

Image Reconstruction in Optical Interferometry

Éric Thiébaud

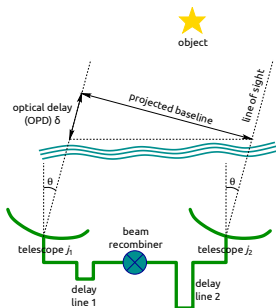
Centre de Recherche Astronomique de Lyon
Université Claude Bernard Lyon I
Ecole Normale Supérieure de Lyon
eric.thiebaut@univ-lyon1.fr

OSA Conference in Toronto, 2011

- 1 Interferometric Observables
- 2 Inverse Approach
- 3 Likelihood of the Data
- 4 Regularization
- 5 Optimization Strategy
- 6 Summary and perspectives

Interferometric Observables

Instantaneous output of an interferometer



instantaneous output = *complex visibility*:

$$V_{j_1, j_2}(\lambda, t) = g_{j_1}^*(\lambda, t) g_{j_2}(\lambda, t) \hat{I}_\lambda(\mathbf{b}_{j_1, j_2}(t)/\lambda)$$

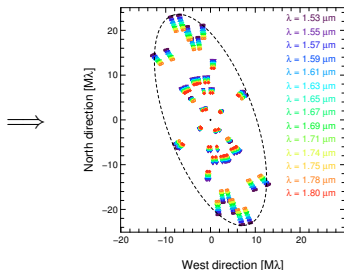
with:

- $g_j(\lambda, t)$ = instantaneous complex amplitude transmission for j th telescope;
- $\hat{I}_\lambda(\boldsymbol{\nu})$ = angular Fourier transform of the specific brightness distribution $I_\lambda(\boldsymbol{\alpha})$ of the observed object in angular direction $\boldsymbol{\alpha}$;
- projected **baseline**:

$$\mathbf{b}_{j_1, j_2}(t) = \mathbf{r}_{j_2}(t) - \mathbf{r}_{j_1}(t)$$

$\mathbf{r}_j(t)$ = position of j th telescope projected on a plane perpendicular to the line of sight;

- λ = wavelength;
- t = time;



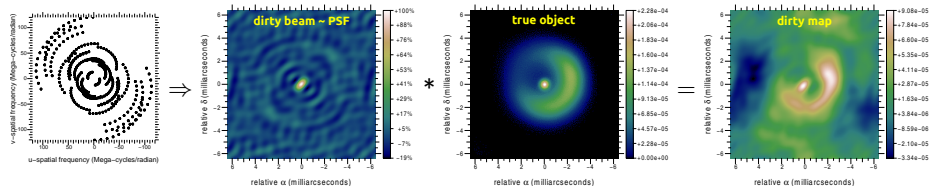
Easy case: image reconstruction \sim deconvolution

At any observed frequency, $\nu_k = \mathbf{b}_{j_1, j_2}(t_m)/\lambda_\ell$, the data is given by:

$$z_k = \hat{h}_k \hat{I}_{\lambda_\ell}(\nu_k) + \text{noise}$$

with the transfer function (the Fourier transform of the *dirty beam*):

$$\hat{h}_k = g_{j_1}^*(\lambda_\ell, t_m) g_{j_2}(\lambda_\ell, t_m)$$



when the complex visibilities and the complex throughput are available:

image reconstruction \sim deconvolution

- ⚠ many missing values (very sparse data)
- \Rightarrow other constraints (**priors**) than the data are required

The effects of turbulence

Because of the **atmospheric turbulence**, averaging during an exposure yields:

$$\langle V_{j_1, j_2}(\lambda, t) \rangle_m = \left\langle g_{j_1}^*(\lambda, t) g_{j_2}(\lambda, t) \widehat{I}_\lambda(\mathbf{v}_{j_1, j_2}(\lambda, t)) \right\rangle_m \quad \langle \dots \rangle_m \text{ means averaging during } m\text{th exposure}$$

$$\approx \underbrace{\langle g_{j_1}(\lambda, t) \rangle_m^*}_{\approx 0} \underbrace{\langle g_{j_2}(\lambda, t) \rangle_m}_{\approx 0} \widehat{I}_\lambda(\mathbf{b}_{j_1, j_2, m}/\lambda)$$

with: $\mathbf{b}_{j_1, j_2, m} \stackrel{\text{def}}{=} \langle \mathbf{r}_{j_2}(t) \rangle_m - \langle \mathbf{r}_{j_1}(t) \rangle_m$ the mean baseline during the exposure

⇒ we need to integrate observables which are **insensitive to phase delay errors**:

- **powerspectrum**

$$\langle |V_{j_1, j_2}(\lambda, t)|^2 \rangle_m \approx \underbrace{\langle |g_{j_1}(\lambda, t)|^2 \rangle_m \langle |g_{j_2}(\lambda, t)|^2 \rangle_m}_{> 0} |\widehat{I}_\lambda(\mathbf{b}_{j_1, j_2, m}/\lambda)|^2$$

- **bispectrum**

$$\langle V_{j_1, j_2}(\lambda, t) V_{j_2, j_3}(\lambda, t) V_{j_3, j_1}(\lambda, t) \rangle_m \approx \underbrace{\langle |g_{j_1}(\lambda, t)|^2 \rangle_m \langle |g_{j_2}(\lambda, t)|^2 \rangle_m \langle |g_{j_3}(\lambda, t)|^2 \rangle_m}_{> 0}$$

$$\widehat{I}_\lambda(\mathbf{b}_{j_1, j_2, m}/\lambda) \widehat{I}_\lambda(\mathbf{b}_{j_2, j_3, m}/\lambda) \widehat{I}_\lambda(\mathbf{b}_{j_3, j_1, m}/\lambda)$$

Issues in image reconstruction from optical interferometry data

- 1 **sparsity of the data**
(holes in the spatial frequency coverage ►)
⇒ additional **prior** needed

- 2 **non-linear data**

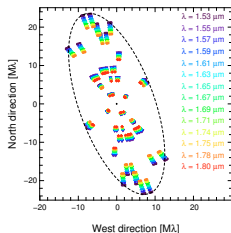
$$\text{powerspectrum} \propto |\widehat{I}_\lambda(\boldsymbol{\nu}_k)|^2$$

$$\text{bispectrum} \propto \widehat{I}_\lambda(\boldsymbol{\nu}_{k_1}) \widehat{I}_\lambda(\boldsymbol{\nu}_{k_2}) \widehat{I}_\lambda^*(\boldsymbol{\nu}_{k_1} + \boldsymbol{\nu}_{k_2})$$

- 3 **calibration** of the effective transfer functions

- 4 **missing Fourier phases**

- powerspectrum provides no phase
- **phase closure** (the phase of the bispectrum)
only provide **1 phase out of 3**



Inverse Approach

Inverse approach for image reconstruction

Inverse approach provides a very general framework to describe most (if not all) image reconstruction algorithms (le Besnerais et al. 2008; Thiébaud 2009; Thiébaud and Giovannelli 2010).

The recipe involves:

- 1 a **direct model**: model of the brightness distribution and its Fourier transform;
- 2 a **criterion** to determine a unique and stable solution;
- 3 an **optimization strategy** to find the solution.

Image and complex visibilities models

Image model

The specific brightness distribution in angular direction α is approximated by:

$$I_\lambda(\alpha) \approx \sum_n b_n(\alpha) x_n \xrightarrow{\text{F.T.}} \hat{I}_\lambda(\nu) \approx \sum_n \hat{b}_n(\nu) x_n$$

with $\{b_n : \mathbb{R}^2 \mapsto \mathbb{R}\}_{n=1}^N$ a basis of functions and $x \in \mathbb{R}^N$ the **image parameters**.

Complex visibility model

For any sampled spatial frequency $\nu_k = \mathbf{b}_{j_1, j_2, m} / \lambda$ the model complex visibility writes:

$$\hat{I}_\lambda(\nu_k) \approx y_k = \sum_n \hat{b}_n(\nu_k) x_n = \sum_n H_{k,n} x_n$$

with $H_{k,n} = \hat{b}_n(\nu_k)$, in matrix notation:

$$\mathbf{y} = \mathbf{H} \cdot \mathbf{x}$$

with $\mathbf{y} \in \mathbb{C}^K$ and $\mathbf{H} \in \mathbb{C}^{K \times N}$ is a sub-sampled Fourier transform operator.

Image constraints

Image reconstruction is a compromise between various constraints (Thiébaud 2009).

Data constraints

The image must be **compatible with the data** z (powerspectrum, bispectrum, etc.):

$$f_{\text{data}}(\mathbf{H}\cdot\mathbf{x}) \stackrel{\text{def}}{=} -\log \text{pdf}(z|\mathbf{H}\cdot\mathbf{x}) + c \leq \eta$$

with $\text{pdf}(z|\mathbf{H}\cdot\mathbf{x})$ the **likelihood** of the data given the model and $\eta > 0$.



Even with $\eta = 0$, this is insufficient to define a unique (and stable) solution, we need additional a priori constraints:

Strict priors

e.g. the image must be non-negative and normalized

$$\forall n, x_n \geq 0 \quad \text{and} \quad \sum_n x_n = 1$$

$$\iff \mathbf{x} \in \mathbb{X} \quad (\text{the feasible set})$$

Loose priors

e.g. the image must be *simple* or *smooth*

$$\min_{\mathbf{x} \in \mathbb{X}} f_{\text{prior}}(\mathbf{x})$$

Inverse problem formulation

We want to follow the priors as far as possible providing the image remains compatible with the data:

$$\mathbf{x}^+ = \arg \min_{\mathbf{x} \in \mathbb{X}} f_{\text{prior}}(\mathbf{x}) \quad \text{s.t.} \quad f_{\text{data}}(\mathbf{H} \cdot \mathbf{x}) \leq \eta$$

which can be solved via the Lagrangian:

$$\mathcal{L}(\mathbf{x}; \ell) = f_{\text{prior}}(\mathbf{x}) + \ell f_{\text{data}}(\mathbf{H} \cdot \mathbf{x})$$

with $\ell \geq 0$ the Lagrange multiplier for the inequality constraint $f_{\text{data}}(\mathbf{H} \cdot \mathbf{x}) \leq \eta$. The inequality constraint must be active, hence $\ell > 0$ and, taking $\mu = 1/\ell$, leads to solve:

Maximum a posteriori solution

$$\mathbf{x}(\mu) = \arg \min_{\mathbf{x} \in \mathbb{X}} f(\mathbf{x}; \mu)$$

$$\text{with: } f(\mathbf{x}; \mu) = \underbrace{f_{\text{data}}(\mathbf{H} \cdot \mathbf{x})}_{\text{likelihood}} + \underbrace{\mu f_{\text{prior}}(\mathbf{x})}_{\text{regularization}}$$

where $\mu > 0$ is tuned so as to match $f_{\text{data}}(\mathbf{H} \cdot \mathbf{x}^+) = \eta$ with $\mathbf{x}^+ = \mathbf{x}(\mu^+)$ and μ^+ the optimal regularization weight.

Likelihood of the Data

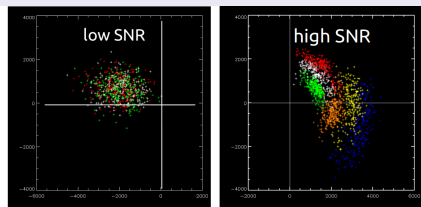
Likelihood of the data

- should be based on the noise statistics of the data:

$$f_{\text{data}}(\mathbf{H}\cdot\mathbf{x}) \stackrel{\text{def}}{=} -\log \text{pdf}(z|\mathbf{H}\cdot\mathbf{x}) + c$$

- can be very complicated (non-convex, phase wrapping, etc.)
 - various approximations have been proposed (e.g., Meimon et al. 2005a)
- ⚠ in general this does not amount to least-squares (even weighted ones!)

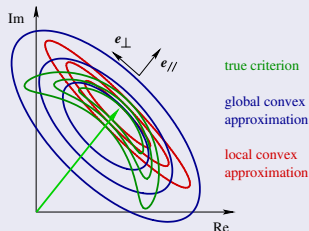
Real complex data



(triple product of FKV0509)

Source : C. Hummel et al. <<http://www.mrao.cam.ac.uk/~jsy1001/exchange/complex/complex.html>>

Approximate cost function

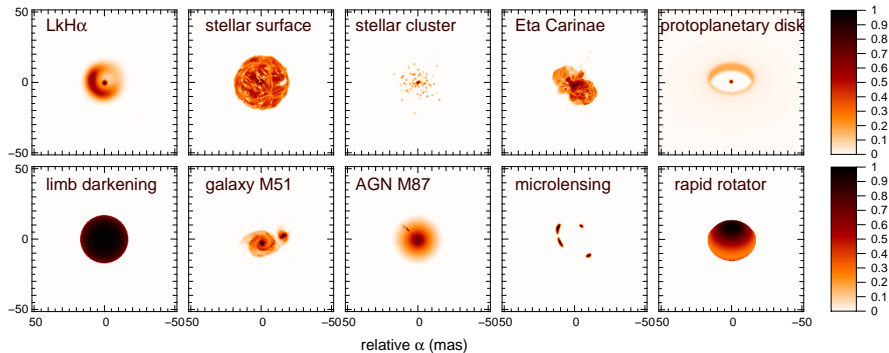


Regularization

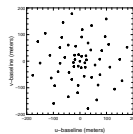
Which are the best regularization methods?

Practical comparison of regularization methods:

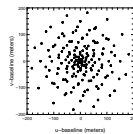
- a study made by S. Renard, É. Thiébaud and F. Malbet (to appear in *Astron. & Astrophys.*, 2011);
- about **1000** simulations:
 - 10 different objects;
 - 11 different regularizations;
 - 3 different (u, v) coverages: *poor* (31 freq.), *medium* (88 freq.), and *rich* (245 freq.);
 - 3 different signal-to-noise ratios (SNR): *high* (1%), *medium* (5%), and *low* (10%);
- assumptions: complex visibilities available
⇒ **convex** constrained non-linear optimization problem;
- algorithm: **MIRA** (Thiébaud, 2008, 2009);



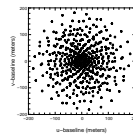
Various objects \blacktriangle and
 (u, v) coverages \blacktriangleright
 (Renard et al. 2011).



31 freq.



88 freq.



245 freq.

Various regularizations

We consider the following regularizations:

1. Quadratic smoothness:

$$f_{\text{prior}}(\mathbf{x}) = \|\mathbf{x} - \mathbf{S} \cdot \mathbf{x}\|^2$$

where \mathbf{S} is a smoothing operator (by finite differences).

2-3. Compactness (le Besnerais et al. 2008):

$$f_{\text{prior}}(\mathbf{x}) = \sum_n w_n^{\text{prior}} \mathbf{x}_n^2$$

with $w_n^{\text{prior}} = \|\boldsymbol{\theta}_n\|^\beta$ and $\beta = 2$ or 3 yields *spectral smoothness*.

4-5. Non-linear smoothness:

$$f_{\text{prior}}(\mathbf{x}) = \sum_n \sqrt{\|\nabla x_n\|^2 + \epsilon^2}$$

where $\|\nabla x_n\|^2$ is the squared magnitude of the spatial gradient in the image at n th pixel and $\epsilon \rightarrow 0$ yields **total variation** (Strong and Chan 2003) while $\epsilon > 0$ yields **edge-preserving smoothness** (Charbonnier et al. 1997).

Various regularizations (continued)

6-8. Separable norms (ℓ_p):

$$f_{\text{prior}}(\mathbf{x}) = \sum_n (x_n^2 + \epsilon^2)^{p/2} \approx \sum_n |x_n|^p$$

where $\epsilon > 0$ and $p = 1.5, 2$, and 3 . Note that $p = 1$ is what is advocated in **compress sensing** (Candes et al. 2006) while $p = 2$ corresponds to regular **Tikhonov regularization**.

9-11. Maximum entropy methods (Narayan and Nityananda 1986):

$$f_{\text{prior}}(\mathbf{x}) = - \sum_n h(x_n; \bar{x}_n).$$

Here the prior is to assume that the image is drawn toward a prior model \bar{x} according to a non quadratic potential h , called the **entropy**:

$$\text{MEM-sqrt: } h(x; \bar{x}) = \sqrt{x};$$

$$\text{MEM-log: } h(x; \bar{x}) = \log(x);$$

$$\text{MEM-prior: } h(x; \bar{x}) = x - \bar{x} - x \log(x/\bar{x}).$$

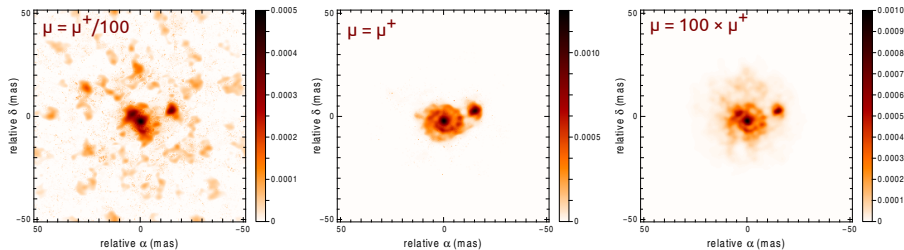
Tuning the regularization level

We choose the regularization level μ^+ by minimizing the *mean squared error* (MSE) of the reconstruction versus the true image:

$$\mu^+ = \arg \min_{\mu > 0} \left\| \mathbf{x}(\mu) - \mathbf{x}^{\text{true}} \right\|_2$$

where

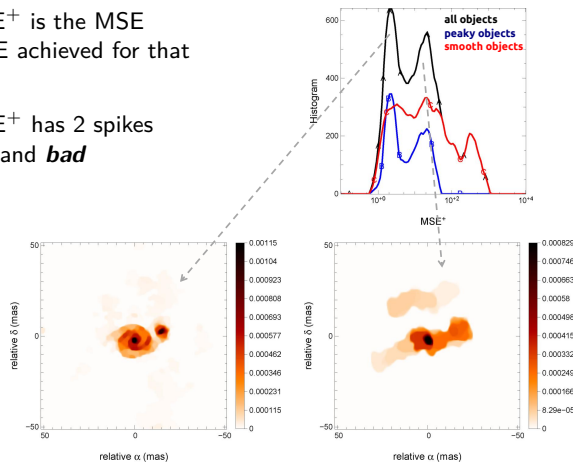
$$\mathbf{x}(\mu) \stackrel{\text{def}}{=} \arg \min_{\mathbf{x} \in \mathbb{X}} \{f_{\text{data}}(\mathbf{H} \cdot \mathbf{x}) + \mu f_{\text{prior}}(\mathbf{x})\}$$



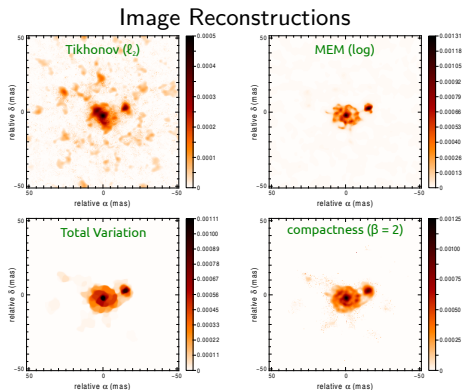
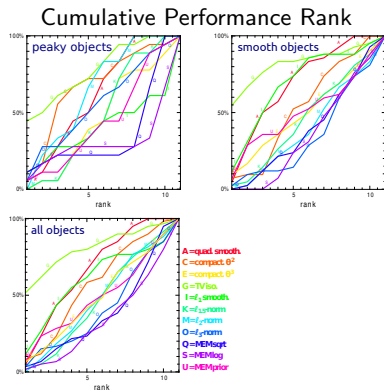
Is the MSE^+ a good figure of merit?

For a given object, MSE^+ is the MSE divided by the best MSE achieved for that object.

The distribution of MSE^+ has 2 spikes corresponding to *good* and *bad* reconstructions.



And the winner is...



Based on cumulative rank, **TV** and **compactness** are the most successful.

However the best prior depend on the particular case (object type, SNR and coverage).

Optimization Strategy

Image reconstruction = optimization problem

Assuming $\mu^+ = 1$, image reconstruction amounts to solve:

$$\min_{\mathbf{x} \in \mathbb{X}} \underbrace{\{f_{\text{prior}}(\mathbf{x}) + f_{\text{data}}(\mathbf{H} \cdot \mathbf{x})\}}_{f(\mathbf{x})}$$

For optical interferometric data, the joint criterion $f(\mathbf{x})$ is:

- highly non-linear (means *non-quadratic*);
- depending on a very large number of parameters (the image pixels);
- multimodal \implies in principle, needs *global optimization* or a good starting point followed by continuous optimization;

Proposed methods:

- matching-pursuit: **CLEAN** (Fomalont 1973; Högbom 1974), the **building-blocks** method (Hofmann and Weigelt 1993)
- self-calibration: **Wisard** (Meimon et al. 2005b);
- direct optimization: **BSMEM** (Baron and Young 2008), **MIRA** (Thiébaud 2008);
- global optimization: **MACIM** (Markov Chain Imager, Ireland et al. 2008);

Self-calibration

Self-calibration (Readhead and Wilkinson 1978; Schwab 1980; Cornwell and Wilkinson 1981) proposed to solve for missing calibration of the transfer function or missing Fourier phases.

Self-calibration algorithm

Choose an initial image $\mathbf{x}^{[0]}$ and repeat the following steps for $k = 0, 1, \dots$ until convergence:

❶ **self-calibration step:**

$$\mathbf{y}^{[k+1]} = \arg \min_{\mathbf{y}} f_{\text{data}}(\mathbf{y}) \quad \text{s.t.} \quad \mathbf{y} \approx \mathbf{H} \cdot \mathbf{x}^{[k]}$$

❷ **image reconstruction step (deconvolution):**

$$\mathbf{x}^{[k+1]} = \arg \min_{\mathbf{x} \in \mathbb{X}} f_{\text{prior}}(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{H} \cdot \mathbf{x} \approx \mathbf{y}^{[k+1]}$$

Issues:

- What is the meaning of \approx (depends on the algorithm)?
- How to consistently tune the balance between prior and data?
- Not rigorously equivalent to minimizing a given criterion.

Augmented Lagrangian approach

Solving the image reconstruction problem by **direct minimization** of the criterion, *i.e.*

$$\min_{\mathbf{x} \in \mathbb{X}} \{f_{\text{prior}}(\mathbf{x}) + f_{\text{data}}(\mathbf{H} \cdot \mathbf{x})\}$$

is exactly the same as solving the **constrained problem**:

$$\min_{\mathbf{x} \in \mathbb{X}, \mathbf{y}} \{f_{\text{prior}}(\mathbf{x}) + f_{\text{data}}(\mathbf{y})\} \quad \text{s.t.} \quad \mathbf{H} \cdot \mathbf{x} = \mathbf{y}$$

where the **model complex visibilities** $\mathbf{y} = \mathbf{H} \cdot \mathbf{x}$ have been explicitly introduced as **auxiliary variables**.

The **augmented Lagrangian** (Boyd et al. 2010) is a practical algorithm to solve this constrained problem:

$$\mathcal{L}_A(\mathbf{x}, \mathbf{y}, \mathbf{u}; \beta) = f_{\text{prior}}(\mathbf{x}) + f_{\text{data}}(\mathbf{y}) - \mathbf{u}^T \cdot [\mathbf{H} \cdot \mathbf{x} - \mathbf{y}] + \frac{\beta}{2} \|\mathbf{H} \cdot \mathbf{x} - \mathbf{y}\|^2,$$

with \mathbf{u} the Lagrange multipliers related to the constraints $\mathbf{H} \cdot \mathbf{x} = \mathbf{y}$ and $\beta > 0$ the weight of the quadratic penalty to reinforce the constraints.

Advantages: explicit update formula for the Lagrange multipliers, strong convergence properties for β large enough (no needs for $\beta \rightarrow \infty$), *etc.*

Augmented Lagrangian approach (continued)

$$\mathcal{L}_A(\mathbf{x}, \mathbf{y}, \mathbf{u}; \beta) = f_{\text{prior}}(\mathbf{x}) + f_{\text{data}}(\mathbf{y}) - \mathbf{u}^T \cdot [\mathbf{H} \cdot \mathbf{x} - \mathbf{y}] + \frac{\beta}{2} \|\mathbf{H} \cdot \mathbf{x} - \mathbf{y}\|^2$$

Augmented Lagrangian algorithm (in our case)

Start with initial multipliers $\mathbf{u}^{[0]}$ and $\beta^{[0]} > 0$ and repeat the following steps for $k = 0, 1, \dots$ until convergence:

- 1 improve the variables:

$$\{\mathbf{x}, \mathbf{y}\}^{[k+1]} \approx \arg \min_{\mathbf{x} \in \mathbb{X}, \mathbf{y}} \mathcal{L}_A(\mathbf{x}, \mathbf{y}, \mathbf{u}^{[k]}; \beta^{[k]})$$

- 2 update the multipliers:

$$\mathbf{u}^{[k+1]} = \mathbf{u}^{[k]} + \beta \left(\mathbf{y}^{[k+1]} - \mathbf{H} \cdot \mathbf{x}^{[k+1]} \right)$$

$$\beta^{[k+1]} = \beta^{[k]}$$

- or strengthen the constraints:

$$\mathbf{u}^{[k+1]} = \mathbf{u}^{[k]}$$

$$\beta^{[k+1]} = \gamma \beta^{[k]} \quad (\text{with } \gamma > 1)$$

Step 1 can be implemented thanks to alternating minimization, e.g.:

$$\mathbf{x}^+ = \arg \min_{\mathbf{x} \in \mathbb{X}} \mathcal{L}_A(\mathbf{x}, \mathbf{y}, \mathbf{u}; \beta) \quad \text{followed by} \quad \mathbf{y}^+ = \arg \min_{\mathbf{y}} \mathcal{L}_A(\mathbf{x}^+, \mathbf{y}, \mathbf{u}; \beta)$$

Image reconstruction step in augmented Lagrangian approach

The augmented Lagrangian can be rewritten as:

$$\begin{aligned}\mathcal{L}_A(\mathbf{x}, \mathbf{y}, \mathbf{u}; \beta) &= f_{\text{prior}}(\mathbf{x}) + f_{\text{data}}(\mathbf{y}) - \mathbf{u}^T \cdot [\mathbf{H} \cdot \mathbf{x} - \mathbf{y}] + \frac{\beta}{2} \|\mathbf{H} \cdot \mathbf{x} - \mathbf{y}\|^2, \\ &= f_{\text{prior}}(\mathbf{x}) + f_{\text{data}}(\mathbf{y}) + \frac{\beta}{2} \|\mathbf{H} \cdot \mathbf{x} - \mathbf{y} - \mathbf{u}/\beta\|^2 - \frac{1}{2\beta} \|\mathbf{u}\|^2.\end{aligned}$$

Improving \mathbf{x} given the other variables writes:

$$\begin{aligned}\mathbf{x}^+ &= \arg \min_{\mathbf{x} \in \mathbb{X}} \mathcal{L}_A(\mathbf{x}, \mathbf{y}, \mathbf{u}; \beta) \\ &= \arg \min_{\mathbf{x} \in \mathbb{X}} \left\{ f_{\text{prior}}(\mathbf{x}) + \frac{\beta}{2} \|\mathbf{H} \cdot \mathbf{x} - \mathbf{v}\|^2 \right\} \quad \text{with} \quad \mathbf{v} = \mathbf{y} + \mathbf{u}/\beta.\end{aligned}$$

which is the analogous of **image reconstruction** given *pseudo-complex visibilities* $\mathbf{v} = \mathbf{y} + \mathbf{u}/\beta$ with white noise of variance $\propto \beta^{-1/2}$ (unlike self-calibration which would try to fit \mathbf{y}).

Calibration step in augmented Lagrangian approach

Recalling that the augmented Lagrangian can be rewritten as:


$$\mathcal{L}_A(\mathbf{x}, \mathbf{y}, \mathbf{u}; \beta) = f_{\text{prior}}(\mathbf{x}) + f_{\text{data}}(\mathbf{y}) + \frac{\beta}{2} \|\mathbf{H} \cdot \mathbf{x} - \mathbf{y} - \mathbf{u}/\beta\|^2 - \frac{1}{2\beta} \|\mathbf{u}\|^2,$$

improving \mathbf{y} given the other variables writes:

$$\begin{aligned} \mathbf{y}^+ &= \arg \min_{\mathbf{y}} \mathcal{L}_A(\mathbf{x}, \mathbf{y}, \mathbf{u}; \beta) \\ &= \arg \min_{\mathbf{y}} \left\{ f_{\text{data}}(\mathbf{y}) + \frac{\beta}{2} \|\mathbf{y} - \mathbf{w}\|^2 \right\} \quad \text{with} \quad \mathbf{w} = \mathbf{H} \cdot \mathbf{x} - \mathbf{u}/\beta. \end{aligned}$$

which is similar to the self-calibration step in self-calibration methods except that the complex visibilities \mathbf{y} are enforced to fit the actual data and the *shifted* model complex visibilities $\mathbf{w} = \mathbf{H} \cdot \mathbf{x} - \mathbf{u}/\beta$ and not just $\mathbf{H} \cdot \mathbf{x}$.

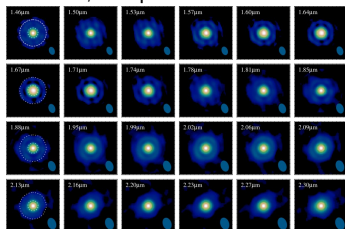
Conclusions about optimization strategy

- **direct optimization** is more consistent (the given criterion is minimized) and much faster and stable than *self-calibration* for finding missing Fourier phases (as in Wisard, Meimon et al. 2005b) or missing parameters in the OTF:
 - imposing $\mathbf{u} = \mathbf{0}$ for the Lagrange multipliers yields the same method as *self-calibration*;
 - exactly matching $\mathbf{H} \cdot \mathbf{x} = \mathbf{y}$ requires $\beta \rightarrow \infty$ which worsen the condition number of the problem and, thus slow down convergence;
 - direct optimization is more consistent (the given criterion is minimized) and much faster and stable;;
- **direct optimization** with ℓ_1 regularization (to impose sparsity) is superior to *matching pursuit* (Marsh and Richardson 1987) for imposing the sparsity in the CLEAN (Fomalont 1973; Högbom 1974) and *building-blocks* (Hofmann and Weigelt 1993) methods;
- the most successful algorithms – e.g. BSMEM (Baron and Young 2008) and MiRA (Thiébaud 2008) – use direct optimization;
-  **global optimization** is however required, e.g. attempt by the Markov Chain Imager (MACIM) algorithm (Ireland et al. 2008);

Summary and perspectives

Summary and perspectives

- general **inverse problem** framework suitable to describe most methods;
- optimization
 - difficulties: **non-linearity**, lots of variables (as many as pixels), **constraints** (non-negativity), etc.
 - **direct optimization** of the criterion is more consistent and probably more efficient
 - **global optimization** is required
- a priori constraints:
 - regularization: **TV** and **compactness** appear to be the most effective ($\ell_2 - \ell_1$ probably a better compromise for astronomical images)
- the future: **multi-spectral data**
 - spectral regularization (Soulez et al. 2008)
 - much more parameters to fit, computational cost will be a big issue



(le Bouquin et al. 2009)

- other links: **medical tomography**, **compressive sensing**, etc.

- Baron, F. and Young, J. S.: 2008, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 7013 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, p. 144
- Boyd, S., Parikh, N., Chu, E., Peleato, B., and Eckstein, J.: 2010, *Foundations and Trends in Machine Learning* **3**, 1
- Candes, E. J., Romberg, J., and Tao, T.: 2006, *IEEE Transactions on Information Theory* **52(2)**, 489
- Charbonnier, P., Blanc-Féraud, L., Aubert, G., and Barlaud, M.: 1997, *IEEE Trans. Image Process.* **6(2)**, 298
- Cornwell, T. J. and Wilkinson, P. N.: 1981, *Monthly Notices of the RAS* **196**, 1067
- Fomalont, E.: 1973, *Proc. IEEE: Special issue on radio and radar astronomy* **61(9)**, 1211
- Hofmann, K.-H. and Weigelt, G.: 1993, *Astron. & Astrophys.* **278(1)**, 328
- Högbom, J. A.: 1974, *Astron. & Astrophys. Suppl.* **15**, 417
- Ireland, M., Monnier, J., and Thureau, N.: 2008, in J. D. Monnier, M. Schöller, and W. C. Danchi (eds.), *Advances in Stellar Interferometry*, Vol. 6268, pp 62681T1–62681T8, SPIE
- le Besnerais, G., Lacour, S., Mugnier, L. M., Thiébaud, E., Perrin, G., and Meimon, S.: 2008, *IEEE Journal of Selected Topics in Signal Processing* **2(5)**, 767
- le Bouquin, J.-B., Lacour, S., Renard, S., Thiébaud, E., and Merand, A.: 2009, *Astron. & Astrophys.* **496**, L1
- Marsh, K. A. and Richardson, J. M.: 1987, *Astron. & Astrophys.* **182**, 174
- Meimon, S., Mugnier, L. M., and le Besnerais, G.: 2005a, *J. Opt. Soc. Am. A* **22**, 2348
- Meimon, S., Mugnier, L. M., and le Besnerais, G.: 2005b, *Optics Letters* **30(14)**, 1809
- Narayan, R. and Nityananda, R.: 1986, *ARA&A* **24**, 127
- Readhead, A. and Wilkinson, P.: 1978, *Astrophys. J.* **223(1)**, 25
- Renard, S., Thiébaud, E., and Malbet, F.: 2011, *Astron. & Astrophys.* p. in press
- Schwab, F.: 1980, in *Proc. SPIE*, Vol. 231, pp 18–25
- Soulez, F., Thiébaud, E., Gressard, A., Dauphin, R., and Bongard, S.: 2008, in *16th European Signal Processing Conference*, EUSIPCO, Lausanne, Suisse
- Strong, D. and Chan, T.: 2003, *Inverse Problems* **19**, S165–S187
- Thiébaud, E.: 2008, in F. D. Markus Schöller, William C. Danchi (ed.), *Astronomical Telescopes and Instrumentation*, Vol. 7013, pp 70131I–1–70131I–12, SPIE
- Thiébaud, E.: 2009, *New Astronomy Reviews* **53**, 312, Our reference: ASTREV1414 PII: S1387-6473(10)00012-6
- Thiébaud, E. and Giovannelli, J.-F.: 2010, *IEEE Signal Processing Magazine* **27(1)**, 97