widehat{y}_u}{|\widehat{h}_u|^2 + \dfrac{\left|\widehat{x}\right|_u^2}{\left|\widehat{e}\right|^2}}\,.
$$

- This is the Wiener filter, in Gaussian case MAP = MMSE.

Smoothness quadratic prior

$$f_{\text{prior}}(\boldsymbol{x}) = \mu \, \|\mathbf{D} \cdot \boldsymbol{x}\|_2^2$$

where $\mathbf{D}$ is a smoothing kernel.

Can be rewritten in:

$$f_{\text{prior}}(\boldsymbol{x}) = \mu \, \boldsymbol{x}^{\mathrm{T}} \cdot \mathbf{D}^{\mathrm{T}} \cdot \mathbf{D} \cdot \boldsymbol{x} \, .$$

where we can identify :

$$\mathbf{C}_{\text{prior}}^{-1} = \mu \, \mathbf{D}^{\mathrm{T}} \cdot \mathbf{D}$$

Can be rewritten as a Wiener-like filter

Restoration with quadratic regularization $\mathrm{MSE} = 23.1\,\mathrm{dB}$

Smoothness quadratic prior

$$f_{\text{prior}}(\boldsymbol{x}) = \mu \, \|\mathbf{D} \cdot \boldsymbol{x}\|_2^2$$

where $\mathbf{D}$ is a smoothing kernel.
Can be rewritten in:

$$f_{\text{prior}}(\boldsymbol{x}) = \mu \, \boldsymbol{x}^{\mathrm{T}} \cdot \mathbf{D}^{\mathrm{T}} \cdot \mathbf{D} \cdot \boldsymbol{x} \, .$$

where we can identify :

$$\mathbf{C}_{\text{prior}}^{-1} = \mu \, \mathbf{D}^{\mathrm{T}} \cdot \mathbf{D}$$

Can be rewritten as a Wiener-like filter

Restoration with quadratic regularization $\mathrm{MSE} = 23.1\,\mathrm{dB}$

Maximum a posteriori: non linear regularization

A vast literature about non linear regularization.

$f_{\text{prior}}(\boldsymbol{x}) =$

- Total variation,
- ℓ_1, $\ell_2 - \ell_1$,
- synthesis priors,
- ...

Good properties:

- convexity,
- derivability,
- closed form solution,
- fast computation.

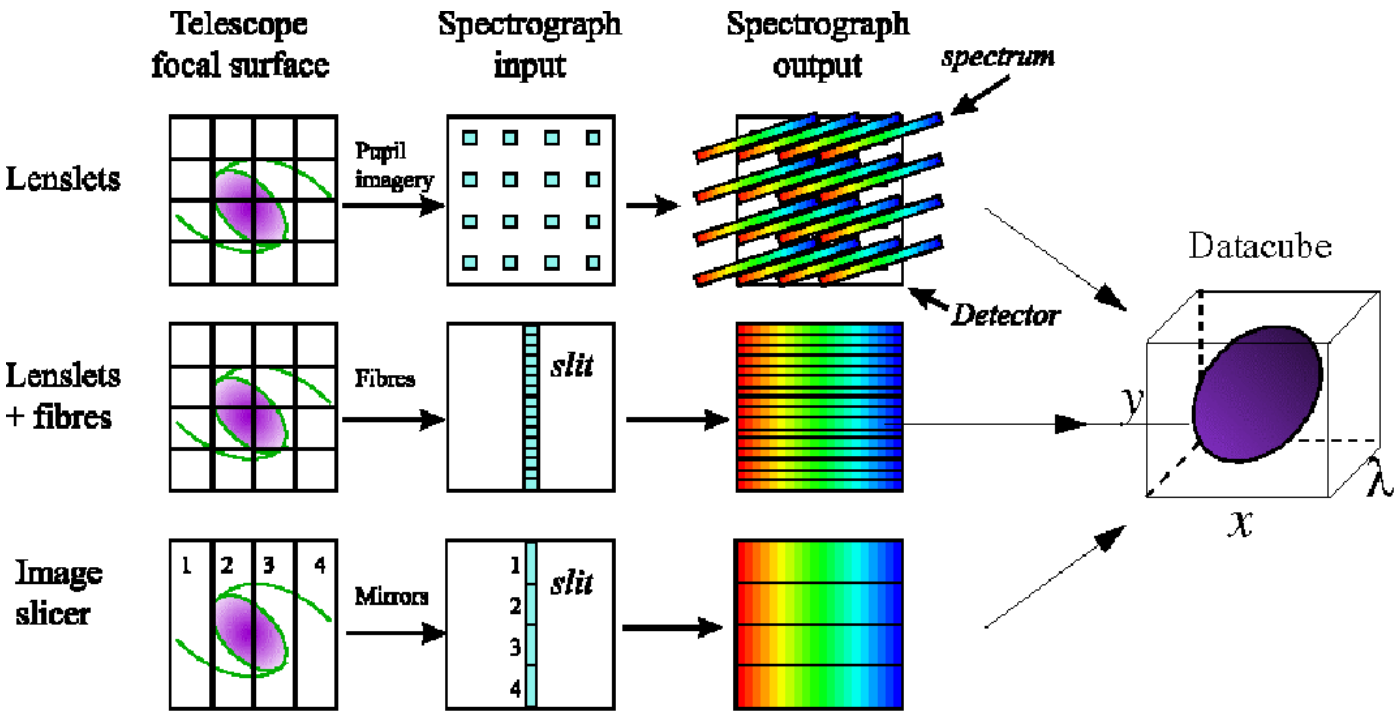
Restoration with TV $\mathrm{MSE} = 21\,\mathrm{dB}$

Inverse problem approach Vs direct inversion

- no coarse approximations (needed to build $\mathcal{H}^{-1}$),
- estimation of non measured areas (border, saturation,...),
- less pre-processing steps,
- priors adapted to the observed objects.

Deconvolution in Integral Field Spectrography

Integral field spectrography in astronomy



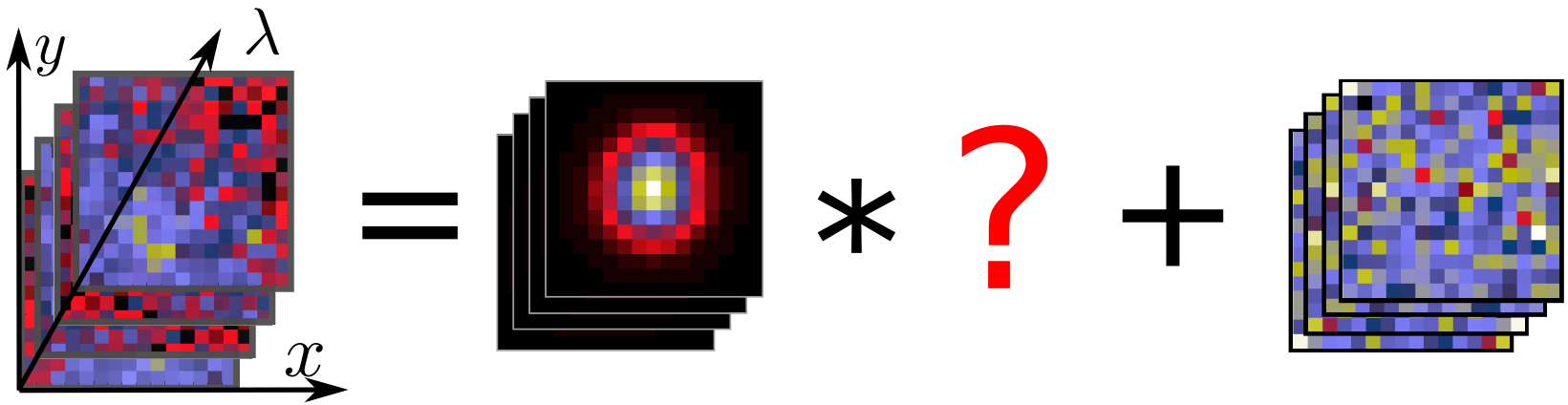
IFU

- Narrow field of view,
- Large number of spectral channels.

Astrophysical data

- Narrow emission and absorption lines,
- Large dynamics,
- Few sources.

Deconvolution



Deconvolution

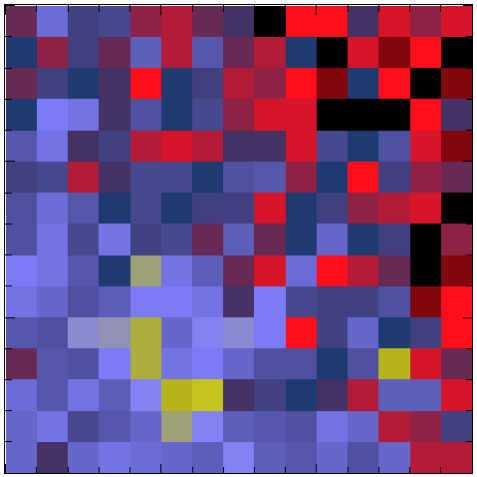
Estimation of the unblurred scene with a known PSF.

Previous work

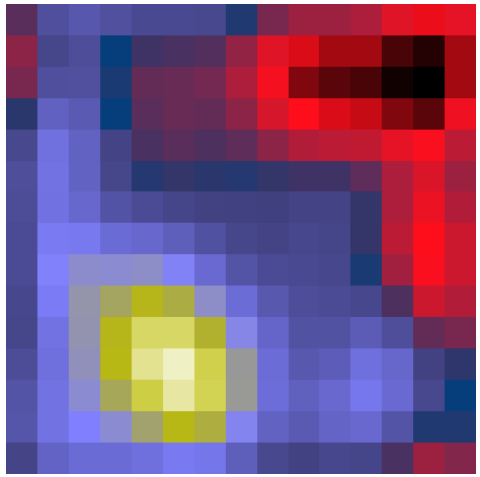
- Inverse filtering (Wiener) (Hunt and Kubler 1984; Galatsanos and Chin 1989),
- Fourier / Wavelet filtering (Benazza-Benyahia and Pesquet 2006; Duijster et al. 2009),
- Maximum a posteriori (Bedet et al. 2008)

Small FoV and boundary artifacts (illustration)

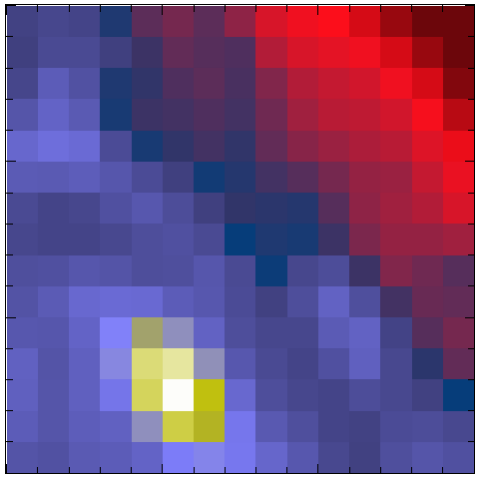
Narrow field of view causes important boundary artifacts



Data (3975Å)



Richardson-Lucy



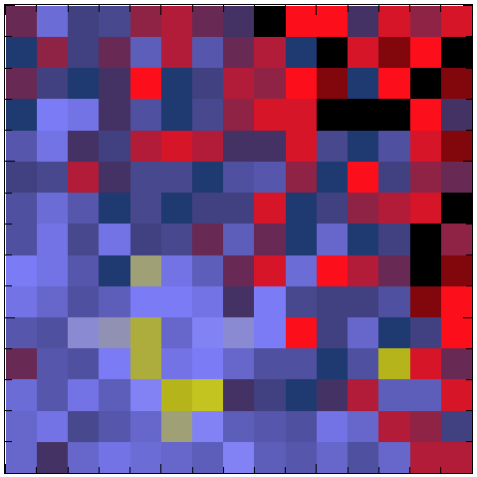
Truth

Illustration using the SuperNovae Integral Field Spectrograph simulation

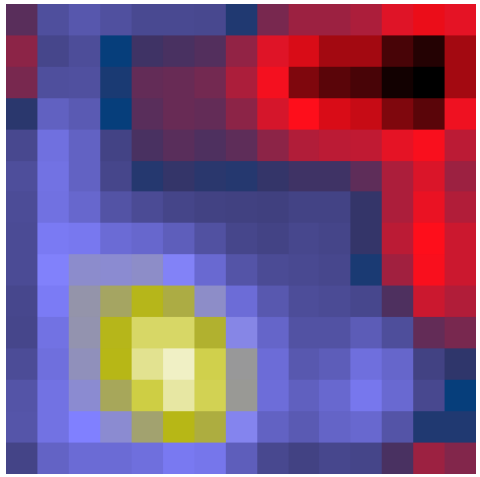
Inverse filtering impossible.

Small FoV and boundary artifacts (illustration)

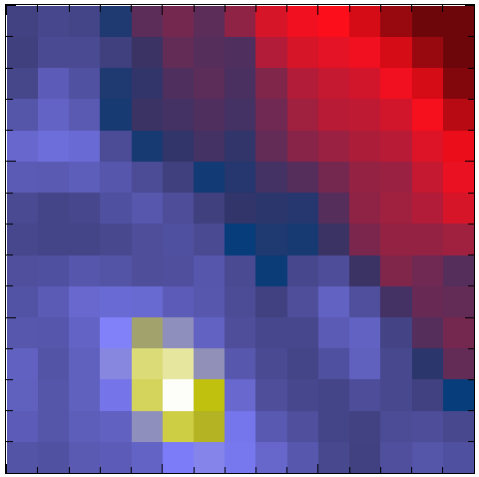
Narrow field of view causes important boundary artifacts



Data (3975Å)



Richardson-Lucy



Truth

Illustration using the SuperNovae Integral Field Spectrograph simulation

Inverse filtering impossible.

Maximum a posteriori approach

Estimation of the most probable object according to data and some priors:

- Few objects on a dark background,
- Similar spatial structure from one channel to its neighbor,
- Narrow emission and absorption lines over a relatively smooth continuum,
- Positivity.

Spectrally varying spatio-spectral PSF:

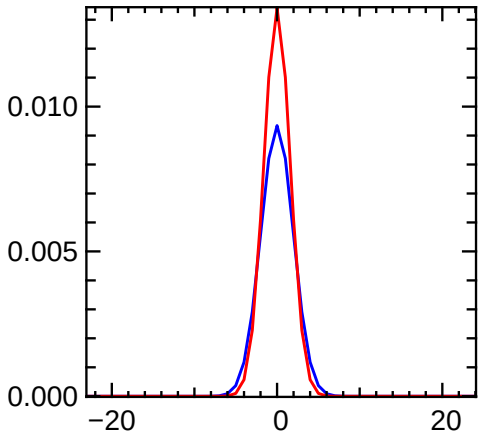
Field Spread Function Line Spread Function

Maximum a posteriori approach

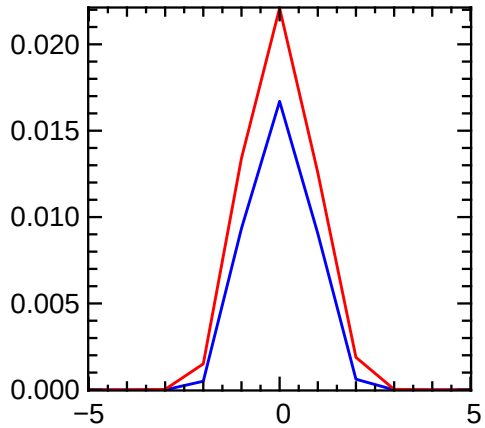
Estimation of the most probable object according to data and some priors:

- Few objects on a dark background,
- Similar spatial structure from one channel to its neighbor,
- Narrow emission and absorption lines over a relatively smooth continuum,
- Positivity.

Spectrally varying spatio-spectral PSF:



Field Spread Function



Line Spread Function

Spatio-spectral regularization

$$f_{\text{prior}}(\boldsymbol{x}) = \Omega_{\text{spatial}}(\boldsymbol{x}) + \Omega_{\text{spectral}}(\boldsymbol{x})$$

Spatial sparsity inducing prior: structured norm [Fornassier (2008)]

$$\Omega_{\text{spatial}}(\boldsymbol{x}) = \mu_{sparsity} \sum_k \sqrt{\sum_\lambda x_{k,\lambda}^2}$$

Ensure spatial sparsity and favour spatial alignment of spectral features.

Spectral prior: Edge preserving spectral smoothing prior

$$\Omega_{\text{spectral}}(\boldsymbol{x}) = \mu_{spectral} \sum_k \sum_\lambda \sqrt{(x_{k,\lambda} - x_{k,\lambda-1})^2 + \epsilon^2},$$

Spatio-spectral regularization

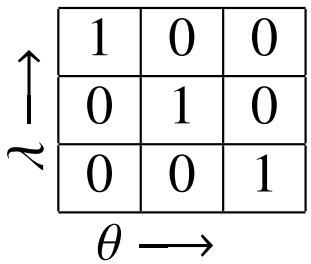
$$f_{\text{prior}}(\boldsymbol{x}) = \Omega_{\text{spatial}}(\boldsymbol{x}) + \Omega_{\text{spectral}}(\boldsymbol{x})$$

Spatial sparsity inducing prior: structured norm [Fornassier (2008)]

$$\Omega_{\text{spatial}}(\boldsymbol{x}) = \mu_{sparsity} \sum_k \sqrt{\sum_\lambda x_{k,\lambda}^2}$$

Ensure spatial sparsity and favour spatial alignment of spectral features.

example with 3 pixels $\times$ 3 wavelengths:



$\Omega_{\text{spatial}}(\boldsymbol{x}) = 3$

$1 + \sqrt{2} \simeq 2.41$

$\sqrt{3} \simeq 1.73$

Spatio-spectral regularization

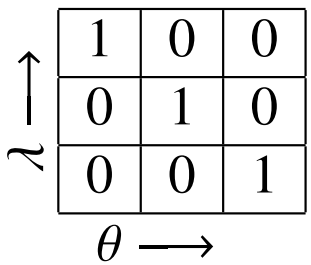
$$f_{\text{prior}}(\boldsymbol{x}) = \Omega_{\text{spatial}}(\boldsymbol{x}) + \Omega_{\text{spectral}}(\boldsymbol{x})$$

Spatial sparsity inducing prior: structured norm [Fornassier (2008)]

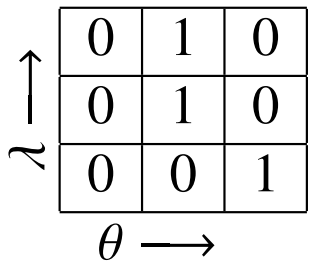
$$\Omega_{\text{spatial}}(\boldsymbol{x}) = \mu_{sparsity} \sum_k \sqrt{\sum_\lambda x_{k,\lambda}^2}$$

Ensure spatial sparsity and favour spatial alignment of spectral features.

example with 3 pixels $\times$ 3 wavelengths:



$\Omega_{\text{spatial}}(\boldsymbol{x}) = 3$



$1 + \sqrt{2} \simeq 2.41$

$\sqrt{3} \simeq 1.73$

Spatio-spectral regularization

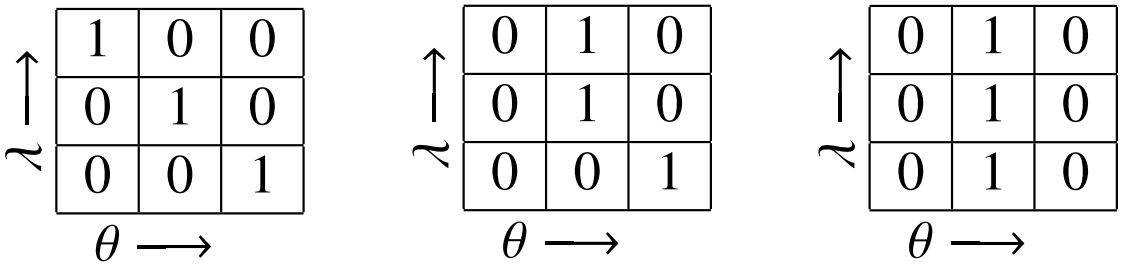
$$f_{\text{prior}}(\boldsymbol{x}) = \Omega_{\text{spatial}}(\boldsymbol{x}) + \Omega_{\text{spectral}}(\boldsymbol{x})$$

Spatial sparsity inducing prior: structured norm [Fornassier (2008)]

$$\Omega_{\text{spatial}}(\boldsymbol{x}) = \mu_{sparsity} \sum_k \sqrt{\sum_\lambda x_{k,\lambda}^2}$$

Ensure spatial sparsity and favour spatial alignment of spectral features.

example with 3 pixels $\times$ 3 wavelengths:



$\Omega_{\text{spatial}}(\boldsymbol{x}) = 3$ $\quad 1 + \sqrt{2} \simeq 2.41$ $\quad \sqrt{3} \simeq 1.73$

Spatio-spectral regularization

$$f_{\text{prior}}(\boldsymbol{x}) = \Omega_{\text{spatial}}(\boldsymbol{x}) + \Omega_{\text{spectral}}(\boldsymbol{x})$$

Spatial sparsity inducing prior: structured norm [Fornassier (2008)]

$$\Omega_{\text{spatial}}(\boldsymbol{x}) = \mu_{sparsity} \sum_{k} \sqrt{\sum_{\lambda} x_{k,\lambda}^2}$$

Ensure spatial sparsity and favour spatial alignment of spectral features.

Spectral prior: Edge preserving spectral smoothing prior

$$\Omega_{\text{spectral}}(\boldsymbol{x}) = \mu_{spectral} \sum_{k} \sum_{\lambda} \sqrt{(x_{k,\lambda} - x_{k,\lambda-1})^2 + \epsilon^2},$$

Spectral renormalization

Large dynamical range between spectral channels lead to:

- Over-regularization of bright features,
- Under-regularization of fainter ones.

Renormalization: Regularization on a spectrally flattened object

$$f_{\text{prior}}(\boldsymbol{x}) = \Omega_{\text{spectral}}(\boldsymbol{x}') + \Omega_{\text{spatial}}(\boldsymbol{x}')$$

with

$$x'_{k,\ell} = \frac{x_{k,\ell}}{\langle y_{j,\ell} \rangle_j}$$

where j = spatial index, k = spectral index and $\langle y_{j,\ell} \rangle_j$ is the mean spectrum from the data.

Spectral renormalization

Large dynamical range between spectral channels lead to:

- Over-regularization of bright features,
- Under-regularization of fainter ones.

Renormalization: Regularization on a spectrally flattened object

$$f_{\text{prior}}(\boldsymbol{x}) = \Omega_{\text{spectral}}(\boldsymbol{x'}) + \Omega_{\text{spatial}}(\boldsymbol{x'})$$

with

$$x'_{k,\ell} = \frac{x_{k,\ell}}{\langle y_{j,\ell} \rangle_j}$$

where j = spatial index, k = spectral index and $\langle y_{j,\ell} \rangle_j$ is the mean spectrum from the data.

Simulated MUSE data

Multi-Units Spectroscopic Explorer (MUSE)

Integral field spectrograph for the Very Large Telescope (VLT)

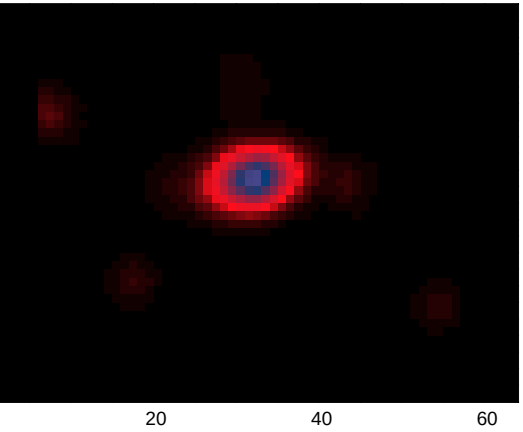
- 3578 spectral chanels from 465 nm to 930 nm,
- 300×300 pixels,
- pixel size $0.2'' \times 0.2''$,
- PSF estimation using adaptive optic data and guiding stars.
- provides variance map.

Test data

- Data, PSFs and variance map provided by the MUSE simulator,
- sky substracted,
- 51×36 pixels, 3578 spectral chanels,
- PSF FWHM varies spectrally from 3.75 $(0.93'')$ to 4.66 $(0.75'')$ pixels.

MUSE simulations (Soulez et al. 2013)

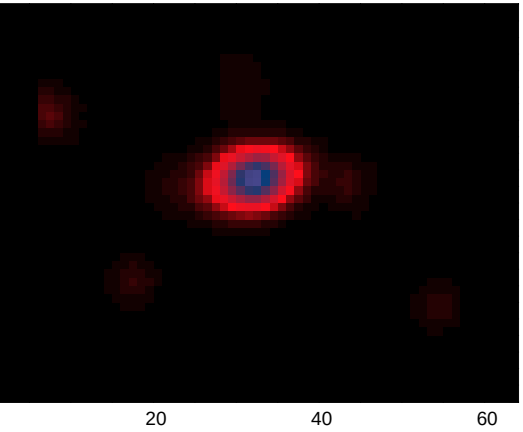
Data



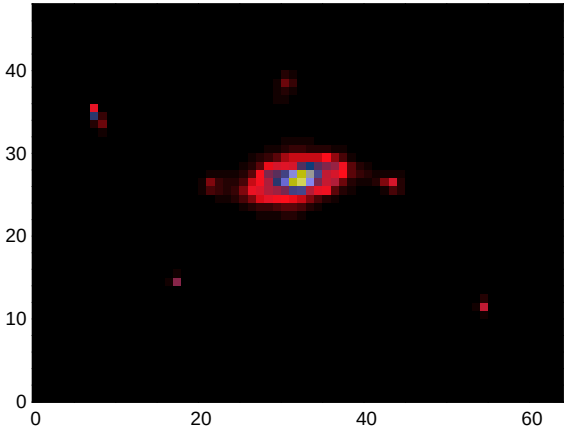
Images integrated over the spectral range.

MUSE simulations (Soulez et al. 2013)

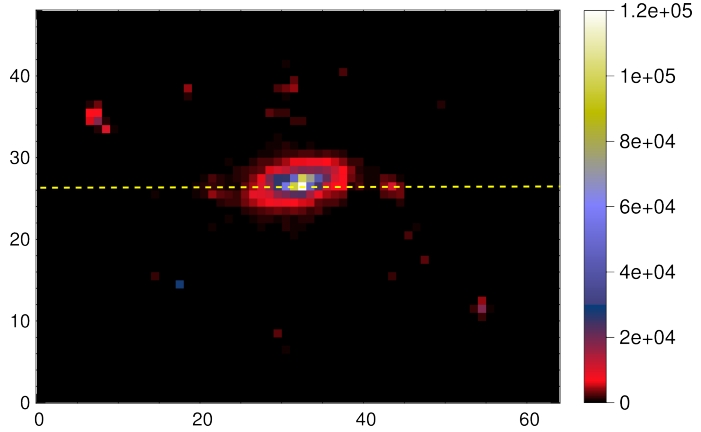
Data



Deconvolution



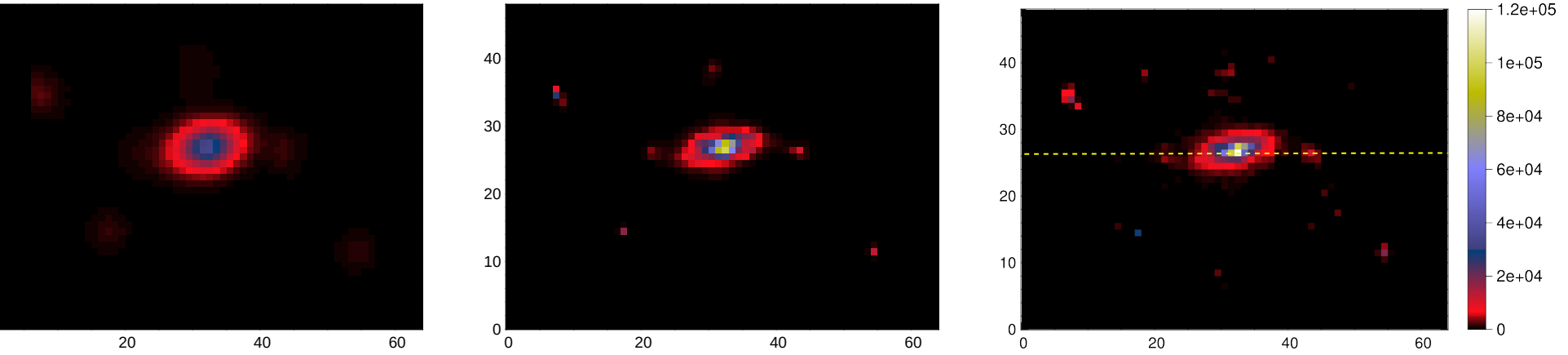
Ground truth



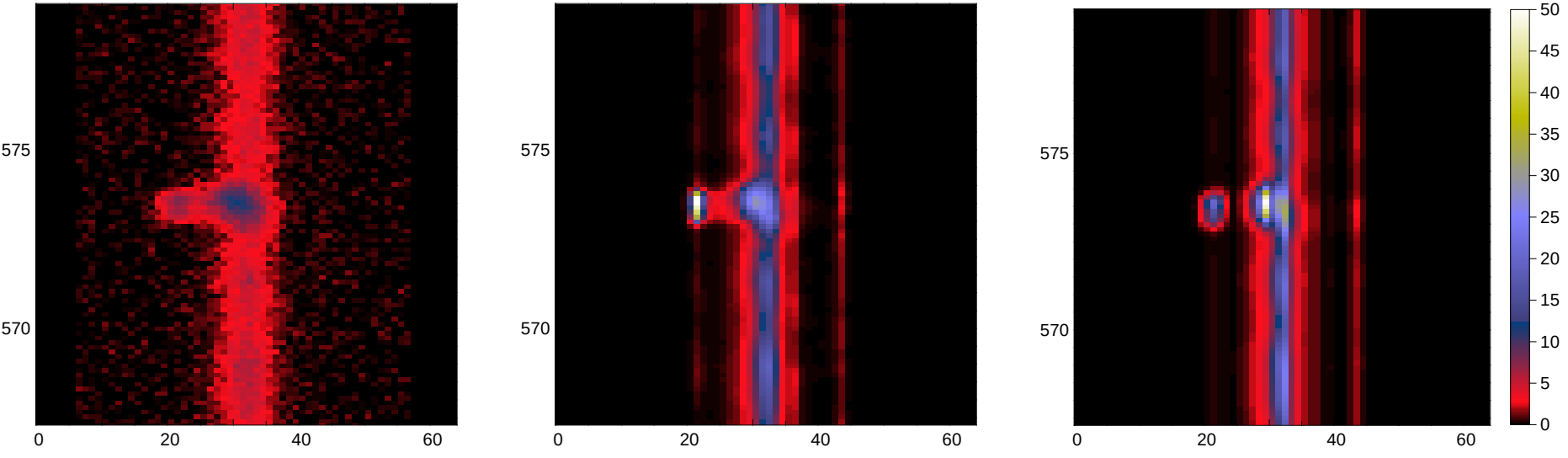
Images integrated over the spectral range.

MUSE simulations (Soulez et al. 2013)

Data | Deconvolution | Ground truth

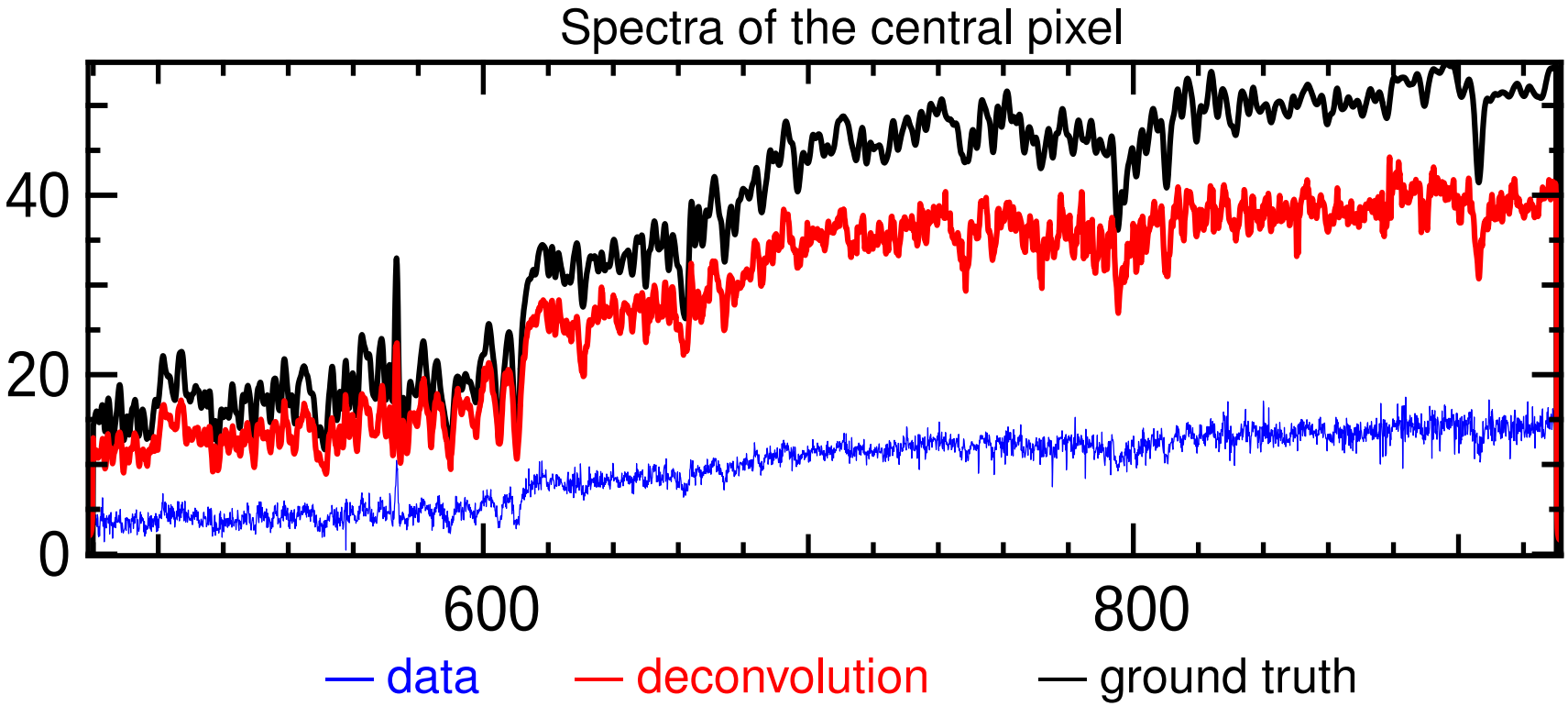


Images integrated over the spectral range.

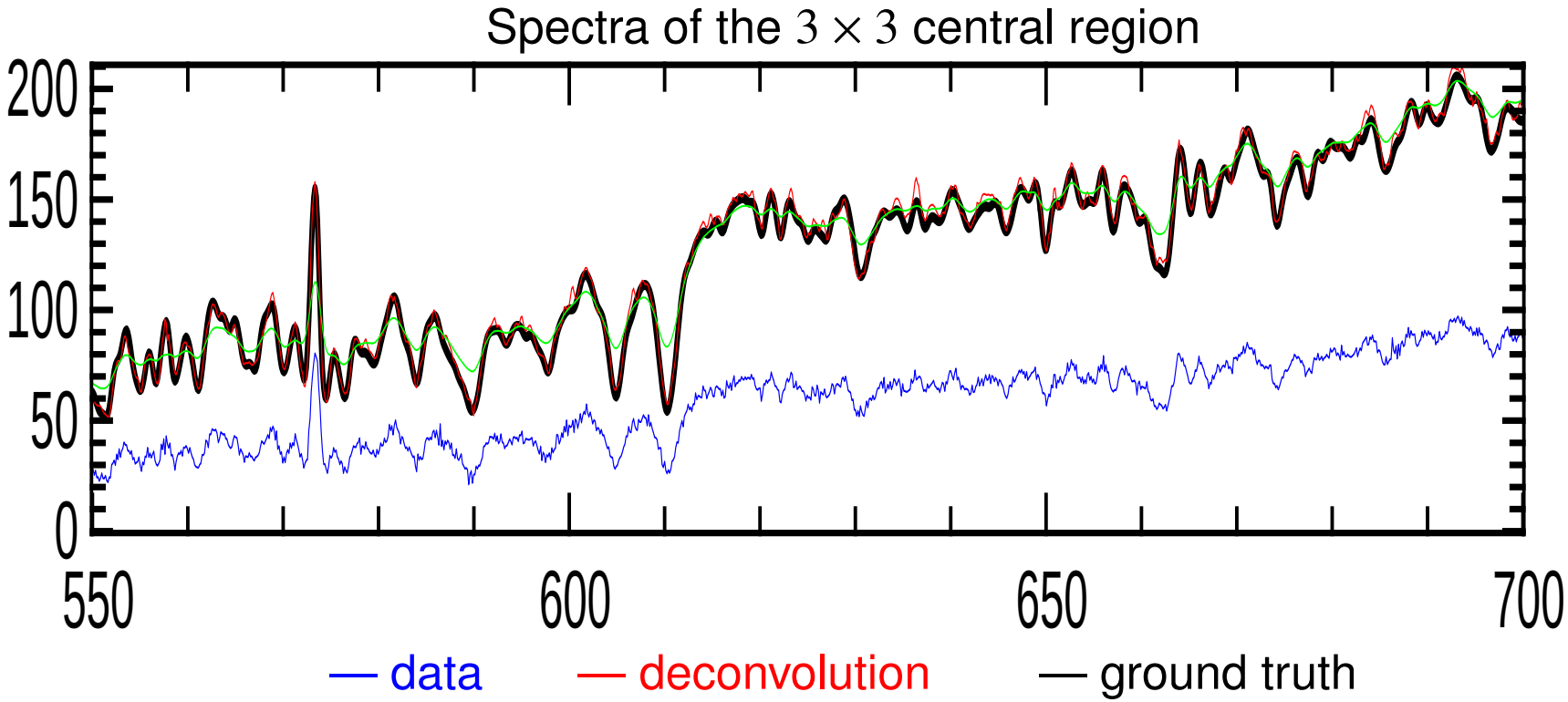


Detail of (x, λ) cuts along the yellow line.

MUSE simulations

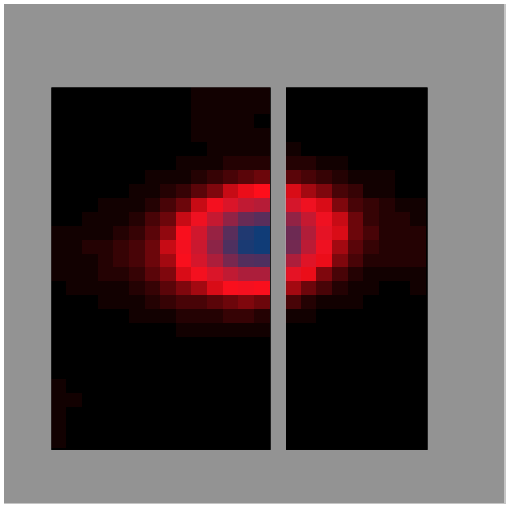


MUSE simulations

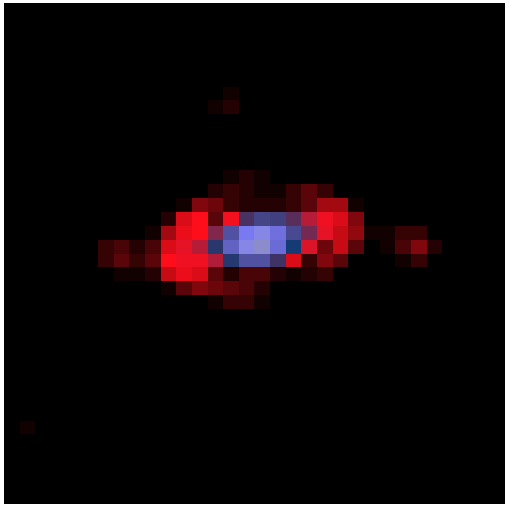


Extrapolation and missing data

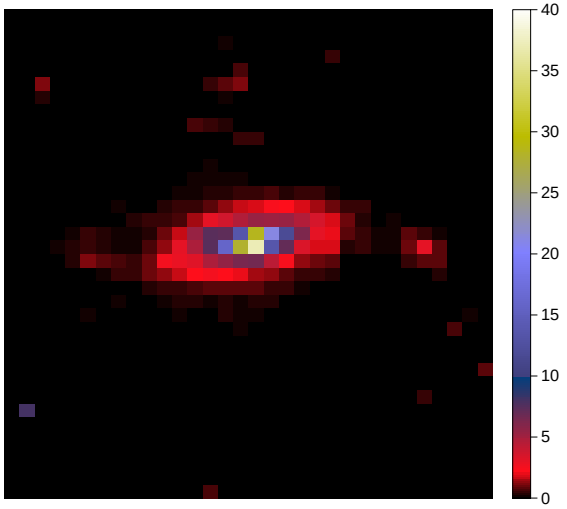
Data



Deconvolution



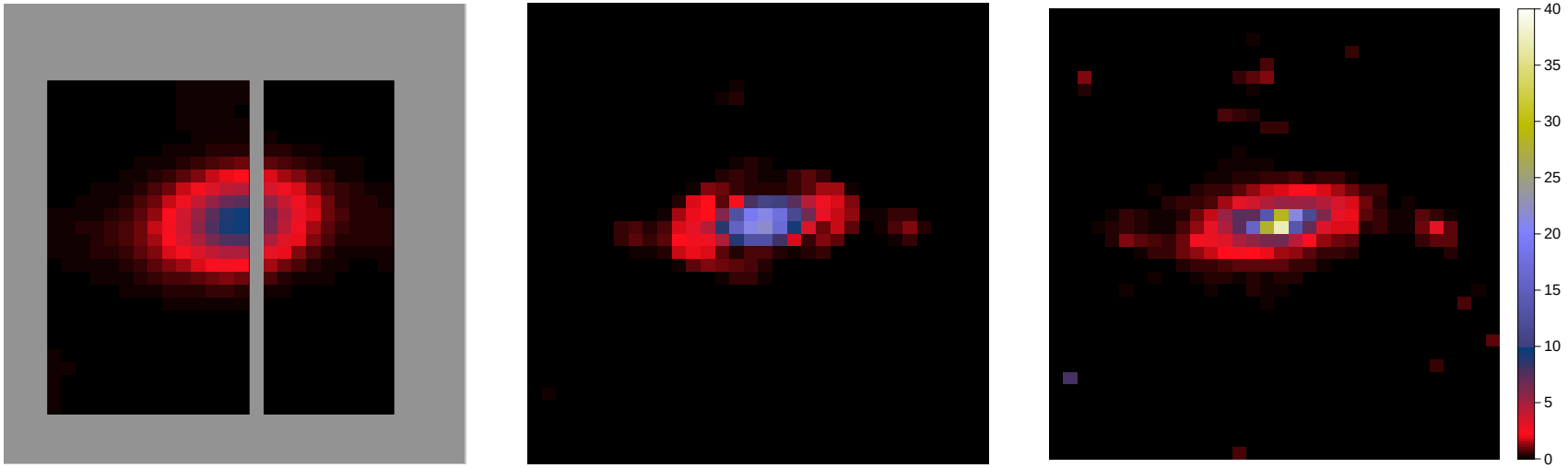
Ground truth



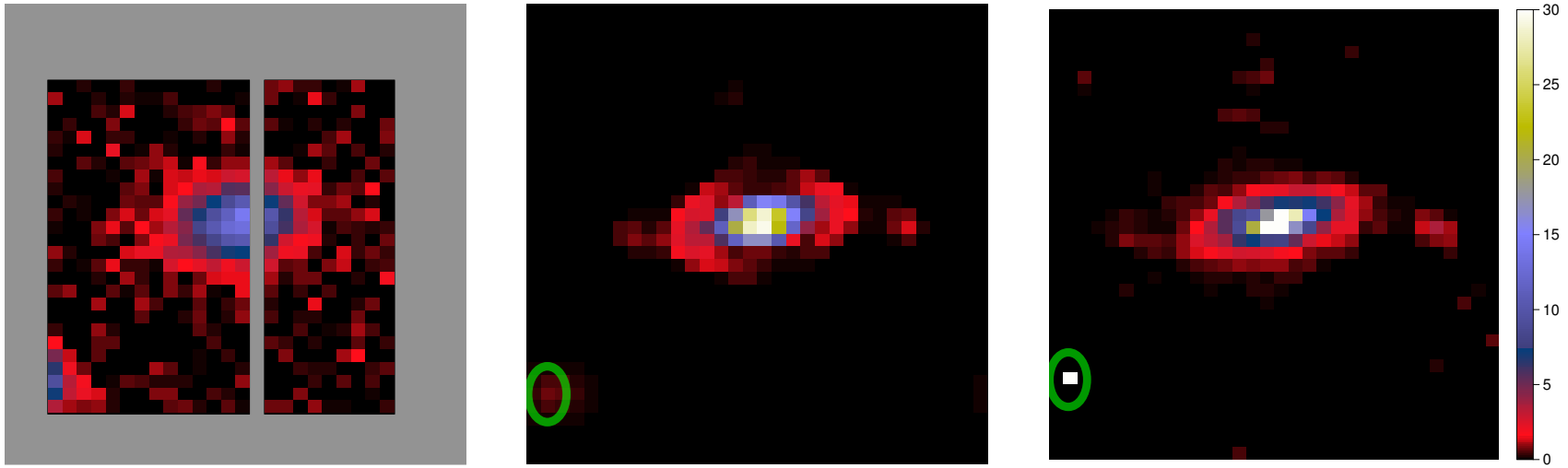
Images integrated over the spectral range.

Extrapolation and missing data

Data | Deconvolution | Ground truth



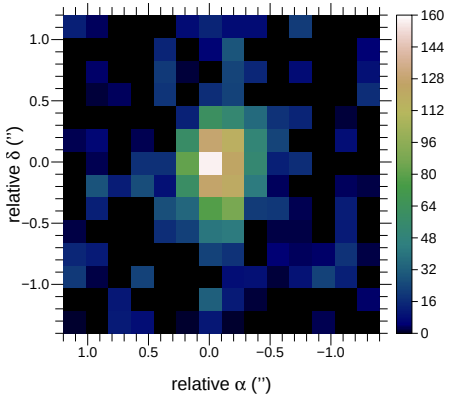
Images integrated over the spectral range.



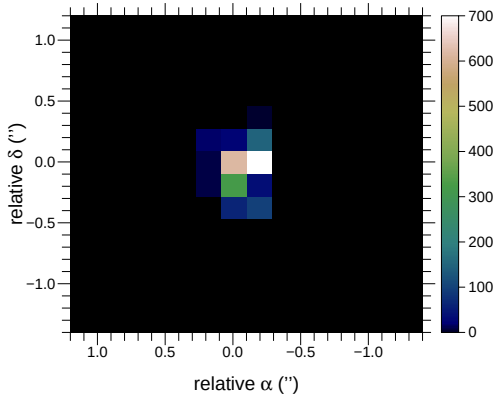
Images at $\lambda = 878.2$nm.

Preliminary result on KMOS data

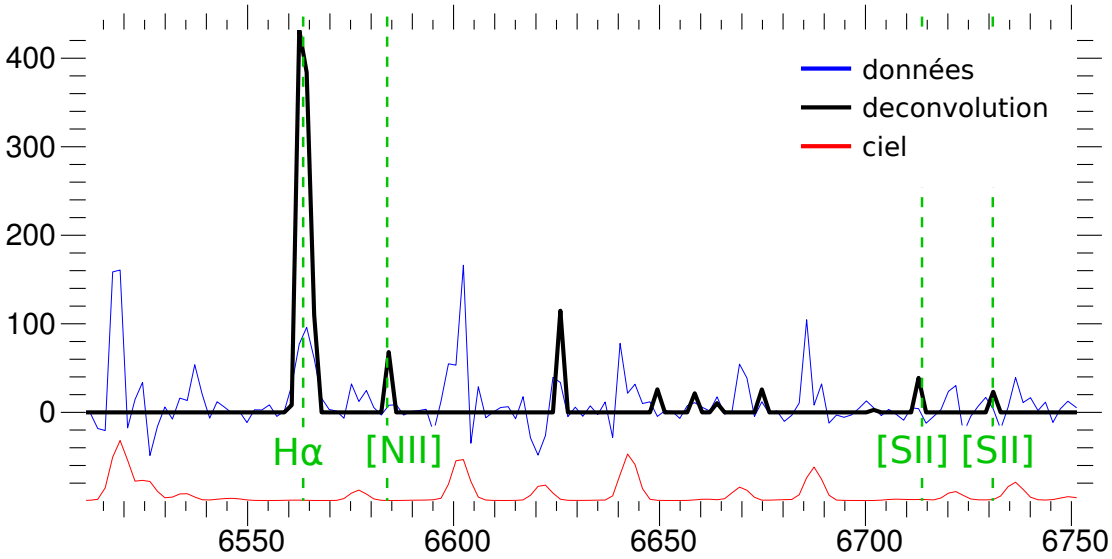
Distant galaxy ($z = 1.56$) lensed by the galaxy cluster Abell 1689.



Data (Hα)



Deconvolution

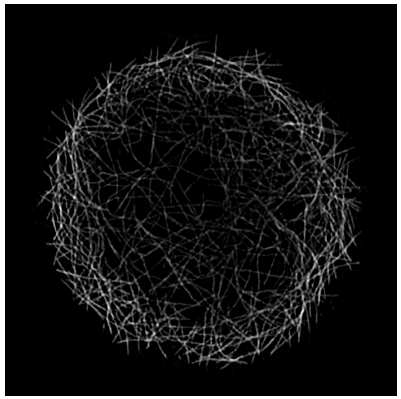


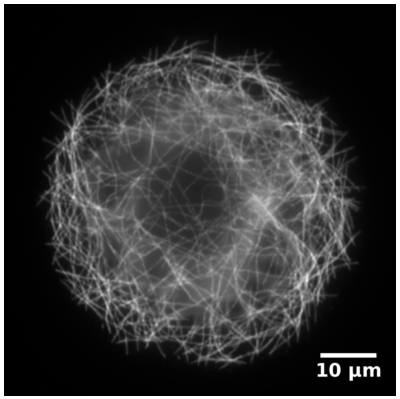
Perspectives

- direct model using intensity distribution on the detector plane (minimal data reducing step),
- blind deconvolution Soulez et al. (2012)
- space varying deconvolution Denis et al. (2011)

Perspectives

- direct model using intensity distribution on the detector plane (minimal data reducing step),
- blind deconvolution Soulez et al. (2012)

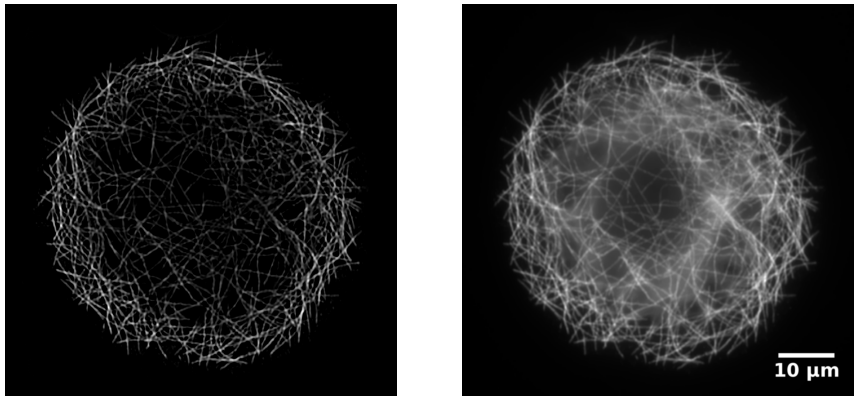




- space varying deconvolution Denis et al. (2011)

Perspectives

- direct model using intensity distribution on the detector plane (minimal data reducing step),
- blind deconvolution Soulez et al. (2012)



- space varying deconvolution Denis et al. (2011)

Demixing in Integral Field Spectrography

Supernovæ factory

SNfactor: a Franco American collaboration (IPNL, LPNHE, CRAL, LBNL)Aldering et al. (2002)

Goals:

- Using type Ia Supernovæ as standard candles,
- build the Hubble diagram for $0.03 < z < 0.08$.

constraints:

- time-lapse observation on several weeks,
- precise spectra estimations,
- precise photometry ($\leq 1\%$).

Instrument:

- Integral field spectrograph (SNIFS) at the hawai university telescope,
- $6'' \times 6''$ (15×15 pixels),
- 2×798 spectral channels,
- PSF estimated using photometric channel.

Data formation (1)

Ideal case

$$I(\boldsymbol{\theta}, \lambda, t) = I_{sky}(\lambda, t) + I_{gal}(\boldsymbol{\theta}, \lambda) + F_{SN}(\lambda, t) * \delta(\boldsymbol{\theta_s}) \,,$$

Sky emission background $I_{ciel}(\lambda, t)$

- spatially flat,
- vary spectrally and temporally.

Host galaxay $I_{gal}(\theta, \lambda)$

- constant in time,
- Spatially and spectrally structured.

Supernova $F_{SN}(\lambda, t)$

- point source centered in θ_s,
- vary spectrally and temporally.

Data formation (1)

Ideal case

$$I(\boldsymbol{\theta}, \lambda, t) = I_{sky}(\lambda, t) + I_{gal}(\boldsymbol{\theta}, \lambda) + F_{SN}(\lambda, t) * \delta(\boldsymbol{\theta_s}) \,,$$

Sky emission background $I_{ciel}(\lambda, t)$

- spatially flat,
- vary spectrally and temporally.

Host galaxay $I_{gal}(\theta, \lambda)$

- constant in time,
- Spatially and spectrally structured.

Supernova $F_{SN}(\lambda, t)$

- point source centered in θ_s,
- vary spectrally and temporally.

Data formation (1)

Ideal case

$$I(\boldsymbol{\theta}, \lambda, t) = I_{sky}(\lambda, t) + I_{gal}(\boldsymbol{\theta}, \lambda) + F_{SN}(\lambda, t) * \delta(\boldsymbol{\theta_s}) \,,$$

Sky emission background $I_{ciel}(\lambda, t)$

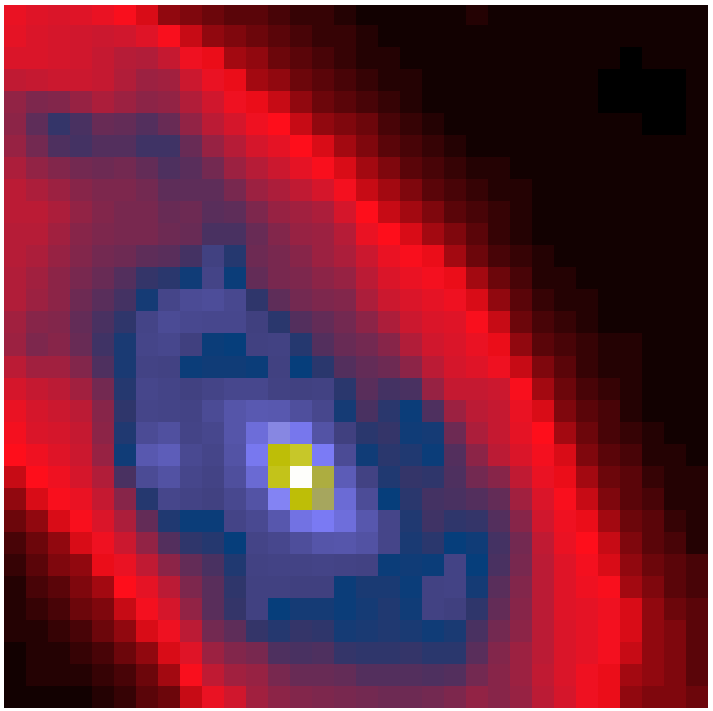
- spatially flat,
- vary spectrally and temporally.

Host galaxay $I_{gal}(\boldsymbol{\theta}, \lambda)$

- constant in time,
- Spatially and spectrally structured.

Supernova $F_{SN}(\lambda, t)$

- point source centered in θ_s,
- vary spectrally and temporally.



Data formation (1)

Ideal case

$$I(\boldsymbol{\theta}, \lambda, t) = I_{sky}(\lambda, t) + I_{gal}(\boldsymbol{\theta}, \lambda) + F_{SN}(\lambda, t) * \delta(\boldsymbol{\theta}_s)\,,$$

Sky emission background $I_{ciel}(\lambda, t)$

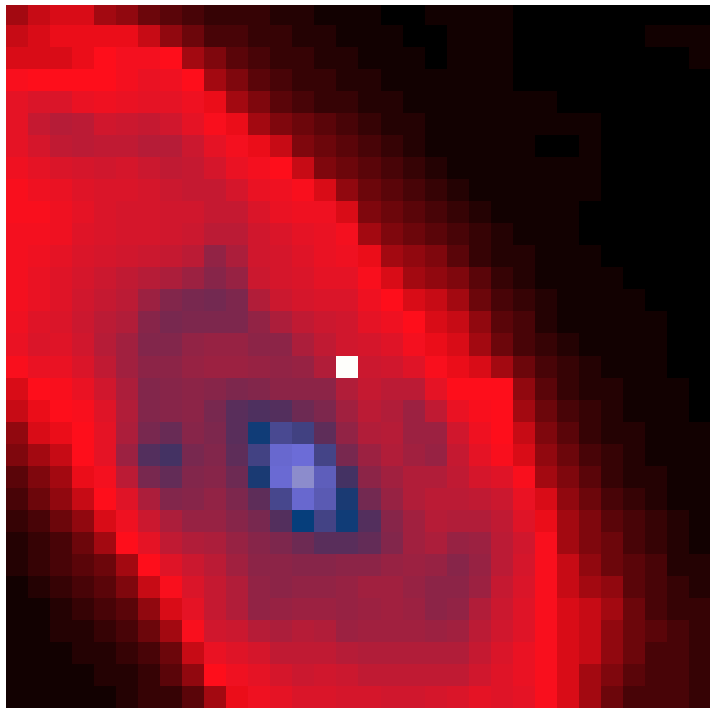
- spatially flat,
- vary spectrally and temporally.

Host galaxay $I_{gal}(\boldsymbol{\theta}, \lambda)$

- constant in time,
- Spatially and spectrally structured.

Supernova $F_{SN}(\lambda, t)$

- point source centered in $\boldsymbol{\theta}_s$,
- vary spectrally and temporally.



Data formation (2)

In practice:

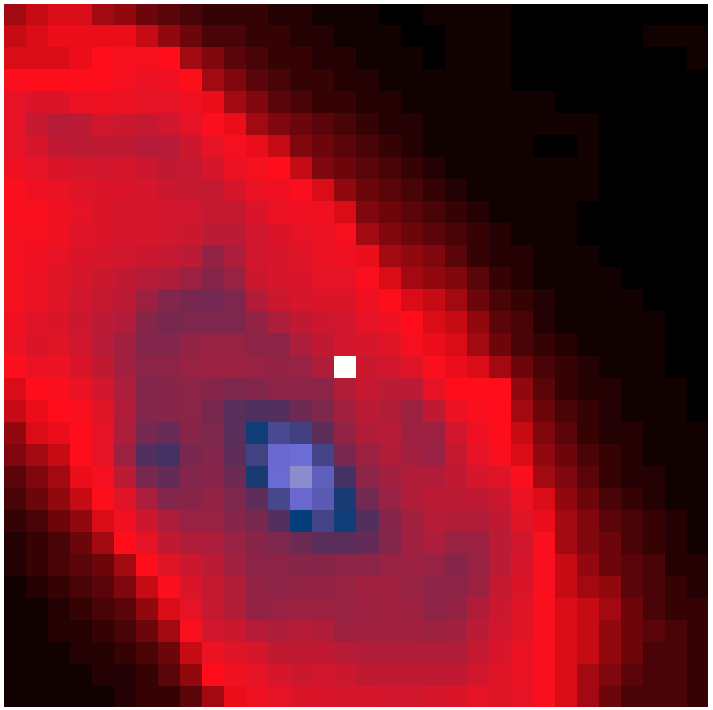
$$
\begin{aligned}
I(\boldsymbol{\theta}, \lambda, t) &= I_{ciel}(\lambda, t) + I_{gal}(\boldsymbol{\theta}, \lambda) + F_{SN}(\lambda, t) * \delta(\boldsymbol{\theta_s}) \,, \\
\boldsymbol{data} &= \boldsymbol{I} * \boldsymbol{psf} * \delta_{\mathrm{ADR}} * \delta_{\mathrm{pointage}} + \boldsymbol{b} \,,
\end{aligned}
$$

Blur (seeing limited) $\boldsymbol{psf}$

Astmospheric differential refraction δ_{ADR} and pointing errors δ_{point}

Field truncation

Noises



Data formation (2)

In practice:

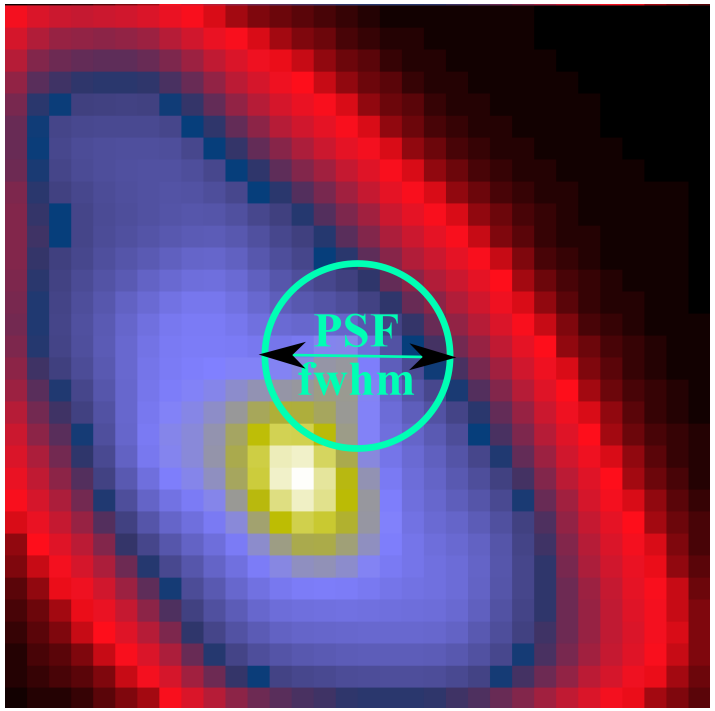
$$
\begin{aligned}
I(\boldsymbol{\theta}, \lambda, t) &= I_{ciel}(\lambda, t) + I_{gal}(\boldsymbol{\theta}, \lambda) + F_{SN}(\lambda, t) * \delta(\boldsymbol{\theta_s}) \,, \\
\boldsymbol{data} &= \boldsymbol{I} * \boldsymbol{psf} * \delta_{\mathrm{ADR}} * \delta_{\mathrm{pointage}} + \boldsymbol{b} \,,
\end{aligned}
$$

Blur (seeing limited) ***psf***

Astmospheric differential refraction δ_{ADR} and pointing errors δ_{point}

Field truncation

Noises



Data formation (2)

In practice:

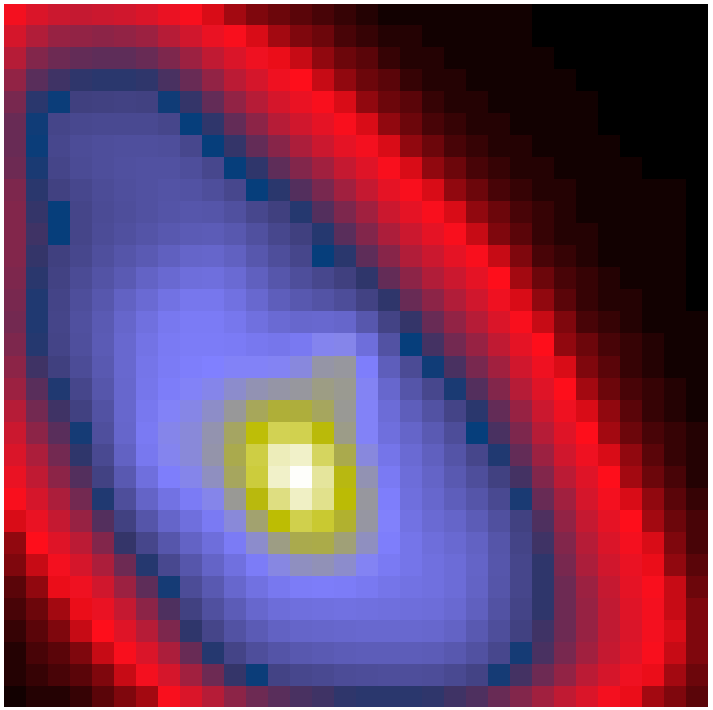
$$
\begin{aligned}
I(\boldsymbol{\theta}, \lambda, t) &= I_{ciel}(\lambda, t) + I_{gal}(\boldsymbol{\theta}, \lambda) + F_{SN}(\lambda, t) * \delta(\boldsymbol{\theta_s}) \,, \\
\boldsymbol{data} &= \boldsymbol{I} * \boldsymbol{psf} * \delta_{\mathrm{ADR}} * \delta_{\mathrm{pointage}} + \boldsymbol{b} \,,
\end{aligned}
$$

Blur (seeing limited) ***psf***

Astmospheric differential refraction δ_{ADR} and pointing errors δ_{point}

Field truncation

Noises



Data formation (2)

In practice:

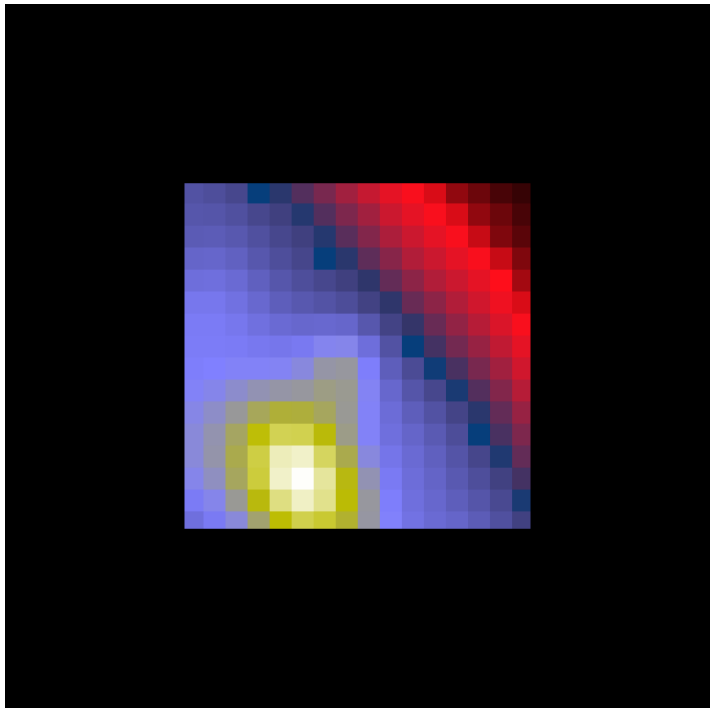
$$
\begin{aligned}
I(\boldsymbol{\theta}, \lambda, t) &= I_{ciel}(\lambda, t) + I_{gal}(\boldsymbol{\theta}, \lambda) + F_{SN}(\lambda, t) * \delta(\boldsymbol{\theta_s}) \,, \\
\boldsymbol{data} &= \boldsymbol{I} * \boldsymbol{psf} * \delta_{\text{ADR}} * \delta_{\text{pointage}} + \boldsymbol{b} \,,
\end{aligned}
$$

Blur (seeing limited) ***psf***

Astmospheric differential refraction δ_{ADR} and pointing errors δ_{point}

Field truncation

Noises



Data formation (2)

In practice:

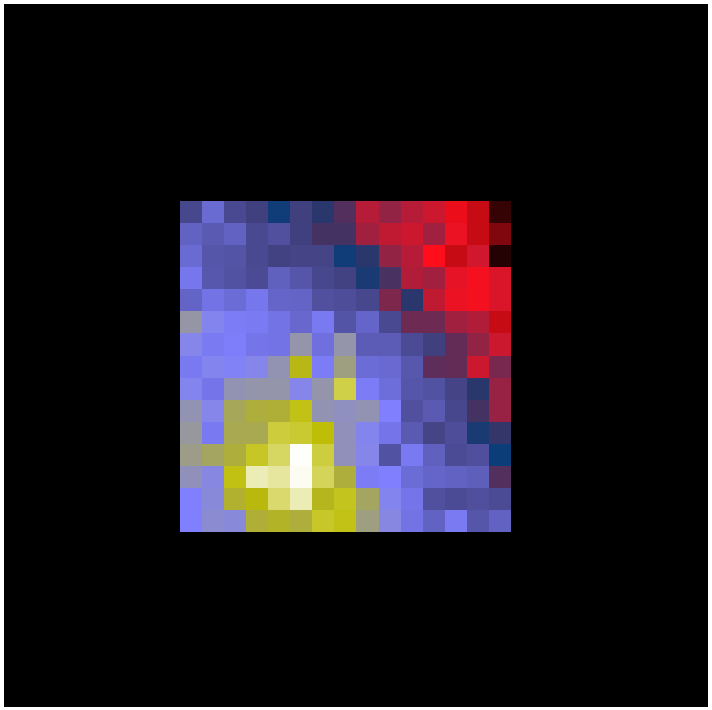
$$
\begin{aligned}
I(\boldsymbol{\theta}, \lambda, t) &= I_{ciel}(\lambda, t) + I_{gal}(\boldsymbol{\theta}, \lambda) + F_{SN}(\lambda, t) * \delta(\boldsymbol{\theta_s}) \,, \\
\boldsymbol{data} &= \boldsymbol{I} * \boldsymbol{psf} * \delta_{\mathrm{ADR}} * \delta_{\mathrm{pointage}} + \boldsymbol{b} \,,
\end{aligned}
$$

Blur (seeing limited) ***psf***

Astmospheric differential refraction δ_{ADR} and pointing errors δ_{point}

Field truncation

Noises



Demixing:

Direct model:

$$
\begin{aligned}
I(\boldsymbol{a}, \lambda, t) &= I_{ciel}(\lambda, t) + I_{gal}(\boldsymbol{a}, \lambda) + F_{SN}(\lambda, t) * \delta(\boldsymbol{a_c}) \,, \\
\boldsymbol{model} &= \boldsymbol{I} * \boldsymbol{psf} * \delta_{\mathrm{ADR}} * \delta_{\mathrm{point}} \,,
\end{aligned}
$$

MAP solution

- Solution $\{\boldsymbol{F}_{SN}, \boldsymbol{I}_{gal}\}^{+}$ that minimizes

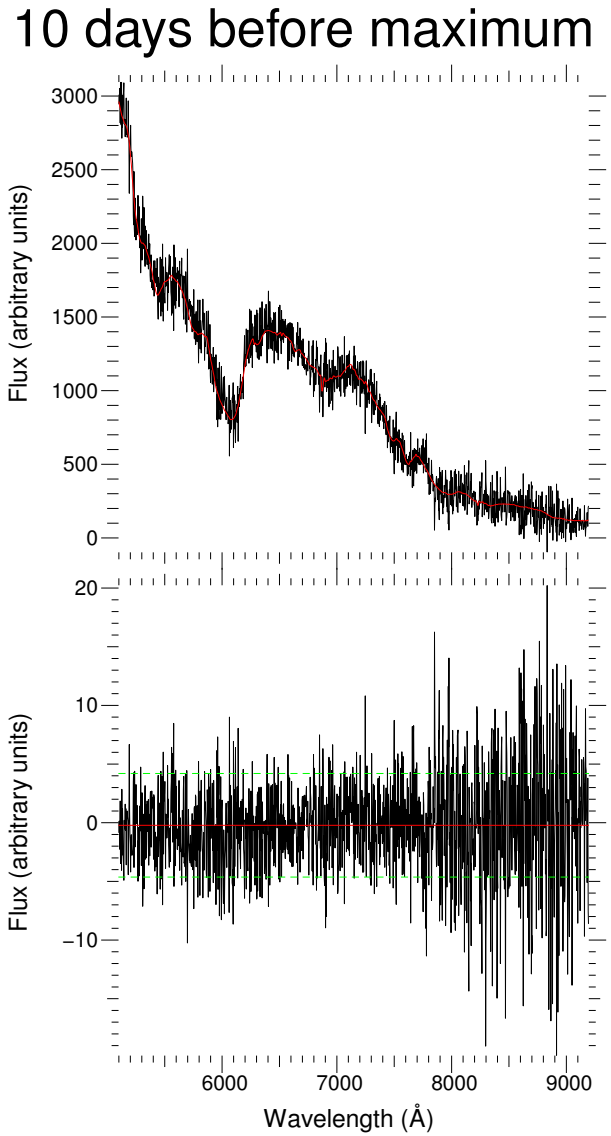
$$
f = ||\boldsymbol{model} - \boldsymbol{data}||^2 + f_{\mathrm{prior}}(\boldsymbol{I}_{gal}) \,,
$$

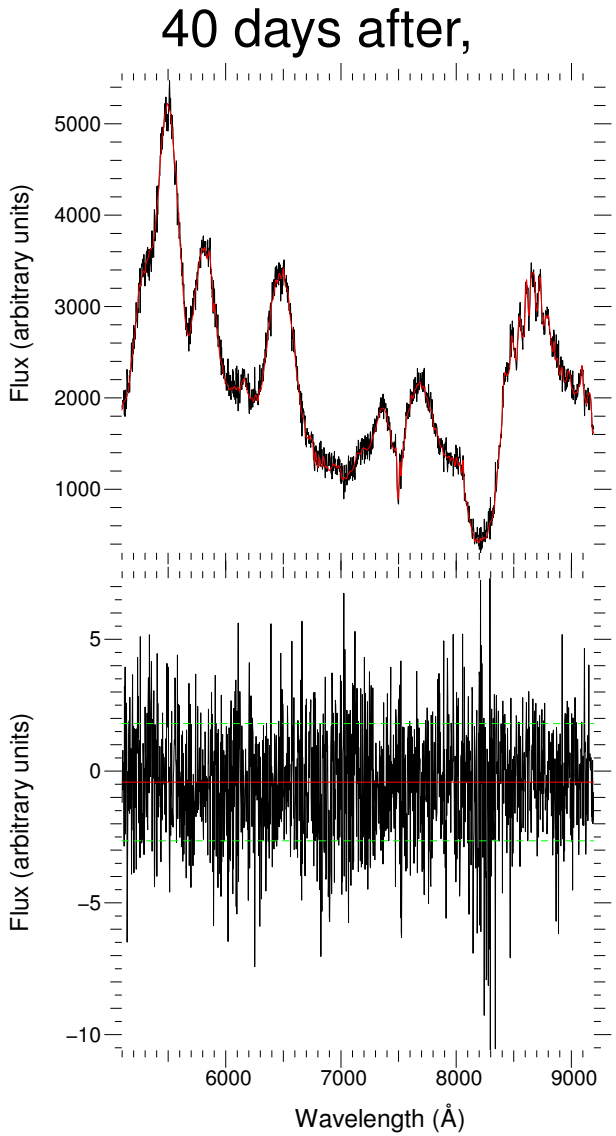
- $f_{\mathrm{prior}}(\boldsymbol{I}_{gal})$: regularization (quadratic smoothness) on the galaxy,
- no priors on the supernova.

Results: simulation

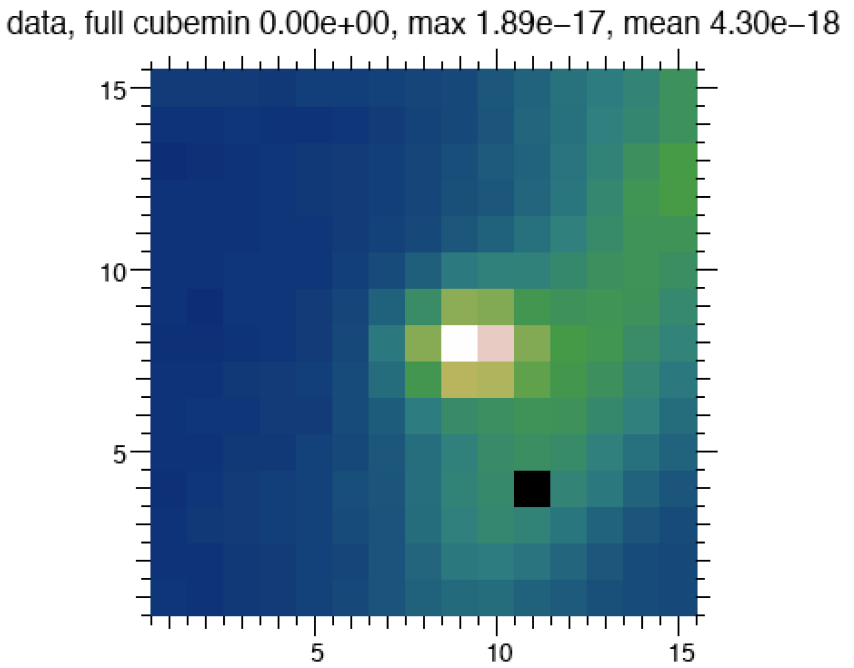
supernova spectra
(truth, reconstruction).

errors

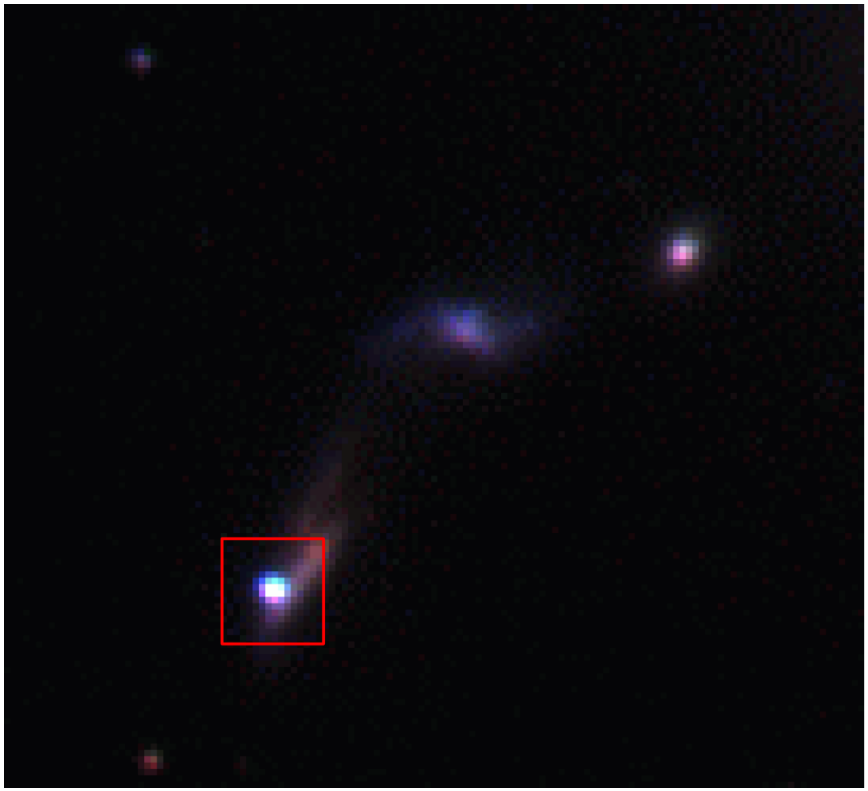




Real data: sn2004gc

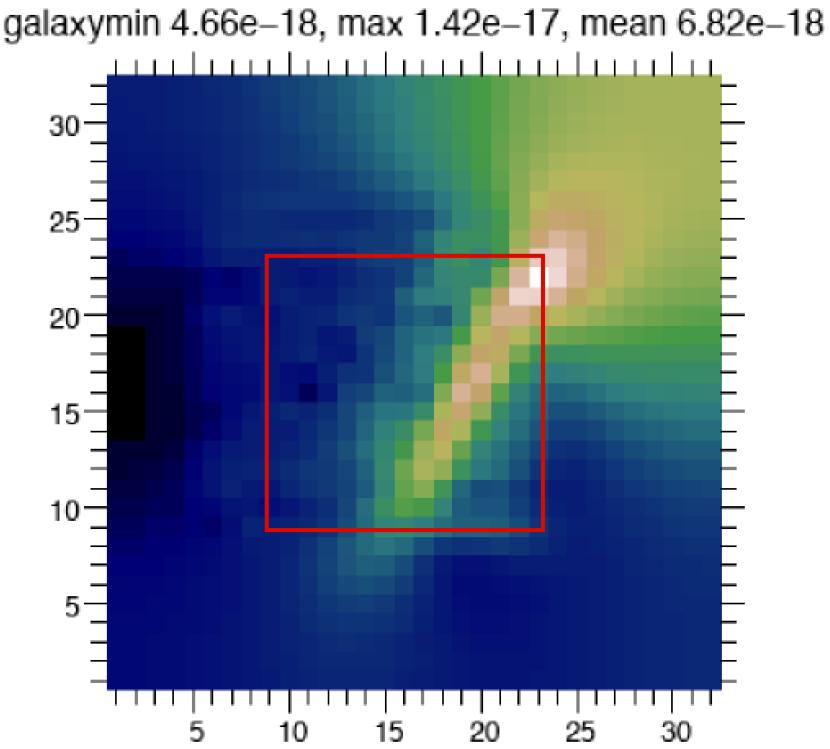


Data (integrated).

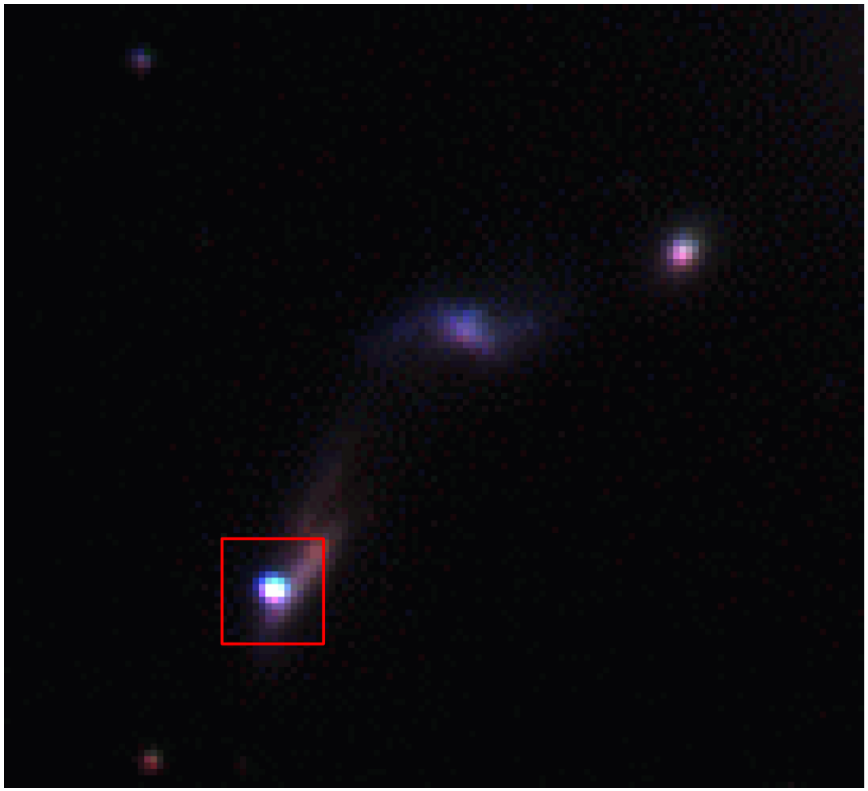


sn2004gc (SDSS).

Real data: sn2004gc

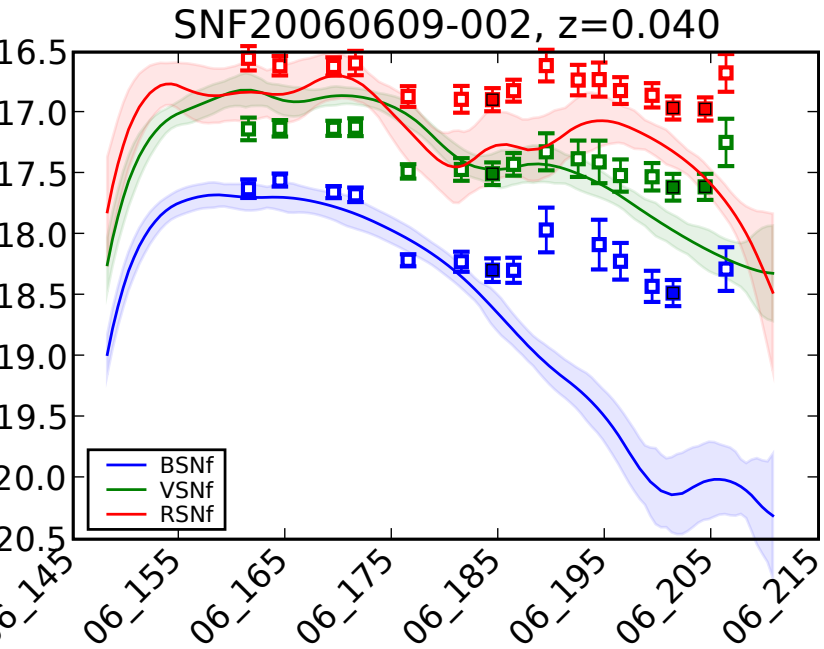


Reconstruction.

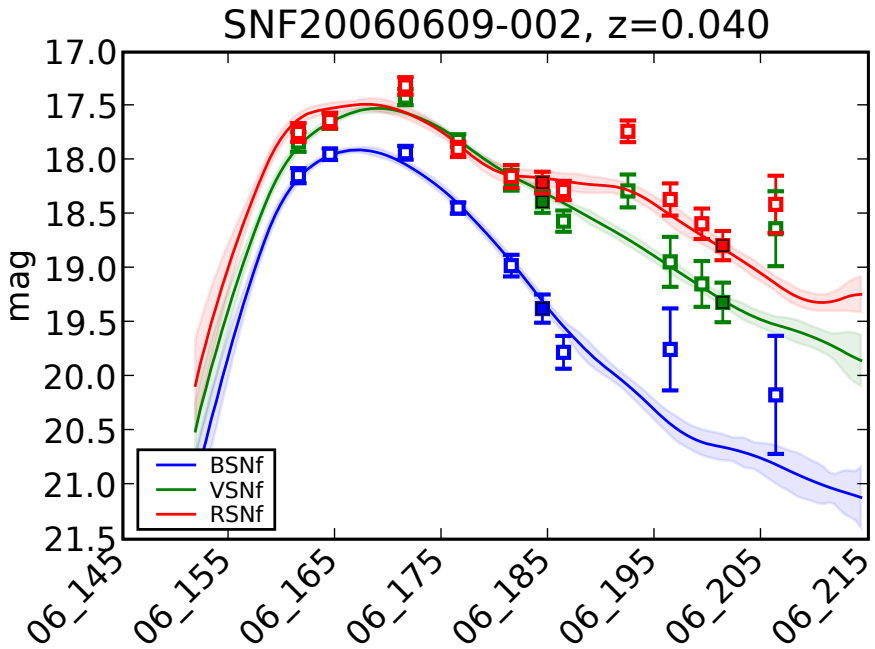


sn2004gc (SDSS).

Real data: light curve of sn2006?



conventional



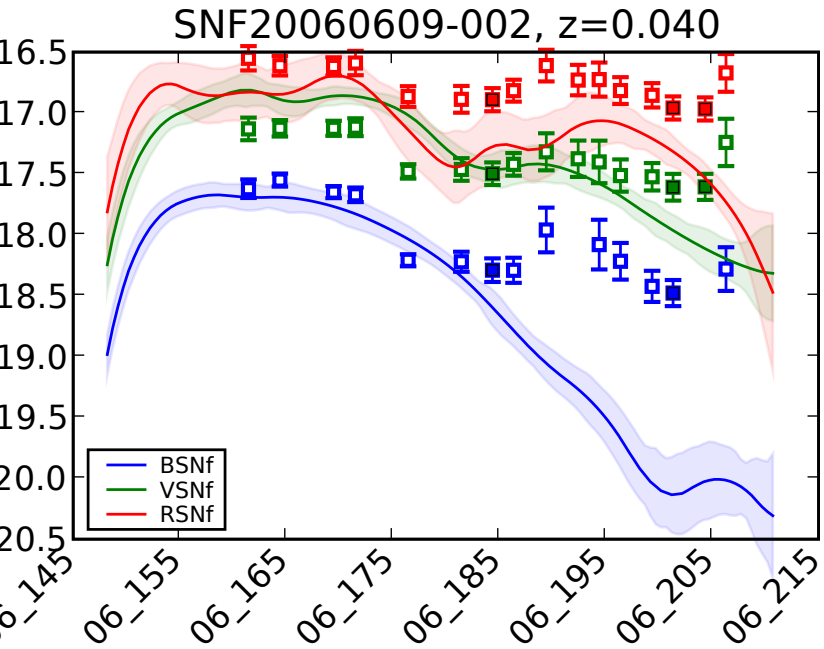
inverse problem approach

SN estimated photometry and light curves fitted by the model SALT2

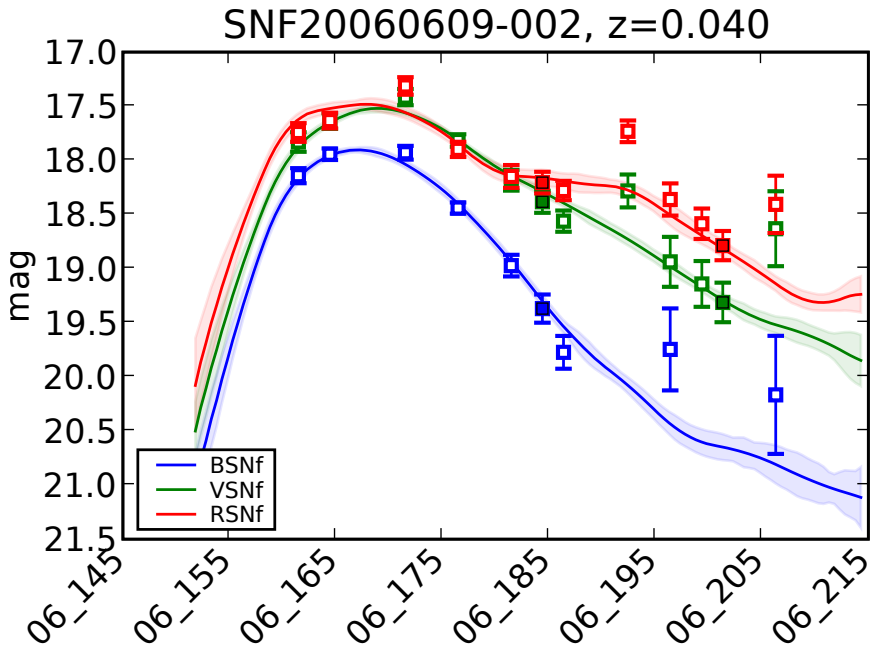
Twice more supernova in the sample after this better decontamination

More supernovæ from central bulges included in the sample.

Real data: light curve of sn2006?



conventional



inverse problem approach

SN estimated photometry and light curves fitted by the model SALT2

Twice more supernova in the sample after this better decontamination

More supernovæ from central bulges included in the sample.

Interferometry

Introduction

- **Context**,
  - there is a need of dedicated image reconstruction algorithms for polychromatic interferometry,
  - spatio-spectral correlations put strong constraints for image reconstruction (SNIFS (Bongard et al. 2011), MUSE (Soulez et al. 2013)),
  - one of the goal of the POLCA project,
- **Outlines**
  - image reconstruction framework,
  - with phase reference (GRAVITY),
  - without phase reference (PIONNIER)

Introduction

- **Context**,
  - there is a need of dedicated image reconstruction algorithms for polychromatic interferometry,
  - spatio-spectral correlations put strong constraints for image reconstruction (SNIFS (Bongard et al. 2011), MUSE (Soulez et al. 2013)),
  - one of the goal of the POLCA project,
- **Outlines**
  - image reconstruction framework,
  - with phase reference (GRAVITY),
  - without phase reference (PIONNIER)

Image reconstruction

- **goal:** *estimate 3D intensity distribution $I(\theta_n, \lambda_\ell)$ of the observed object*
- **we have** (measurements):
  - interferometric measurements (visibilities, powerspectrum, closure. . .),
- **we know** (priors):
  - positivity,
  - normalization,
  - type of the object $\longrightarrow$ regularization function.

Image reconstruction

- **goal:** *estimate 3D intensity distribution $I(\theta_n, \lambda_\ell)$ of the observed object*
- **we have** (measurements):
  - interferometric measurements (visibilities, powerspectrum, closure...),
- **we know** (priors):
  - positivity,
  - normalization,
  - type of the object $\longrightarrow$ regularization function.

Image reconstruction

- **goal:** *estimate 3D intensity distribution* $I(\theta_n, \lambda_\ell)$ *of the observed object*
- **we have** (measurements):
  - interferometric measurements (visibilities, powerspectrum, closure...),
- **we know** (priors):
  - positivity,
  - normalization,
  - type of the object $\longrightarrow$ regularization function.

Inverse problem framework

- Reconstructed image:

$$\boldsymbol{x}^{+} = \arg\min_{\boldsymbol{x}\in\mathbb{X}} \underbrace{f_{\text{prior}}(\boldsymbol{x})}_{\text{regularization}} \quad \text{s.t.} \quad \underbrace{f_{\text{data}}(\boldsymbol{x}|\boldsymbol{m})}_{\text{likelihood}} \leq \eta_1 \quad \text{and} \quad \underbrace{\sum_{\lambda} \boldsymbol{x}_{\lambda} = 1}_{\text{normalization}}$$

- with:
  - η_1 a tolerance level (hyperparameter);
  - $\mathbb{X}$ feasible set *e.g.* $\mathbb{X} = \{\boldsymbol{x} \in \mathbb{R}_{+}^{N}\}$

Interferometric likelihood (with phase reference)

- **Sampling**:

$$\boxed{x_{n,\ell} = I(\boldsymbol{\theta}_n, \lambda_\ell)} \quad \text{avec:} \quad \begin{cases} \lambda_\ell & \text{wavelength of } \ell\text{pspectral channel} \\ \boldsymbol{\theta}_n & \text{angular position of pixel } n \\ \boldsymbol{B}_b & \text{projected position of baseline } b \end{cases}$$

- **Direct model**:

$$m_{p,b,\ell} = \sum\nolimits_n H_{p,b,n,\ell}\, x_{n,\ell} + e_{p,b,\ell} \quad \text{in brief:} \quad \boxed{\boldsymbol{m} = \mathbf{H} \cdot \boldsymbol{x} + \boldsymbol{e}}$$

with:

$$H_{p,b,n,\ell} = \begin{cases} +\cos(\boldsymbol{\theta}_n^{\mathrm{T}} \cdot \boldsymbol{B}_b / \lambda_\ell) & \text{for } p = 1 \text{ (real part)} \\ -\sin(\boldsymbol{\theta}_n^{\mathrm{T}} \cdot \boldsymbol{B}_b / \lambda_\ell) & \text{for } p = 2 \text{ (imaginary part)} \end{cases}$$

- **Interferometric likelihood term** (assuming Gaussian statistics for the errors)

$$\boxed{f_{\text{data}}(\boldsymbol{x}|\boldsymbol{m}) = \frac{1}{2} (\mathbf{H} \cdot \boldsymbol{x} - \boldsymbol{m})^{\mathrm{T}} \cdot \mathbf{W} \cdot (\mathbf{H} \cdot \boldsymbol{x} - \boldsymbol{m})} \quad \text{with:} \quad \mathbf{W} = \mathrm{Cov}(\boldsymbol{e})^{-1}$$

Interferometric likelihood (with phase reference)

- **Sampling**:

$$\boxed{x_{n,\ell} = I(\boldsymbol{\theta}_n, \lambda_\ell)} \quad \text{avec:} \quad \begin{cases} \lambda_\ell & \text{wavelength of } \ell\text{pspectral channel} \\ \boldsymbol{\theta}_n & \text{angular position of pixel } n \\ \boldsymbol{B}_b & \text{projected position of baseline } b \end{cases}$$

- **Direct model**:

$$m_{p,b,\ell} = \sum\nolimits_n H_{p,b,n,\ell}\, x_{n,\ell} + e_{p,b,\ell} \quad \text{in brief:} \quad \boxed{\boldsymbol{m} = \mathbf{H} \cdot \boldsymbol{x} + \boldsymbol{e}}$$

with:

$$H_{p,b,n,\ell} = \begin{cases} +\cos(\boldsymbol{\theta}_n^{\mathrm{T}} \cdot \boldsymbol{B}_b / \lambda_\ell) & \text{for } p = 1 \text{ (real part)} \\ -\sin(\boldsymbol{\theta}_n^{\mathrm{T}} \cdot \boldsymbol{B}_b / \lambda_\ell) & \text{for } p = 2 \text{ (imaginary part)} \end{cases}$$

- **Interferometric likelihood term** (assuming Gaussian statistics for the errors)

$$\boxed{f_{\text{data}}(\boldsymbol{x}|\boldsymbol{m}) = \frac{1}{2} (\mathbf{H} \cdot \boldsymbol{x} - \boldsymbol{m})^{\mathrm{T}} \cdot \mathbf{W} \cdot (\mathbf{H} \cdot \boldsymbol{x} - \boldsymbol{m})} \quad \text{with:} \quad \mathbf{W} = \mathrm{Cov}(\boldsymbol{e})^{-1}$$

Interferometric likelihood (with phase reference)

- **Sampling**:

$$\boxed{x_{n,\ell} = I(\boldsymbol{\theta}_n, \lambda_\ell)} \quad \text{avec:} \quad \begin{cases} \lambda_\ell & \text{wavelength of } \ell\text{pspectral channel} \\ \boldsymbol{\theta}_n & \text{angular position of pixel } n \\ \boldsymbol{B}_b & \text{projected position of baseline } b \end{cases}$$

- **Direct model**:

$$m_{p,b,\ell} = \sum\nolimits_n H_{p,b,n,\ell}\, x_{n,\ell} + e_{p,b,\ell} \quad \text{in brief:} \quad \boxed{\boldsymbol{m} = \mathbf{H} \cdot \boldsymbol{x} + \boldsymbol{e}}$$

with:

$$H_{p,b,n,\ell} = \begin{cases} +\cos(\boldsymbol{\theta}_n^{\mathrm{T}} \cdot \boldsymbol{B}_b / \lambda_\ell) & \text{for } p = 1 \text{ (real part)} \\ -\sin(\boldsymbol{\theta}_n^{\mathrm{T}} \cdot \boldsymbol{B}_b / \lambda_\ell) & \text{for } p = 2 \text{ (imaginary part)} \end{cases}$$

- **Interferometric likelihood term** (assuming Gaussian statistics for the errors)

$$\boxed{f_{\mathrm{data}}(\boldsymbol{x}|\boldsymbol{m}) = \frac{1}{2} (\mathbf{H} \cdot \boldsymbol{x} - \boldsymbol{m})^{\mathrm{T}} \cdot \mathbf{W} \cdot (\mathbf{H} \cdot \boldsymbol{x} - \boldsymbol{m})} \quad \text{with:} \quad \mathbf{W} = \mathrm{Cov}(\boldsymbol{e})^{-1}$$

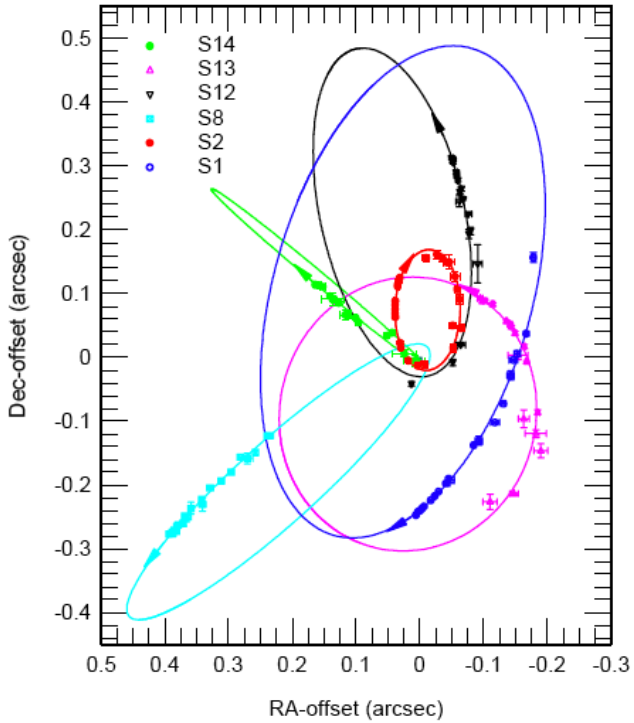
GRAVITY

GRAVITY

2nd generation VLTI instrument

- 4 telescope,
- up to $R = 4000$ in K band,
- phase referenced imaging

Main goal: probing the vicinity of Sgr A*.

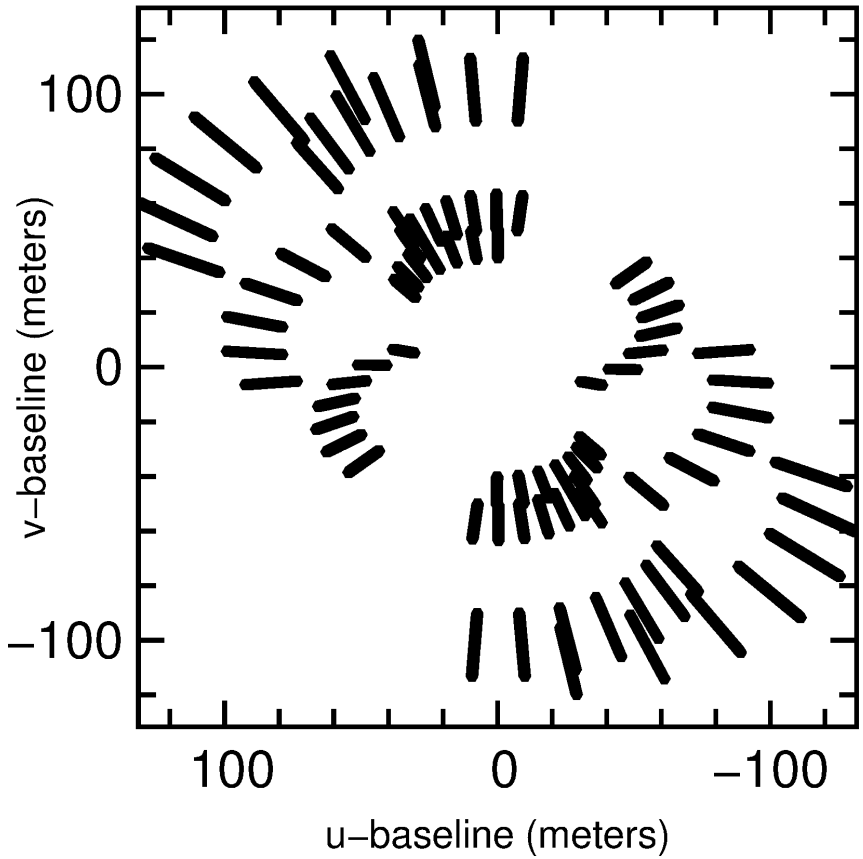


Eisenhauer *et al.* 2005

Simulation

Data simulated by GRAVITY consortium

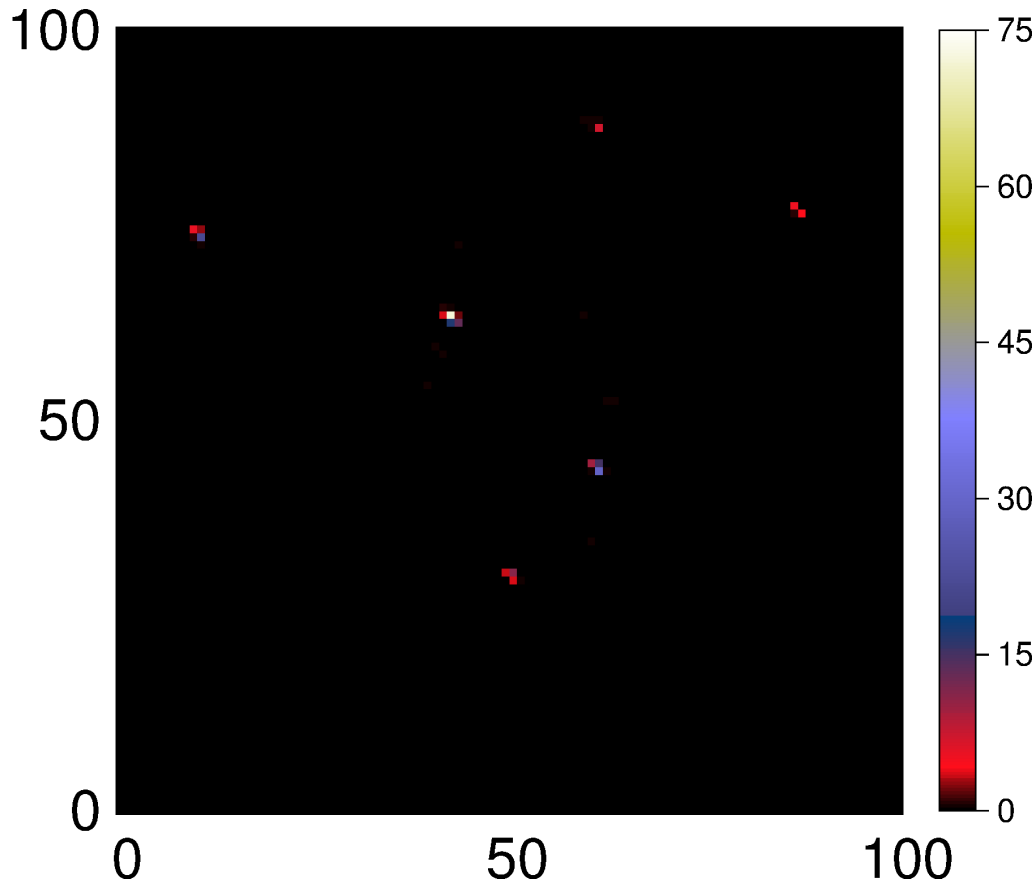
- 6 black bodies from 2000 ° to 15000 °,
- 240 spectral channels from 1.95 µm to 2.45 µm,
- good (u, v) coverage (42 baselines),
- 10080 measurements (complex visibility),
- very good SNR.



(u, v) coverage.

Results

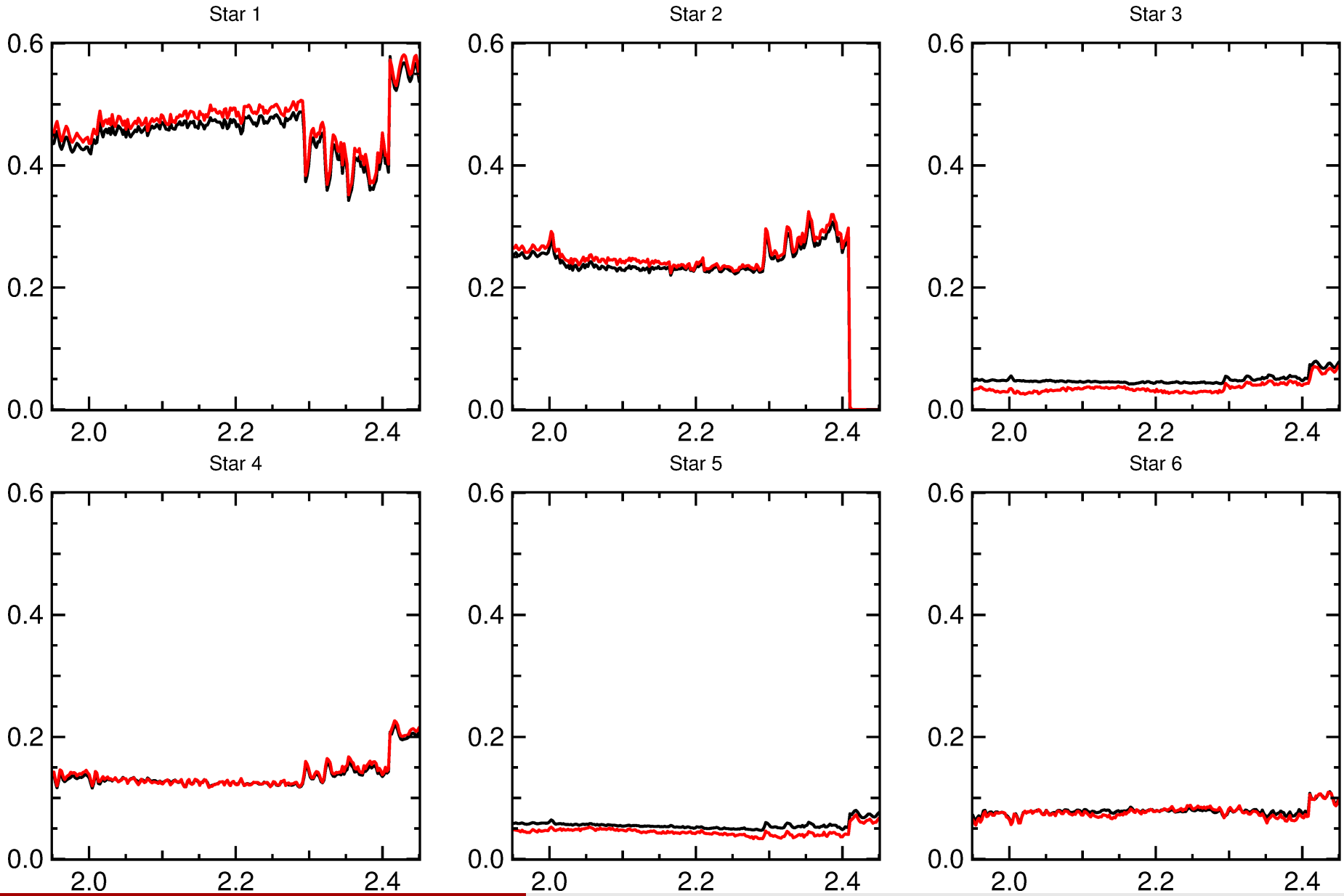
- Few unresolved point sources,
  - $\longrightarrow$ regularization function: structured norm.
- Convex problem,
- Solved using an ADMM based algorithm



Reconstruction (spectrally integrated, axis are in mas).

Positionning errors ≤ 0.1 **mas** ($\lambda/B_{\text{max}} \sim 4$ mas).

Reconstructed spectra



No phase reference

No more complex visibility $\widehat{I}_\ell(\boldsymbol{\nu}_m)$ but:

- powerspectrum:
$$P_m(\boldsymbol{\nu}_m) = \left|\widehat{I}_\ell(\boldsymbol{\nu}_m)\right|^2 ,$$
- bispectrum:
$$B_m(\boldsymbol{\nu}_m, \boldsymbol{\nu}'_m) = \widehat{I}_\ell(\boldsymbol{\nu}_m)\widehat{I}_\ell(\boldsymbol{\nu}'_m)\widehat{I}^*_\ell(\boldsymbol{\nu}_m + \boldsymbol{\nu}'_m) ,$$
- phase closure
$$\Phi_m = \arg B_m ,$$
- differential phase and visibility

Non convex problem solved using slack variables $\boldsymbol{y}$ by alternating minimization

$$\boldsymbol{x}^+ = \arg\min_{\boldsymbol{x}} \mu f_{\text{prior}}(\boldsymbol{x}) + f_{\text{data}}(\boldsymbol{y}|\boldsymbol{m}) \quad \text{s.t.} \begin{cases} \boldsymbol{y} & = \mathbf{H}\cdot\boldsymbol{x} , \\ \boldsymbol{x} & \geq 0 , \\ \sum_\ell \boldsymbol{x} & = 1 . \end{cases}$$

No phase reference

No more complex visibility $\widehat{I_\ell}(\boldsymbol{\nu}_m)$ but:

- powerspectrum:
$$P_m(\boldsymbol{\nu}_m) = \left|\widehat{I_\ell}(\boldsymbol{\nu}_m)\right|^2 ,$$
- bispectrum:
$$B_m(\boldsymbol{\nu}_m, \boldsymbol{\nu}'_m) = \widehat{I_\ell}(\boldsymbol{\nu}_m)\widehat{I_\ell}(\boldsymbol{\nu}'_m)\widehat{I^*_\ell}(\boldsymbol{\nu}_m + \boldsymbol{\nu}'_m) ,$$
- phase closure
$$\Phi_m = \arg B_m ,$$
- differential phase and visibility

Non convex problem solved using slack variables $\boldsymbol{y}$ by alternating minimization

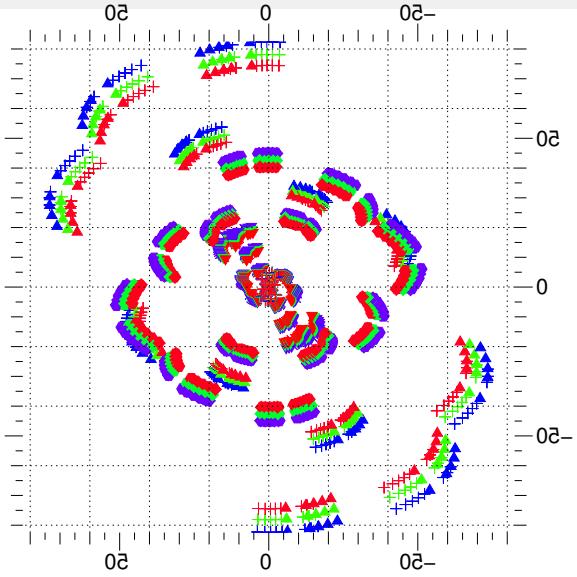
$$\boldsymbol{x}^+ = \arg\min_{\boldsymbol{x}} \mu f_{\text{prior}}(\boldsymbol{x}) + f_{\text{data}}(\boldsymbol{y}|\boldsymbol{m}) \quad \text{s.t.} \begin{cases} \boldsymbol{y} & = \mathbf{H}\cdot\boldsymbol{x} , \\ \boldsymbol{x} & \geq 0 , \\ \sum_\ell \boldsymbol{x} & = 1 . \end{cases}$$

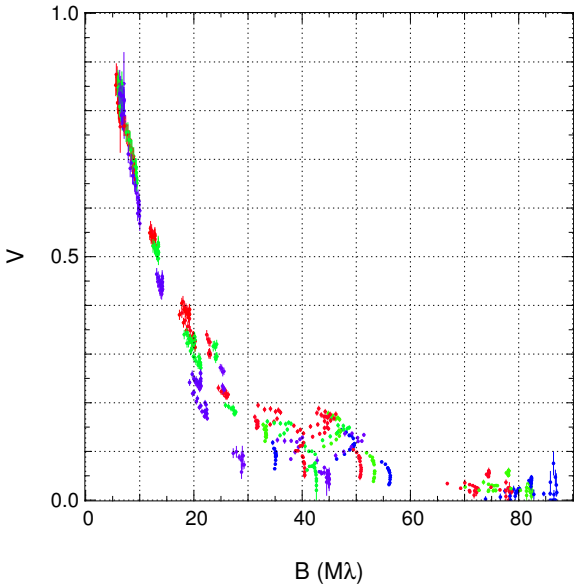
Reconstructing a Mira with MiRA3D

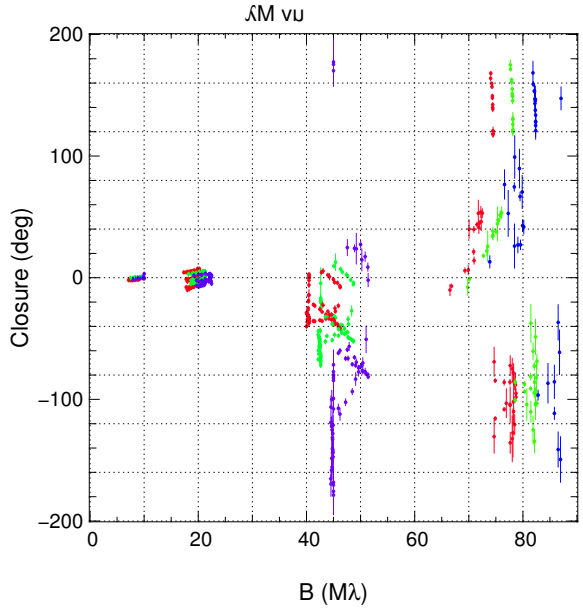
R For observations

(C. Paladini J.B. Le Bouquin *et al.*)

- carbon rich Mira,
- dust obscuration events,
- observed with PIONIER (H band),
- 6 half night within 3 weeks,
- 3 spectral channel



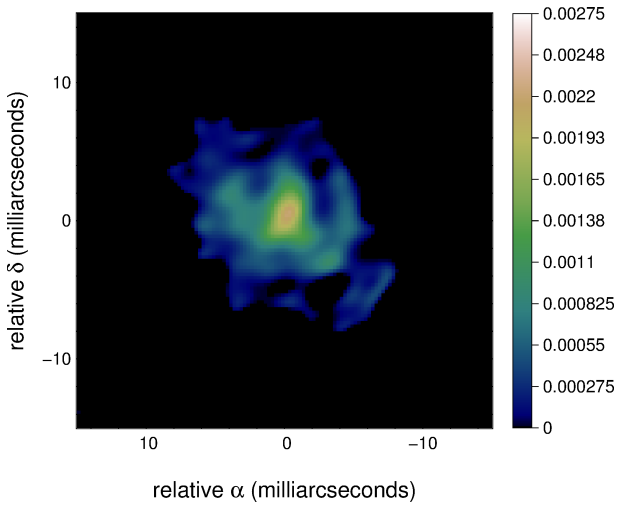




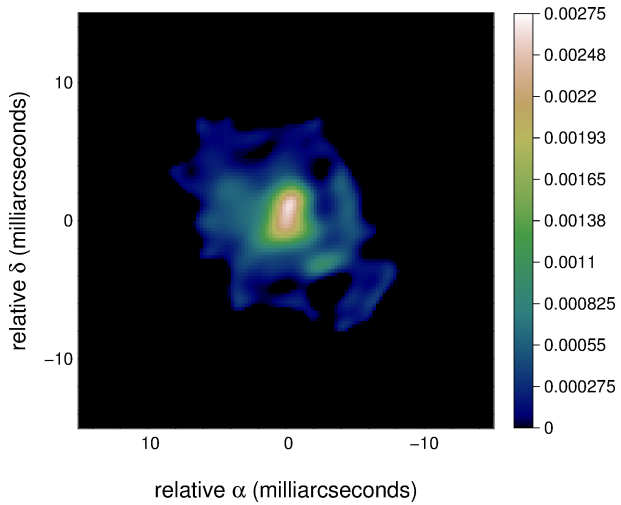
Reconstructing a Mira with MiRA3D

Polychromatic reconstruction with structured norm regularization.

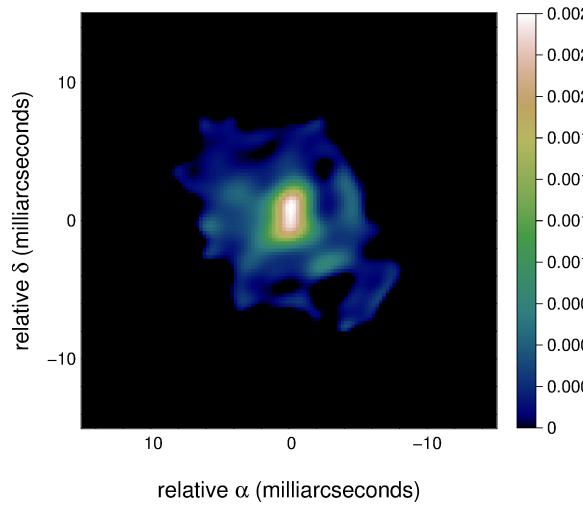
1.59 µm



1.68 µm



1.76 µm



Comparison (credits J-B Le Bouquin)

Once more (chi2~1)

Dynamical model best matching the spectra

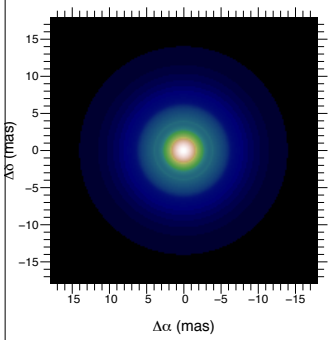
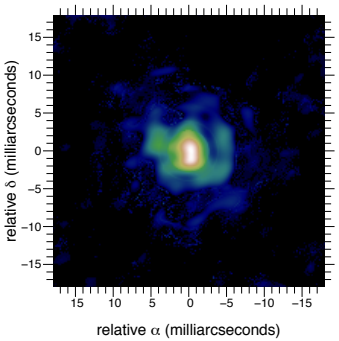


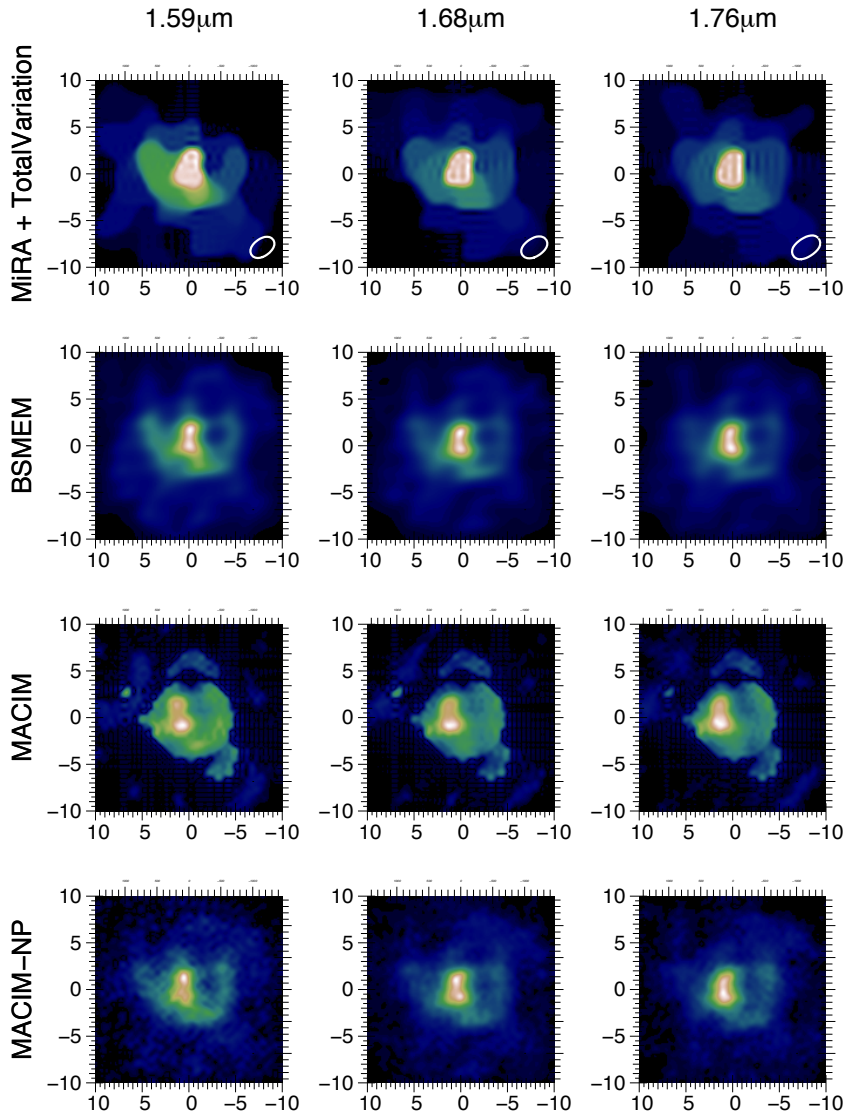
Image reconstruction with prior = model



Somehow, this is the minimal amount of perturbations needed in the model to fit the interferometric data.

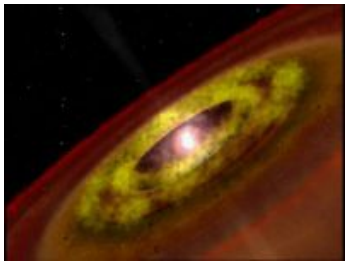
Why/How to deal with the 3 channels together ?

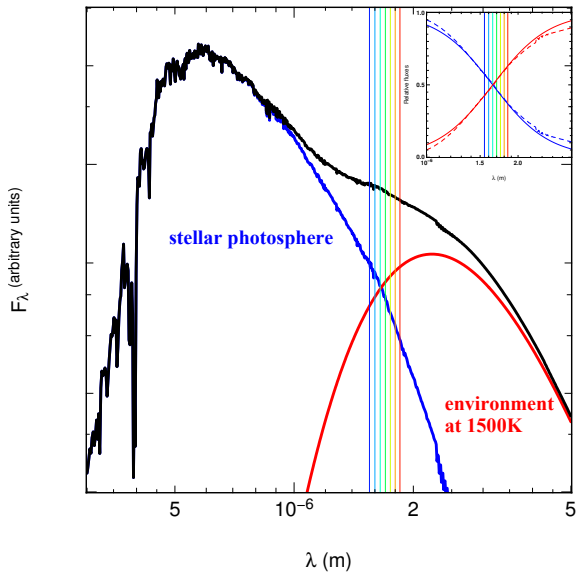
- Improve uv-coverage = reduce degeneracy.
- Force the maps to be as similar as possible.
- ...



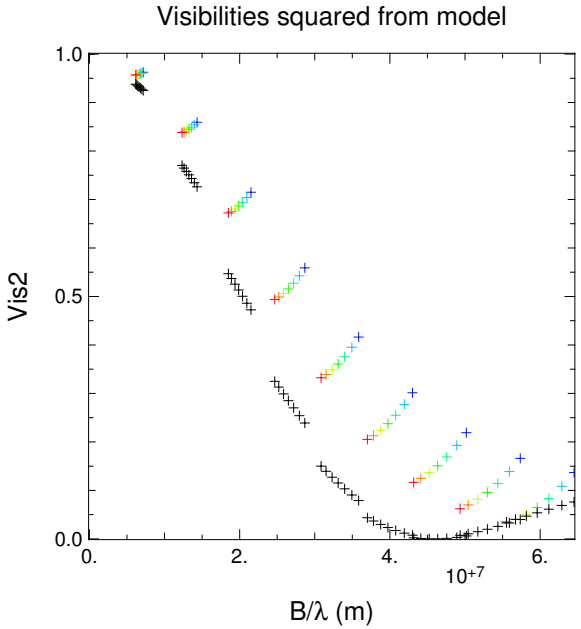
YSO reconstruction (by J. Kluska)

- PIONIER survey of Herbig Ae/Be,
- star + environment,
- strong chromatic effects.





SED of a YSO and PIONIER spectral channels



Chromaticity effect on V2

Strict priors on SED

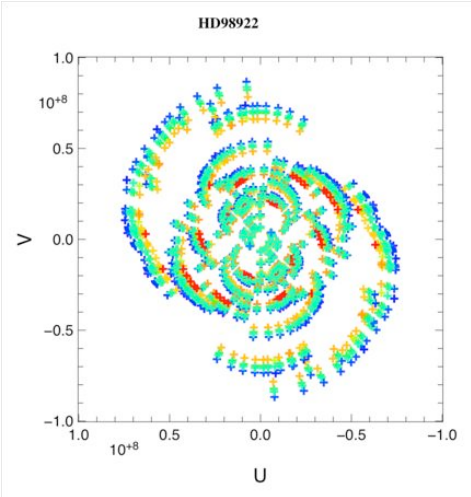
Fitting a MAP of approximate black bodies

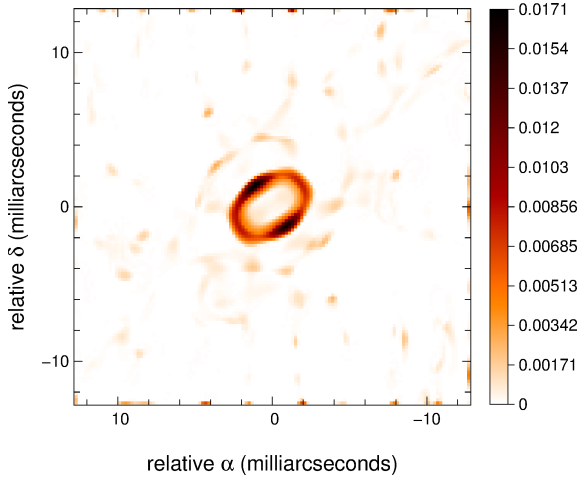
$$\boldsymbol{x}^{+} = \arg\min_{\boldsymbol{x}} f_{\text{data}}(\boldsymbol{y}|\boldsymbol{m}) \quad \text{s.t.} \begin{cases} \boldsymbol{y} & = \mathbf{H} \cdot \boldsymbol{x}\,, \\ \boldsymbol{x} & \geq 0\,, \\ \sum_{\ell} \boldsymbol{x} & = 1\,, \\ (\nabla_{\ell})^3 \boldsymbol{x} & = 0\,. \end{cases}$$

approximation of the SED by a parabola.

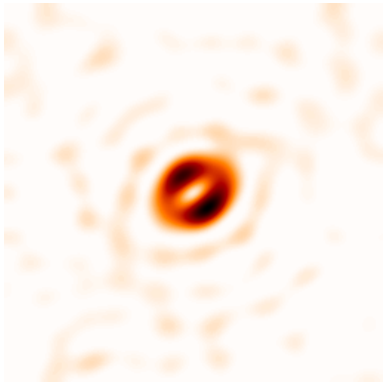
HD98922

- Herbig star type B9Ve,
- Good (u,v) coverage,
- up to 7 spectral channels,





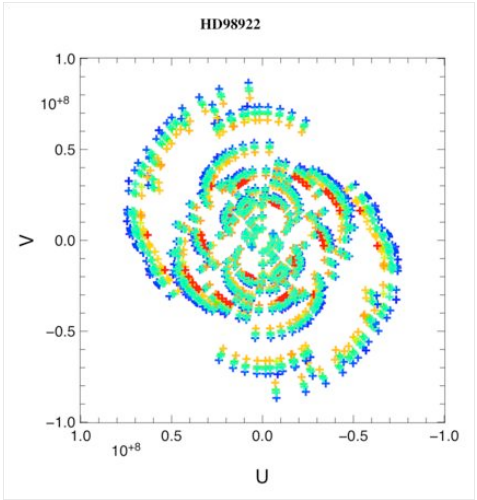
Reconstruction star subtracted

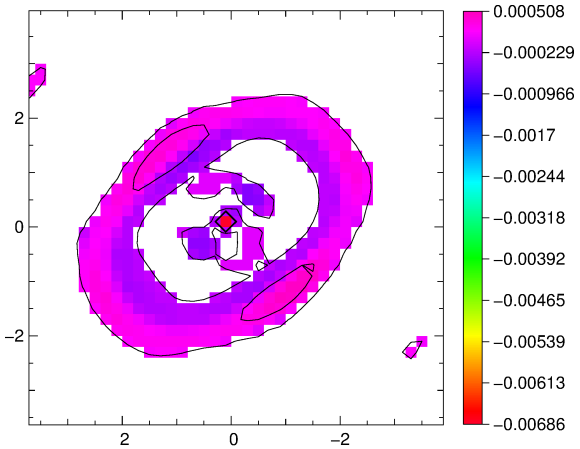


Reconstruction with SPARCO (Kluska *et al.* sub.)

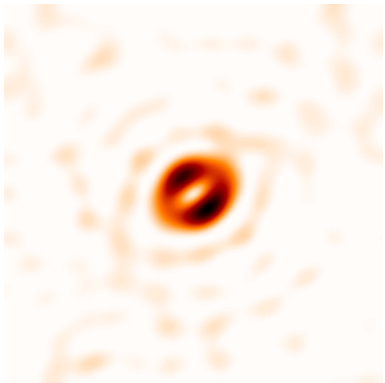
HD98922

- Herbig star type B9Ve,
- Good (u,v) coverage,
- up to 7 spectral channels,





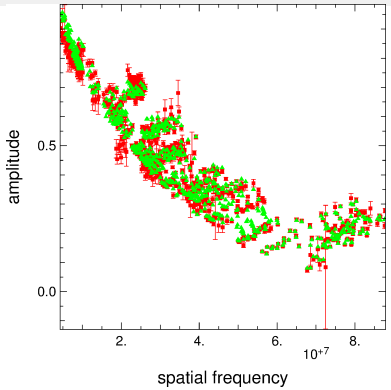
$\propto$ spectral index estimation



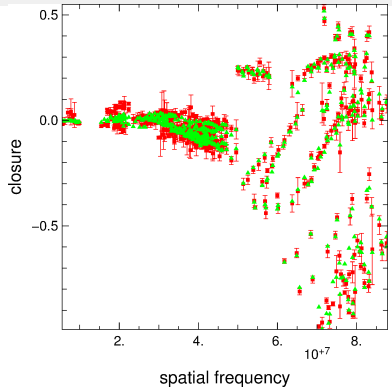
Reconstruction with SPARCO (Kluska *et al.* sub.)

The enigmatic MWC 158

- Herbig Be or B[e] star
- variable,
- subject to outflow and shell ejection?

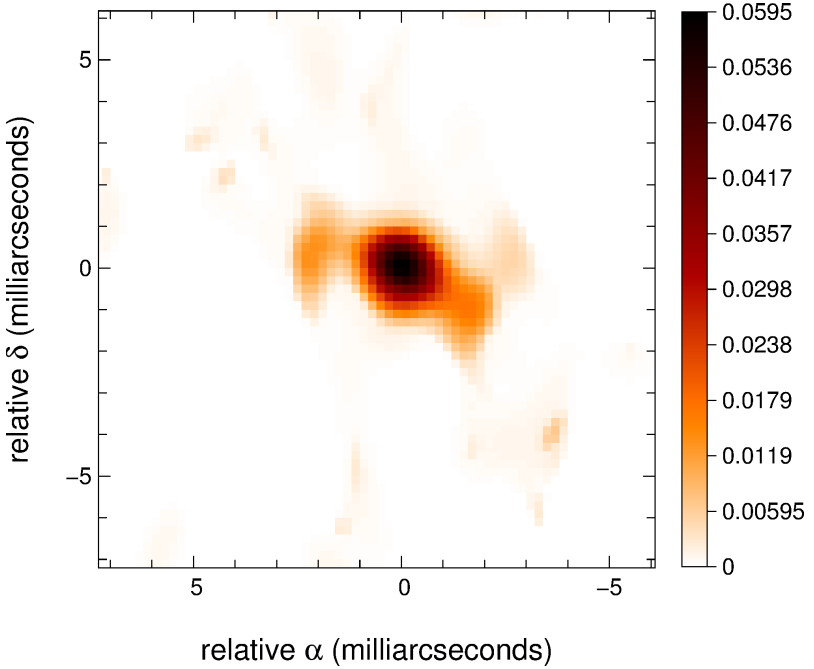


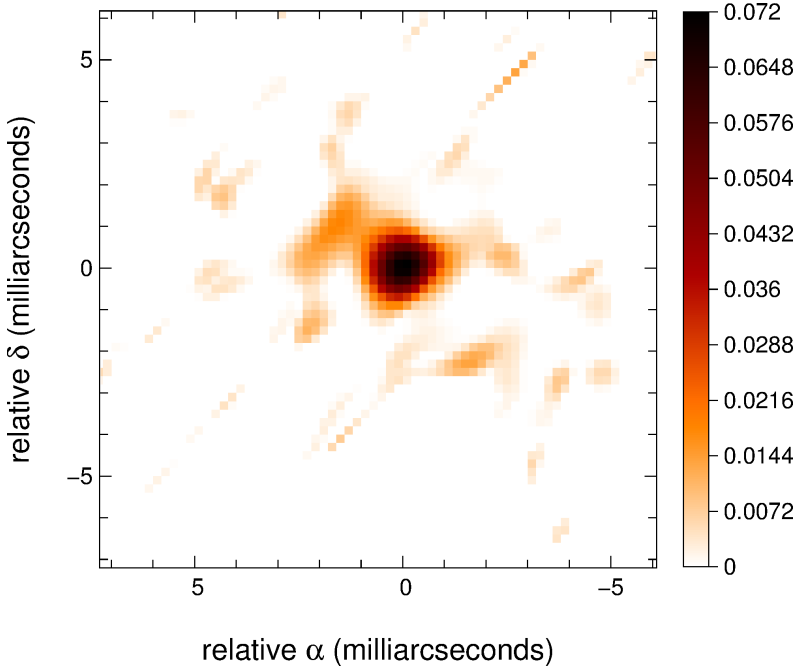
powerspectrum



closure

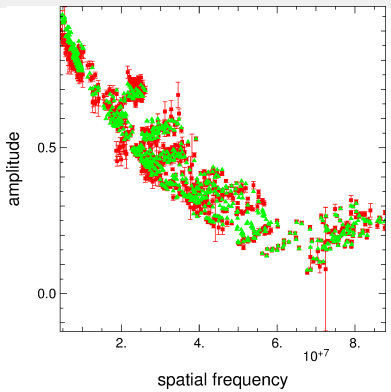
reconstruction H band



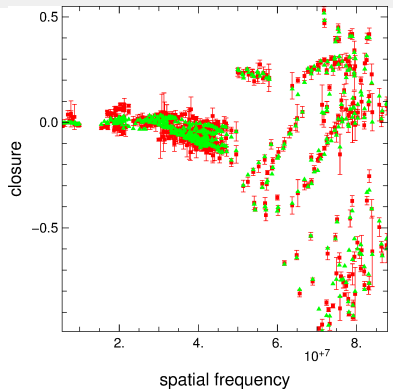


The enigmatic MWC 158

- Herbig Be or B[e] star
- variable,
- subject to outflow and shell ejection?

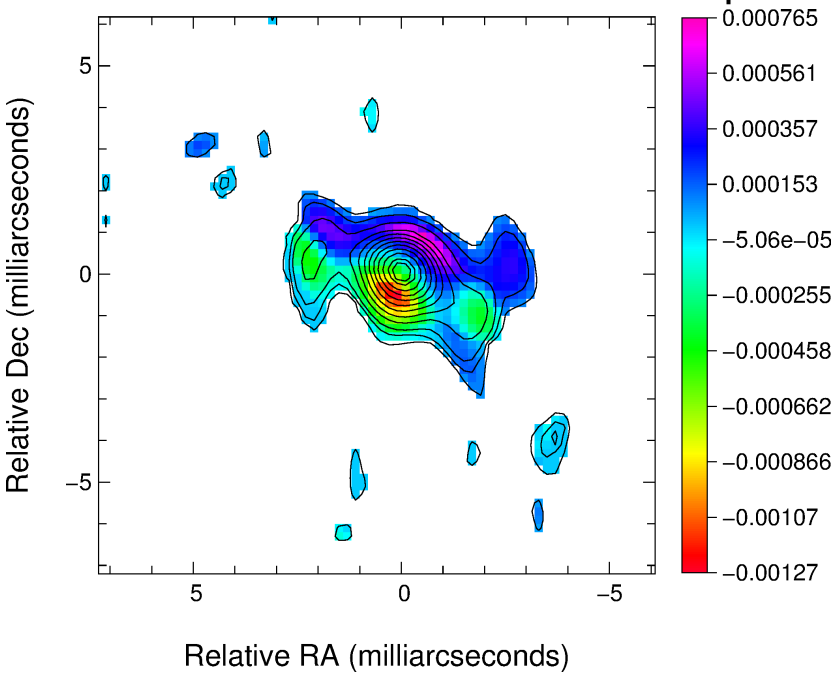


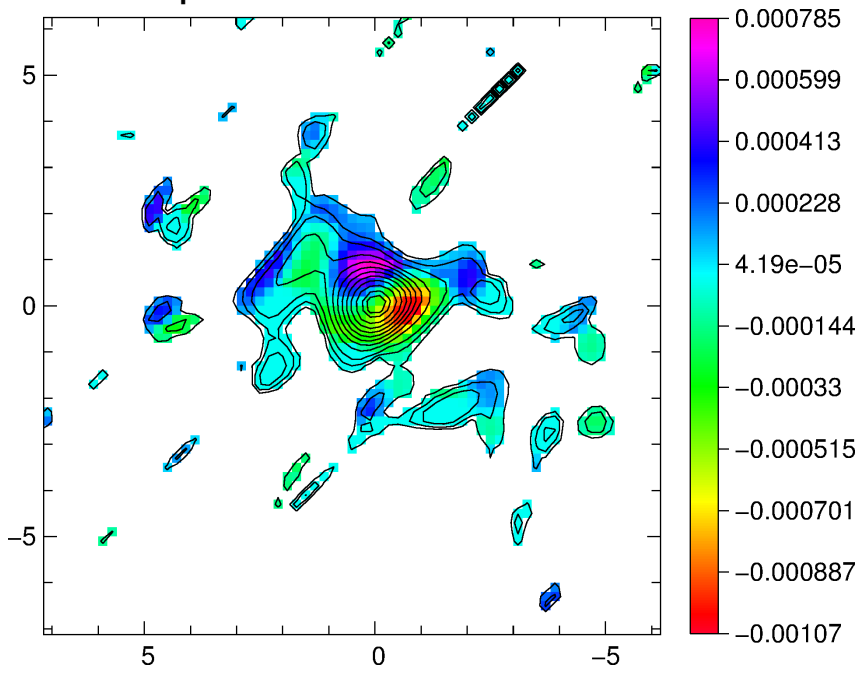
powerspectrum



closure

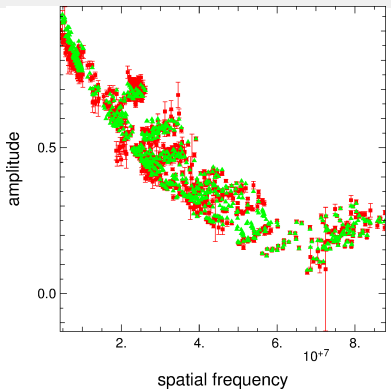
$\propto$ spectral index map



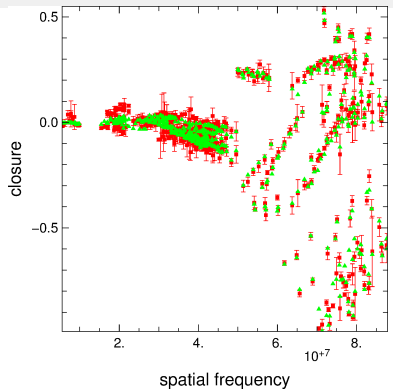


The enigmatic MWC 158

- Herbig Be or B[e] star
- variable,
- subject to outflow and shell ejection?

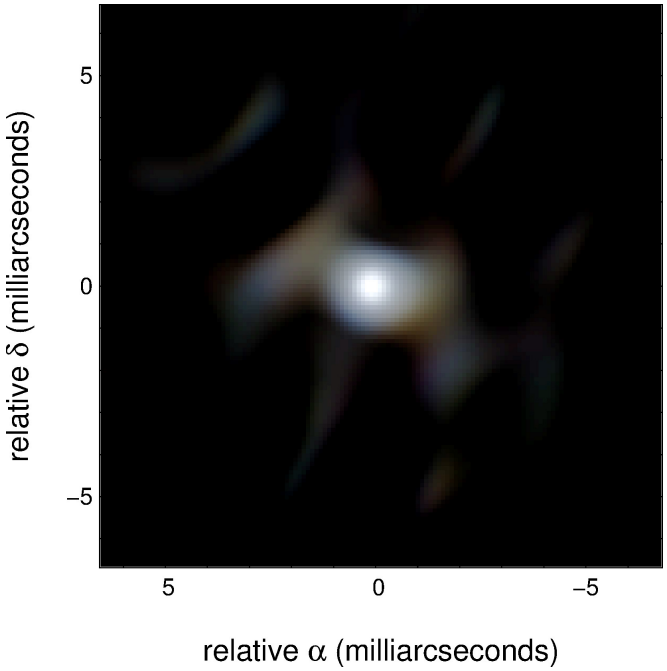


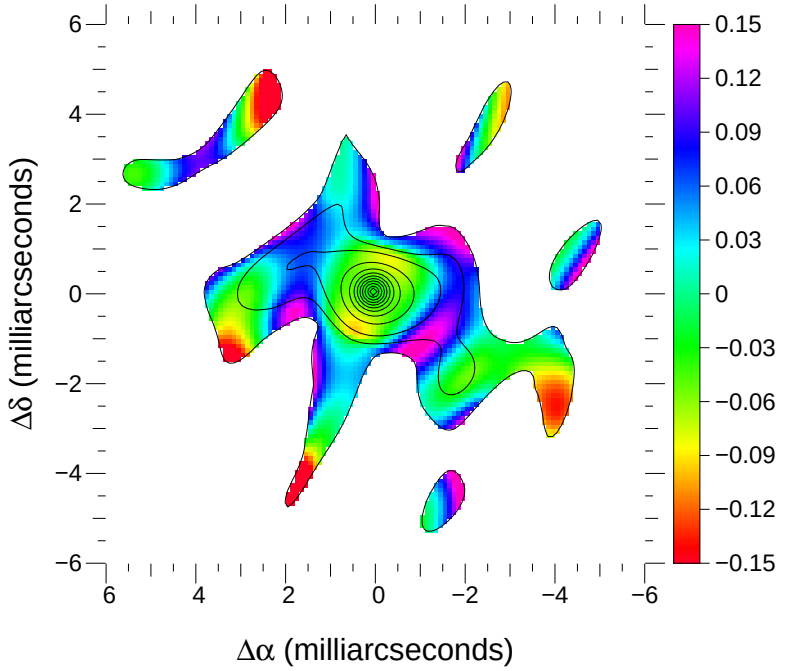
powerspectrum



closure

Both epoch





Conclusion

work still in progress

- MiRA3D: the first reconstruction algorithm dedicated for polychromatic interferometry,
- very flexible, can be adapted to many kind of objects

future works

- evaluation on data with more spectral channels than PIONIER,
- using atmospheric differential refraction law to estimate complex visibilities (autocalibration).

Conclusion

work still in progress

- MiRA3D: the first reconstruction algorithm dedicated for polychromatic interferometry,
- very flexible, can be adapted to many kind of objects

future works

- evaluation on data with more spectral channels than PIONIER,
- using atmospheric differential refraction law to estimate complex visibilities (autocalibration).

General conclusions

Astrophysical image reconstruction:

- increase resolution and SNR ($\propto$ 1 order of magnitude),
- a relatively unexplored field,
- room for improvements,
- a quite cheap way to improve instrument capability

Thank you

Aldering, G., Adam, G., Antilogus, P., Astier, P., Bacon, R., Bongard, S., Bonnaud, C., Copin, Y., Hardin, D., Henault, F., Howell, D. A., Lemonnier, J.-P., Levy, J.-M., Loken, S. C., Nugent, P. E., Pain, R., Pecontal, A., Pecontal, E., Perlmutter, S., Quimby, R. M., Schahmaneche, K., Smadja, G., and Wood-Vasey, W. M. (2002). Overview of the nearby supernova factory. In Tyson, J. A. and Wolff, S., editors, *Survey and Other Telescope Technologies and Discoveries. Edited by Tyson, J. Anthony; Wolff, Sidney. Proceedings of the SPIE, Volume 4836, pp. 61-72 (2002).*, volume 4836 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 61–72.

Benazza-Benyahia, A. and Pesquet, J.-C. (2006). Multichannel image deconvolution in the wavelet transform domain. In *EUSIPCO*.

Bongard, S., Soulez, F., Thiébaut, E., and Pecontal, E. (2011). 3d deconvolution of hyper-spectral astronomical data. *Monthly Notices of the Royal Astronomical Society*, 418(1):258–270.

Daube-Witherspoon, M. E. and Muehllehner, G. (1986). An iterative image space reconstruction algorithm suitable for volume ECT. *IEEE Trans. Med. Imaging*, 5(2):61–66.

Denis, L., Thiébaut, E., and Soulez, F. (2011). Fast model of space-variant blurring and its application to deconvolution in astronomy. In *18th IEEE International Conference on Image Processing*, pages 2873–2876, Bruxelles, France.

Duijster, A., Scheunders, P., and Backer, S. D. (2009). Wavelet-Based EM algorithm for Multispectral-Image restoration. 47(11):3892–3898.

Galatsanos, N. and Chin, R. (1989). Digital restoration of multichannel images. 37(3):415–421.

Hunt, B. R. and Kubler, O. (1984). Karhunen-Loeve multispectral image restoration. i: Theory. 32(3):592–600.

Rodet, T., Orieux, F., Giovannelli, J., and Abergel, A. (2008). Data inversion for Over-Resolved spectral imaging in astronomy. *IEEE Journal of Selected Topics in Signal Processing*, 2(5):802–811.

Soulez, F., Denis, L., Thiébaut, E., and Tourneur, Y. (2012). Blind deconvolution of 3d data in wide field fluorescence microscopy. In *proceedings of the 9th International Symposium on Biological Imaging*, Barcelona.

Soulez, F., Thiébaut, E., and Denis, L. (2013). Restoration of hyperspectral astronomical data with spectrally varying blur. *EAS Publications Series*, 59:403–416.