Improving the performances of current optical interferometers & future designs
Proceedings of Haute Provence Observatory Colloquium (23-27 September 2013)
Edited by L. Arnold, H. Le Coroller & J. Surdej

An image reconstruction framework for polychromatic interferometry

F. Soulez, É. Thiébaut

Centre de Recherche Astrophysique de Lyon, 9 avenue Charles André, 69561 Saint Genis Laval cedex, France; Université de Lyon; Université Lyon 1; CNRS, UMR 5574; École Normale Supérieure de Lyon

Abstract. We describe a new approach to implement multi-wavelength image reconstruction in the case where the observed scene is a collection of point-like sources. We show the gain in image quality (both spatially and spectrally) achieved by globally taking into account all the data instead of dealing with independent spectral slices. This is achieved thanks to a regularization which favors spatial sparsity and spectral grouping of the sources. Since the objective function is not differentiable, we had to develop a specialized optimization algorithm.

1. Introduction

Optical interferometers provide multiple wavelength measurements. In order to fully exploit the spectral and spatial resolution of these instruments, new algorithms for image reconstruction have to be developed. Early attempts to deal with multi-chromatic interferometric data have consisted in recovering a gray image of the object or independent monochromatic images in some spectral bandwidths. The main challenge is now to recover the full 3-D (spatio-spectral) brightness distribution of the astronomical target given all the available data.

2. Direct model

The complex visibility measured at the wavelength ℓ and for the baseline b can be discomposed in a real $m_{b,\ell,0}$ and a imaginary part $m_{b,\ell,1}$ that can be modeled as:

$$m_{b,\ell,c} \approx \sum_n H_{b,c,n,\ell}\, x_{n,\ell}\,. \tag{1}$$

where $x_{n,\ell}$ is the value of the pixel n in the spectral channel ℓ. The operator $\mathbf{H}$ is separable along the spectral dimension:

$$H_{b,0,n,\ell} = +\cos(\boldsymbol{\theta}_n^\top \cdot \boldsymbol{B}_b/\lambda_\ell)\,, \tag{2}$$

$$H_{b,1,n,\ell} = -\sin(\boldsymbol{\theta}_n^\top \cdot \boldsymbol{B}_b/\lambda_\ell)\,, \tag{3}$$

where $\boldsymbol{B}_b$ is the bth baseline, and $\boldsymbol{\theta}_n$ the position of the n^{th} pixel. As the problem size can be hudge, we use a fast approximation of $\mathbf{H}$ based on the non uniform

fast Fourier transform (Keiner et al. 2009). In this approximation $\mathbf{H}$ has the following structure:

$$\mathbf{H} = \mathbf{R} \cdot \mathbf{F} \cdot \mathbf{S} \tag{4}$$

where $\mathbf{F}$ is the discrete Fourier transform computed by the mean of the fast Fourier transform, $\mathbf{R}$ interpolate the gridded spatial frequencies computed by the DFT to the observed spatial frequencies. $\mathbf{S}$ is an operator that precompensates the convolution by the interpolation kernel used in $\mathbf{R}$.

3. Inverse problem framework

To solve this problem of hyperspectral image reconstruction, we follow an *inverse problem* approach where the image is estimated according to the measurements, its noise and some priors about the observed object. It can be written as a constrained optimization problem:

$$\boldsymbol{x}^+ = \underset{\boldsymbol{x}\in\mathbb{X}}{\arg\min}\, f_{\mathsf{prior}}(\boldsymbol{x}) \quad \text{s.c.} \begin{cases} f_{\mathsf{data}}(\boldsymbol{x}) \leq \eta_1 \\ \mathbf{P}\cdot\boldsymbol{x} = \mathbf{1} \end{cases} \tag{5}$$

where

- $\boldsymbol{x} \in \mathbb{X}$ is the vector of parameters in the image domain. It is angularly sampled by $\boldsymbol{\theta}_n \in \mathbb{A}$ and spectraly sampled in $\lambda_\ell \in \mathbb{L}$
- $\mathbb{X}$ is the subspace of $\mathbb{R}^{\mathrm{Card}(\boldsymbol{m})+}$ where lie the positive object parameters $\boldsymbol{x}$,
- $\boldsymbol{m} \in \mathbb{R}^{\mathrm{Card}(\boldsymbol{m})}$ is the vector of complex visibility,
- $f_{\mathsf{data}}(\boldsymbol{x})$ is the likelihood term ensuring the agreement between parameters $\boldsymbol{x}$ and measurements $\boldsymbol{m}$.
- $\mathbf{P}\cdot\boldsymbol{x} = \mathbf{1}$ is the spatial normalization constraint with the spatial integration operator $\mathbf{P}$,
- $f_{\mathsf{prior}}(\boldsymbol{x})$ is a regularization function that enforces some priors.

3.1 Likelihood

The likelihood or data fidelity function $f_{\mathsf{data}}(\boldsymbol{x})$ is built according to the direct model and the noise statistics. If the noise is Gaussian, this function is a weighted least square function:

$$f_{\mathsf{data}}(\boldsymbol{x}) = (\mathbf{H}\cdot\boldsymbol{x} - \boldsymbol{y})^\top\, \mathbf{C}^{-1}\, (\mathbf{H}\cdot\boldsymbol{x} - \boldsymbol{y})\ , \tag{6}$$

where $\mathbf{C}$ is the noise covariance matrix. Complex visibilities are mutually independant and only real and complex parts of the same visibility are correlated. Thus, we can write:

$$f_{\mathsf{data}}(\boldsymbol{x}) = \sum_{b,\ell,n} (\mathbf{H}_{b,n,\ell}\cdot\boldsymbol{x}_{n,\ell} - \boldsymbol{m}_{b,\ell})^\top\, \mathbf{C}_{b,\ell}^{-1}\, (\mathbf{H}_{b,n,\ell}\cdot\boldsymbol{x}_{n,\ell} - \boldsymbol{m}_{b,\ell})\ , \tag{7}$$

where $\mathbf{C}_{b,\ell}$ is the 2×2 covariance matrix associated with the measured complex visibility $\boldsymbol{m}_{b,\ell}$.

3.2 Regularization

In interferometry, the normalization occurring in the visibility estimation process naturally leads to a strict normalization prior. The parameters space $\mathbb{X}$ is then the subspace where:

$$\boldsymbol{x} \in \mathbb{X} \equiv \left\{ \begin{array}{ll} \boldsymbol{x} & \geq 0 \\ \sum_n x_{n,\ell} & = 1\,. \end{array} \right. \tag{8}$$

More specifically, the main scientific goal of the VLTI GRAVITY instrument is to follow stars in the vicinity of the galactic center. At such a distance, those stars are not resolved by the VLTI baselines and there are only few point-like sources in the instrument field of view. Hence non-negativity and spatial sparsity seem to be adapted priors as this will favor having as few as possible bright point-like sources to explain data. In addition, imposing some spectral continuity while favoring spatial sparsity, structured sparsity is an very well adapted prior for such objects. As we have already shown in integral field spectrography (Soulez et al. 2011) or in optical interferometry (Thiébaut et al. 2013), the structured sparsity prior can be enforced by the mean of mixed norms (Fornasier & Rauhut 2008, Kowalski 2009). In hyperspectral case, it writes:

$$f_{\mathsf{prior}}(\boldsymbol{x}) = \sum\nolimits_n \left(\sum\nolimits_\ell x_{n,\ell}^2 \right)^{1/2} . \tag{9}$$

with n the spatial index (pixel) and ℓ the spectral channel. The fact that such a regularization favors spatial sparsity and spectral grouping is a consequence of the triangular inequality (Fornasier & Rauhut 2008).

4. Algorithm

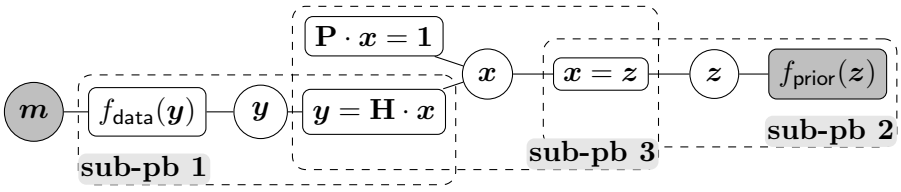

Figure 1.: *Structure of the MiRA3D algorithm*

We propose to solve the image reconstruction problem using an *Alternating Direction of Multipliers Method* (ADMM) (Boyd et al. 2010). Following ADMM strategy, we introduce auxiliary variables $\boldsymbol{y}$ and $\boldsymbol{z}$ to split the complexity of the original problem Eq. (5) which we recast as:

$$\min_{\boldsymbol{x}\geq 0, \boldsymbol{y}, \boldsymbol{z}} \mu\, f_{\mathsf{prior}}(\boldsymbol{z}) + f_{\mathsf{data}}(\boldsymbol{y}) \text{s.t.} \left\{ \begin{array}{ll} \mathbf{H} \cdot \boldsymbol{x} & = \boldsymbol{y}\,, \\ \boldsymbol{x} & = \boldsymbol{z}\,, \\ \mathbf{P} \cdot \boldsymbol{x} & = \mathbf{1}\,, \\ \boldsymbol{x} & \geq 0\,. \end{array} \right. \tag{10}$$

Equality constraints can be enforced by means of the augmented Lagrangian which, for Eq. (10), writes:

$$\begin{aligned}\mathcal{L}_\rho(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) = f_{\mathsf{data}}(\boldsymbol{y}) + \mu\, f_{\mathsf{prior}}(\boldsymbol{z}) \\ + \boldsymbol{u}^\top\cdot(\mathbf{H}\cdot\boldsymbol{x}-\boldsymbol{y}) + \frac{\rho_1}{2}\left\|\mathbf{H}\cdot\boldsymbol{x}-\boldsymbol{y}\right\|_2^2 \\ + \boldsymbol{v}^\top\cdot(\boldsymbol{x}-\boldsymbol{z}) + \frac{\rho_2}{2}\left\|\boldsymbol{x}-\boldsymbol{z}\right\|_2^2, \\ + \boldsymbol{w}^\top\cdot(\mathbf{P}\cdot\boldsymbol{x}-\mathbf{1}) + \frac{\rho_3}{2}\left\|\mathbf{P}\cdot\boldsymbol{x}-\mathbf{1}\right\|_2^2, \end{aligned} \tag{11}$$

where $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$ are Lagrangian parameters associated with the constraints $\boldsymbol{y} = \mathbf{H}\cdot\boldsymbol{x}$, $\boldsymbol{x} = \boldsymbol{z}$ and $\mathbf{P}\cdot\boldsymbol{x} = \mathbf{1}$ respectively. $\rho_1 > 0$, $\rho_2 > 0$ and $\rho_3 > 0$ are quadratic weights.

The ADMM consists on minimizing the augmented lagrangian $\mathcal{L}_\rho(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{u},\boldsymbol{v})$ in an alternating manner with respect to each variable $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ and then updating Lagrangian multipliers $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$:

Algorithm: *Solving the problem (10) using ADMM*

Initialization of the object parameters $\boldsymbol{x}^{(0)}$, Lagrangian parameters $\boldsymbol{u}^{(0)}$, $\boldsymbol{v}^{(0)}$ and $\boldsymbol{w}^{(0)}$ and weight ρ_1, ρ_2 and ρ_3. Then set $t = 1$ and repeat until convergence:

Sub-pb 1: updating $\boldsymbol{y}$:

$$\begin{aligned}\boldsymbol{y}^{(t)} &= \arg\min_{\boldsymbol{y}} \mathcal{L}\big(\boldsymbol{x}^{(t-1)},\boldsymbol{y},\boldsymbol{z}^{(t-1)},\boldsymbol{u}^{(t-1)},\boldsymbol{v}^{(t-1)},\boldsymbol{w}^{(t-1)}\big) \\ &= \arg\min_{\boldsymbol{y}}\Big\{ f_{\mathsf{data}}(\boldsymbol{y}) + \frac{\rho_1}{2}\left\|\boldsymbol{y}-\widetilde{\boldsymbol{y}}^{(t)}\right\|_2^2\Big\}\end{aligned} \tag{12}$$

with:

$$\widetilde{\boldsymbol{y}}^{(t)} = \mathbf{H}\cdot\boldsymbol{x}^{(t-1)} + \boldsymbol{u}^{(t-1)}/\rho_1\,; \tag{13}$$

Sub-pb 2: updating $\boldsymbol{z}$:

$$\begin{aligned}\boldsymbol{z}^{(t)} &= \arg\min_{\boldsymbol{z}} \mathcal{L}\big(\boldsymbol{y}^{(t)},\boldsymbol{x}^{(t-1)},\boldsymbol{z},\boldsymbol{u}^{(t-1)},\boldsymbol{v}^{(t-1)},\boldsymbol{w}^{(t-1)}\big) \\ &= \arg\min_{\boldsymbol{z}}\Big\{ \mu f_{\mathsf{prior}}(\boldsymbol{z}) + \frac{\rho_2}{2}\left\|\boldsymbol{z}-\widetilde{\boldsymbol{z}}^{(t)}\right\|_2^2\Big\}\end{aligned} \tag{14}$$

with:

$$\widetilde{\boldsymbol{z}}^{(t)} = \boldsymbol{x}^{(t-1)} + \boldsymbol{v}^{(t-1)}/\rho_2\,; \tag{15}$$

Sub-pb 3: updating $\boldsymbol{x}$:

$$\begin{aligned}\boldsymbol{x}^{(t)} &= \arg\min_{\boldsymbol{x}\geq 0} \mathcal{L}\big(\boldsymbol{x},\boldsymbol{y}^{(t)},\boldsymbol{z}^{(t)},\boldsymbol{u}^{(t-1)},\boldsymbol{v}^{(t-1)},\boldsymbol{w}^{(t-1)}\big) \\ &= \arg\min_{\boldsymbol{x}\geq 0}\Big\{ \frac{\rho_1}{2}\left\|\mathbf{H}\cdot\boldsymbol{x}-\widetilde{\boldsymbol{\gamma}}^{(t)}\right\|_2^2 + \frac{\rho_2}{2}\left\|\boldsymbol{x}-\widetilde{\boldsymbol{x}}^{(t)}\right\|^2 + \frac{\rho_3}{2}\left\|\mathbf{P}\cdot\boldsymbol{x}-\widetilde{\boldsymbol{\beta}}^{(t)}\right\|_2^2\Big\}\end{aligned} \tag{16}$$

avec:

$$\widetilde{\boldsymbol{\gamma}}^{(t)} = \boldsymbol{y}^{(t)} - \boldsymbol{u}^{(t-1)}/\rho_1\,; \tag{17}$$

$$\widetilde{\boldsymbol{x}}^{(t)} = \boldsymbol{z}^{(t)} - \boldsymbol{v}^{(t-1)}/\rho_2\,; \tag{18}$$

$$\widetilde{\boldsymbol{\beta}}^{(t)} = \mathbf{1} - \boldsymbol{w}^{(t-1)}/\rho_3\,; \tag{19}$$

4: updating multipliers $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$:

$$\boldsymbol{u}^{(t)} = \boldsymbol{u}^{(t-1)} + \rho_1 \left(\mathbf{H}\cdot\boldsymbol{x}^{(t)} - \boldsymbol{y}^{(t)}\right)\,, \tag{20}$$

$$\boldsymbol{v}^{(t)} = \boldsymbol{v}^{(t-1)} + \rho_2 \left(\boldsymbol{x}^{(t)} - \boldsymbol{z}^{(t)}\right)\,. \tag{21}$$

$$\boldsymbol{w}^{(t)} = \boldsymbol{w}^{(t-1)} + \rho_3 \left(\boldsymbol{x}^{(t)} - \mathbf{1}\right)\,. \tag{22}$$

This algorithm is illustrated by the schema Fig. 1. It splits the global minimization problem (Eq. (5)) in three successive simpler sub-problems Eq. (12), Eq. (14) and Eq. (16).

4.1 Sub-problem 1:

The sub-problem 1 defined by Eq. (12) is a denoising problem with Gaussian prior which have an analytical closed form solution. As complex visibilites are independant it is separable and consist on solving several 2×2 linear systems:

$$\boldsymbol{y}_{b,\ell}^{(t)} = \left(\mathbf{C}_{b,\ell}^{-1} + \rho_1\mathbf{I}\right)^{-1}\cdot\left(\mathbf{C}_{b,\ell}^{-1}\cdot\boldsymbol{m}_{b,\ell} + \rho_1\widetilde{\boldsymbol{y}}_{b,\ell}^{(t)}\right)\,, \tag{23}$$

4.2 Sub-problem 2

The sub-problem 2 defined by Eq. (14) is a denoising problem under a group sparsity priors (Thiébaut et al. 2013). For each spectrum, it can be solved independantly applying its so called "proximity operator" (Combettes & Pesquet 2011):

$$\operatorname{prox}_{\alpha\, f}(\widetilde{\boldsymbol{z}}) \stackrel{\text{def}}{=} \underset{\boldsymbol{z}\in\mathbb{R}^{\mathbb{N}}}{\arg\min}\left\{\alpha\, f(\boldsymbol{z}) + \frac{1}{2}\,\|\boldsymbol{z}-\widetilde{\boldsymbol{z}}\|_2^2\right\}\,. \tag{24}$$

which is in the case of group sparsity

$$\operatorname{prox}_{\alpha\, f_{\mathsf{joint}}}(\widetilde{\boldsymbol{z}})_{n,\ell} = \begin{cases} \left(1-\dfrac{\alpha}{\widetilde{\beta}_n}\right)\,\widetilde{z}_{n,\ell} & \text{if } \widetilde{\beta}_n > \alpha\,; \\ 0 & \text{else} \end{cases} \tag{25}$$

with $\widetilde{\beta}_n \stackrel{\text{def}}{=} \left(\sum_\ell \widetilde{\boldsymbol{z}}_{n,\ell}^2\right)^{1/2}$ and $\alpha = \mu/\rho_2$.

4.3 Sub problem 3

The sub-problem 3 defined Eq. (16) is not separable and does not have a closed form solution. However it is a classical quadratic problem under positivity constraint. We solve it by the mean of the VMLMB algorithm (Thiébaut 2002) that can account for bounds. From the Eckstein-Bertsekas theorem (Eckstein & Bertsekas 1992), it is not necessary to solve this sub-problem exactly to ensure global convergence as long as error on $\boldsymbol{z}$ can be absolutely summed. In practice less than a dozen of iteration is necessary.

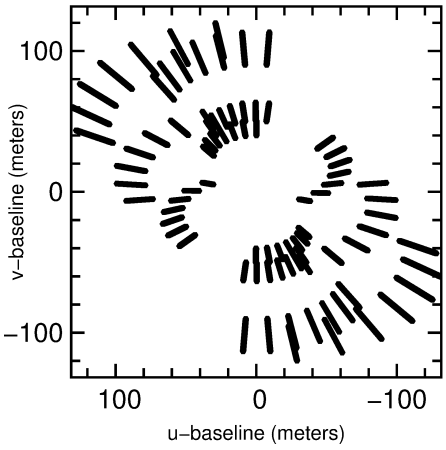

Figure 2.: (u, v) *coverage*

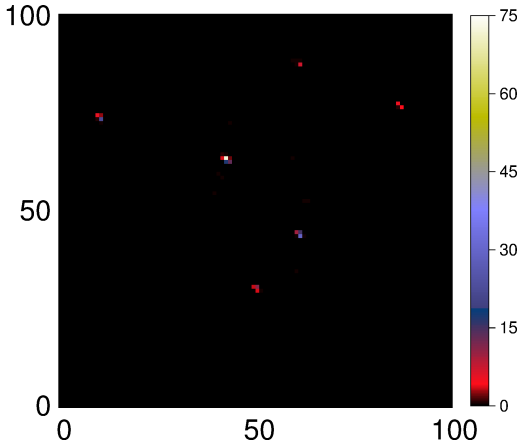

Figure 3.: *Reconstructed image (spectraly integrated)*

5. Results

We have tested our method on GRAVITY simulation. It consists on six unresolved stars with different spectra observed by the 4 UTs with 240 spectral channels from 1.95 μm to 2.45 μm and 42 baselines (about 10080 complex visibilities). (u, v) coverage is presented Fig. 2.

We reconstruct an hyperspectral image with 240 spectral channels and 100×100 pixels of size 1×1 mas. The reconstructed image spectraly integrated in the K band is shown Fig. 3. The six star are recovered and there is not any false detection. The shape of reconstructed star is due to the beam and its centroide indicated its position with an error lower than 0.15 mas. The 6 reconstructed spectra presented in Fig. 4 are very close to the theoretical spectra.

Acknowledgements. The authors thank Thibaut Paumard who kindly provides the GRAVITY simulated data. Our algorithm has been implemented and tested with YORICK (`http://yorick.sourceforge.net/`) which is freely available. This work is supported by the POLCA project (*Percées astrophysiques grâce au traitement de données interférométriques polychromatiques* funded by the French ANR: ANR-10-BLAN-0511).

References

S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3:1-122, 2010. doi: 10.1561/2200000016. URL `http://www.stanford.edu/~boyd/papers/pdf/admm_distr_stats.pdf`.

P. L. Combettes and J.-C. Pesquet. *Proximal splitting methods in signal processing*, chapter Fixed-Point Algorithms for Inverse Problems in Science and Engineering, pages 185-212. Springer, New York, 2011.

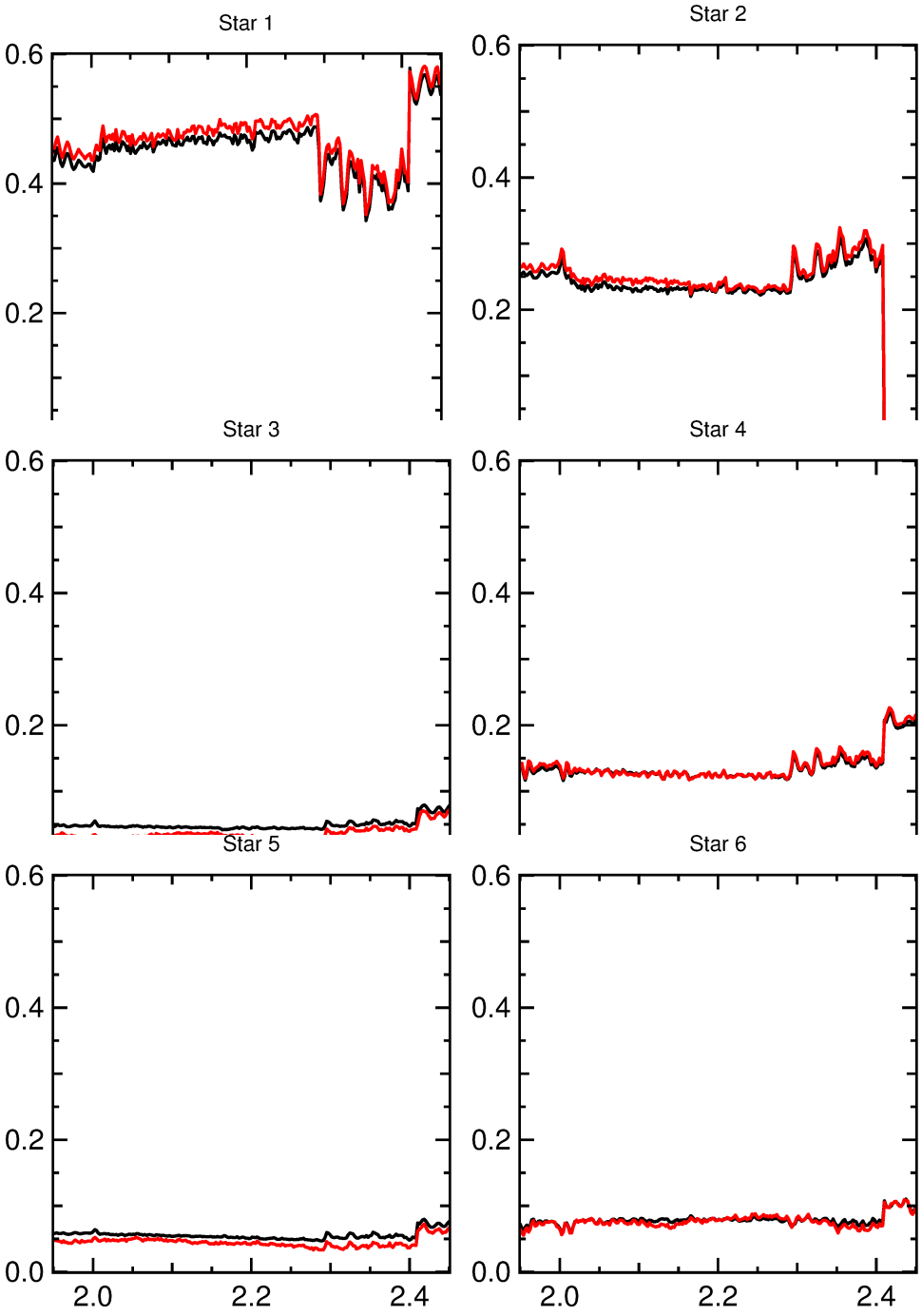

Figure 4.: *Estimated (in black) and theoretical (in red) spectra for the 6 stars.*

J. Eckstein and D. Bertsekas. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming*, 55(1):293-318, 1992.

M. Fornasier and H. Rauhut. Recovery algorithms for vector valued data with joint sparsity constraints. *SIAM Journal on Numerical Analysis*, 46 (2):577-613, 2008.

J. Keiner, S. Kunis, and D. Potts. Using nfft 3-a software library for various nonequispaced fast fourier transforms. *ACM Transactions on Mathematical Software (TOMS)*, 36(4):19, 2009.

M. Kowalski. Sparse regression using mixed norms. *Applied and Computational Harmonic Analysis*, 27(3):303-324, 2009. ISSN 1063-5203. doi: 10.1016/j.acha.2009.05.006. URL `http://www.sciencedirect.com/science/article/pii/S1063520309000608`.

F. Soulez, É. Thiébaut, S. Bongard, and R. Bacon. Restoration of hyperspectral astronomical data from integral field spectrograph. In *3rd Workshop on Hyperspectral Image and Signal Processing (WHISPERS)*, Lisbon, Portugal, 2011. URL `http://core.ieee-whispers.com/`.

É. Thiébaut. Optimization issues in blind deconvolution algorithms. In J.-L. Starck and F. D. Murtagh, editors, *Astronomical Data Analysis II*, volume 4847, pages 174-183, Bellingham, Washington, 2002. SPIE. doi: 10.1117/12.461151.

É. Thiébaut, F. Soulez, and L. Denis. Exploiting spatial sparsity for multi-wavelength imaging in optical interferometry. *Journal of the Optical Society of America A*, 30(2):160-170, Feb. 2013.
URL `http://hal.archives-ouvertes.fr/hal-00771800`.