

Détection de sources en interférométrie optique hyperspectrale

Ferréol SOULEZ, Éric THIÉBAUT

Centre de Recherche Astrophysique de Lyon
9 avenue Charles André, 69561 Saint Genis Laval cedex, France*
ferreol.soulez@univ-lyon1.fr, eric.thiebaut@univ-lyon1.fr

Résumé – En faisant interférer la lumière provenant de plusieurs télescopes, l’interférométrie optique fournit des mesures à très haute résolution angulaire (de l’ordre de la milliseconde d’arc). Chaque mesure estime la valeur en une fréquence spatiale de la transformée de Fourier de la distribution spatiale d’intensité émise par l’objet observé dans chacun des canaux spectraux. Le problème traité ici est la détection, la localisation précise et l’extraction sans biais du spectre de chacune des étoiles d’un amas observé en interférométrie. C’est un verrou important pour l’étude des étoiles au voisinage du trou noir central de notre galaxie, but scientifique du futur instrument GRAVITY[5] du VLTI.

A la suite de nos précédents travaux [6, 7], nous présentons ici une méthode de reconstruction basée sur la méthode de multiplicateur à directions alternées (ADMM) [15]. Cela permet d’utiliser dans le même temps les données interférométriques et photométriques. L’introduction de variables auxiliaires permet de découper le problème de reconstruction en sous problèmes plus faciles à traiter.

Des tests sur des simulations montrent que la méthode proposée permet de détecter toutes les étoiles d’un amas et de d’estimer leurs spectres avec un biais négligeable [7].

Abstract – 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 band- widths. The main challenge is now to recover the full 3-D (spatio-spectral) brightness distribution of the astronomical target given all the available data. 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

L’interférométrie optique permet d’estimer la distribution spatiale d’intensité $I_\lambda(\theta)$ de l’objet d’intérêt astrophysique avec une haute résolution angulaire d’après la mesure de sa transformée de Fourier en un relativement faible nombre de fréquences spatiales. Si de nombreux algorithmes ont été proposés pour la reconstruction d’image dans le cas monochromatique (MiRA[1], BSMEM[2, 3], WISARD[4], *etc.*), aucun n’est spécifiquement adapté aux mesures polychromatiques.

Dans ce travail, nous traitons le cas simple où l’objet observé est une collection de sources ponctuelles de spectres différents et les mesures sont des visibilitées complexes auxquelles est ajouté la photométrie, c’est-à-dire le flux total à chaque longueur d’onde. Nous reconstruisons cet objet en suivant une approche *maximum a posteriori* avec un *a priori* de type parcimonie structurée. En pratique, nous résolvons ce problème avec la méthode de multiplicateurs à directions alternées (ADMM) en introduisant des variables auxiliaires. Cela permet de remplacer un problème complexe par une succession de problèmes plus simples.

2 Modèle direct

La mesure $m_{b,\ell}$ pour la b ème ligne de base et le ℓ ème canal spectral est une valeur complexe pouvant être décomposée en une partie réelle $m_{b,\ell,0}$ et une partie imaginaire $m_{b,\ell,1}$. Le modèle de ces visibilitées complexes s’écrit :

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

où $x_{n,\ell}$ est la valeur de l’objet au pixel n et dans le canal spectral ℓ . L’opérateur \mathbf{H} est séparable spectralement :

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

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

où \mathbf{B}_b est la b ème ligne de base. Toutefois, étant donnée la taille des problèmes à traiter, nous proposons d’utiliser une approximation rapide de \mathbf{H} basée sur la transformée de Fourier rapide non uniforme[8]. Dans ce cas l’opérateur \mathbf{H} a la structure suivante :

$$\mathbf{H} = \mathbf{R} \cdot \mathbf{F} \cdot \mathbf{S} \quad (4)$$

où \mathbf{F} est la transformée de Fourier discrète (DFT) calculée par FFT, \mathbf{R} interpole les fréquences spatiales calculées par la DFT

*Ce travail a été financé par le projet ANR POLCA (ANR-10-BLAN-0511).

aux fréquences spatiales observées et \mathbf{S} est un opérateur de pré-compensation de la convolution par le noyau d'interpolation utilisé dans \mathbf{R} .

3 Méthode

S'inscrivant dans une perspective ‘‘approche inverse’’, nous résolvons ce problème de reconstruction d'image comme un problème d'optimisation sous contraintes :

$$\mathbf{x}^+ = \arg \min_{\mathbf{x} \geq 0} f_{\text{prior}}(\mathbf{x}) \quad \text{s.c.} \begin{cases} f_{\text{data}}(\mathbf{x}|\mathbf{m}) \leq \eta_1 \\ f_{\text{phot}}(\mathbf{x}|\mathbf{s}) \leq \eta_2 \end{cases} \quad (5)$$

où

- $\mathbf{x} \in \mathbb{R}_+^{\text{Card}(\mathbf{x})}$ sont les paramètres dans l'espace image échantillonné angulairement en $\theta_n \in \mathbb{A}$ (la liste des positions angulaires des pixels) et spectralement en $\lambda_\ell \in \mathbb{L}$ (la liste des longueurs d'onde) ;
- $\text{Card}(\mathbf{x})$ est le nombre de ces paramètres ;
- $\mathbf{m} \in \mathbb{R}^{\text{Card}(\mathbf{m})}$ sont les mesures interférométriques ;
- $f_{\text{data}}(\mathbf{x}|\mathbf{m})$ est un terme de vraisemblance assurant l'accord entre le modèle et les données interférométriques ;
- $\mathbf{s} \in \mathbb{R}^{\text{Card}(\mathbf{s})}$ est le flux photométrique mesuré pour chaque canal spectral ;
- $f_{\text{phot}}(\mathbf{x}|\mathbf{s})$ est un terme de vraisemblance assurant l'accord entre le modèle et les flux photométriques mesurés ;
- et $f_{\text{prior}}(\mathbf{x})$ est une fonction de régularisation introduisant les *a priori*.

3.1 Terme de régularisation

Dans ce papier, nous nous souhaitons reconstruire des images hyperspectrales composées d'une collection de sources ponctuelles. Des *a priori* de positivité et de parcimonie spatiale semblent donc tout à fait adaptés à ce type d'objet. Nous proposons ici d'utiliser la régularisation par une norme structurée [11] qui favorise la parcimonie spatiale tout en assurant que les sources soient à la même position spatiale quelque soit le canal spectral [10] :

$$f_{\text{prior}}(\mathbf{x}) = \sum_n \left(\sum_\ell x_{n,\ell}^2 \right)^{1/2}. \quad (6)$$

3.2 Termes d'attache aux données

Dans l'hypothèse d'un bruit de mesure gaussien, la fonction d'attache aux données interférométrique s'écrit :

$$f_{\text{data}}(\mathbf{x}) = (\mathbf{H} \cdot \mathbf{x} - \mathbf{m})^\top \mathbf{C}^{-1} (\mathbf{H} \cdot \mathbf{x} - \mathbf{m}), \quad (7)$$

où \mathbf{C} est la matrice de covariance des mesures interférométriques.

La fonction d'attache aux données photométriques s'écrit

$$f_{\text{phot}}(\mathbf{x}) = \sum_\ell w_\ell \left(\sum_n x_{n,\ell} - s_\ell \right)^2. \quad (8)$$

où w_ℓ est l'inverse de la variance associée avec la photométrie s_ℓ mesurée dans le canal ℓ .

4 Algorithme

Nous proposons de résoudre ce problème en utilisant la méthode ADMM (pour *Alternating Direction of Multipliers Method*) [15]. Dans ce cadre, l'équation (5) peut être réécrite en introduisant des variables auxiliaires \mathbf{z} et \mathbf{y} sous la forme :

$$\mathbf{x}^+ = \arg \min_{\mathbf{x} \geq 0, \mathbf{y}, \mathbf{z}} \mu f_{\text{prior}}(\mathbf{z}) + f_{\text{data}}(\mathbf{y}|\mathbf{m}) + f_{\text{phot}}(\mathbf{x}|\mathbf{s})$$

tels que $\begin{cases} \mathbf{y} = \mathbf{H} \cdot \mathbf{x}, \\ \mathbf{x} = \mathbf{z}. \end{cases} \quad (9)$

Les deux contraintes d'égalité peuvent être imposées au moyen d'un lagrangien augmenté qui dans le cas de Éq. (9) s'écrit :

$$\begin{aligned} \mathcal{L}_\rho(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}, \mathbf{v}) &= f_{\text{data}}(\mathbf{y}) + \mu f_{\text{prior}}(\mathbf{z}) + f_{\text{phot}}(\mathbf{x}) \\ &\quad + \mathbf{u}^\top \cdot (\mathbf{H} \cdot \mathbf{x} - \mathbf{y}) + \frac{\rho_1}{2} \|\mathbf{H} \cdot \mathbf{x} - \mathbf{y}\|_2^2 \\ &\quad + \mathbf{v}^\top \cdot (\mathbf{x} - \mathbf{z}) + \frac{\rho_2}{2} \|\mathbf{x} - \mathbf{z}\|_2^2, \end{aligned} \quad (10)$$

où \mathbf{u} et \mathbf{v} sont les multiplicateurs de Lagrange associés respectivement aux contraintes $\mathbf{y} = \mathbf{H} \cdot \mathbf{x}$ et $\mathbf{x} = \mathbf{z}$. $\rho_1 > 0$ et $\rho_2 > 0$ sont les poids quadratiques.

La méthode ADMM consiste à minimiser alternativement le lagrangien augmenté $\mathcal{L}_\rho(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}, \mathbf{v})$ suivant chacune des variables \mathbf{x} , \mathbf{y} et \mathbf{z} puis à mettre à jour les paramètres de Lagrange \mathbf{u} et \mathbf{v} :

Algorithme 1. Résolution du problème (9) par ADMM.

Initialiser les variables $\mathbf{x}^{(0)}$, les paramètres de lagrange $\mathbf{u}^{(0)}$ et $\mathbf{v}^{(0)}$, ρ_1 et ρ_2 . Puis répéter pour $t = 1, 2, \dots$ et jusqu'à convergence :

Sous-pb 1. : mettre à jour les variables \mathbf{y} :

$$\begin{aligned} \mathbf{y}^{(t)} &= \arg \min_{\mathbf{y}} \mathcal{L}(\mathbf{x}^{(t-1)}, \mathbf{y}, \mathbf{z}^{(t-1)}, \mathbf{u}^{(t-1)}, \mathbf{v}^{(t-1)}) \\ &= \arg \min_{\mathbf{y}} \left\{ f_{\text{data}}(\mathbf{y}) + \frac{\rho_1}{2} \left\| \mathbf{y} - \tilde{\mathbf{y}}^{(t)} \right\|_2^2 \right\} \end{aligned} \quad (11)$$

avec :

$$\tilde{\mathbf{y}}^{(t)} = \mathbf{H} \cdot \mathbf{x}^{(t-1)} + \mathbf{u}^{(t-1)} / \rho_1; \quad (12)$$

Sous-pb 2. : mettre à jour les variables \mathbf{z} :

$$\begin{aligned} \mathbf{z}^{(t)} &= \arg \min_{\mathbf{z}} \mathcal{L}(\mathbf{y}^{(t)}, \mathbf{x}^{(t-1)}, \mathbf{z}, \mathbf{u}^{(t-1)}, \mathbf{v}^{(t-1)}) \\ &= \arg \min_{\mathbf{z}} \left\{ \mu f_{\text{prior}}(\mathbf{z}) + \frac{\rho_2}{2} \left\| \mathbf{z} - \tilde{\mathbf{z}}^{(t)} \right\|_2^2 \right\} \end{aligned} \quad (13)$$

with :

$$\tilde{\mathbf{z}}^{(t)} = \mathbf{x}^{(t)} + \mathbf{v}^{(t-1)} / \rho_2; \quad (14)$$

Sous-pb 3. : mettre à jour les variables \mathbf{x} :

$$\begin{aligned} \mathbf{x}^{(t)} &= \arg \min_{\mathbf{x} \geq 0} \mathcal{L}(\mathbf{x}, \mathbf{y}^{(t)}, \mathbf{z}^{(t)}, \mathbf{u}^{(t-1)}, \mathbf{v}^{(t-1)}) \\ &= \arg \min_{\mathbf{x} \geq 0} \left\{ f_{\text{phot}}(\mathbf{x}) + \frac{\rho_1}{2} \left\| \mathbf{H} \cdot \mathbf{x} - \tilde{\mathbf{y}}^{(t)} \right\|_2^2 \right. \\ &\quad \left. + \frac{\rho_2}{2} \left\| \mathbf{x} - \tilde{\mathbf{x}}^{(t)} \right\|_2^2 \right\} \end{aligned} \quad (15)$$

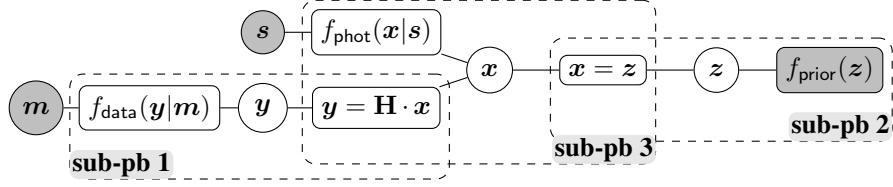


Fig. 1 Schéma illustrant le double découpage proposé

avec :

$$\tilde{\gamma}^{(t)} = \mathbf{y}^{(t)} - \mathbf{u}^{(t-1)} / \rho_1; \quad (16)$$

$$\tilde{\mathbf{x}}^{(t)} = \mathbf{z}^{(t)} - \mathbf{v}^{(t-1)} / \rho_2; \quad (17)$$

4. Mise à jour des multiplicateurs u et v :

$$\mathbf{u}^{(t)} = \mathbf{u}^{(t-1)} + \rho_1 \left(\mathbf{H} \cdot \mathbf{x}^{(t)} - \mathbf{y}^{(t)} \right), \quad (18)$$

$$\mathbf{v}^{(t)} = \mathbf{v}^{(t-1)} + \rho_2 \left(\mathbf{x}^{(t)} - \mathbf{z}^{(t)} \right). \quad (19)$$

Comme il est illustré sur le schéma Fig. 1, cet algorithme revient donc à résoudre successivement trois sous problèmes. Dans les prochains paragraphes, nous montrons que chacun de ces sous problèmes est plus simple à résoudre que le problème global donné par l'Éq. (5).

4.1 Sous problème 1 :

Le sous problème 1, défini par l'équation (11) est un problème de débruitage avec *a priori* gaussien. Les mesures complexes interférométriques $\mathbf{m}_{b,\ell}$ sont indépendantes et seules les parties réelle et imaginaire d'une même mesure sont corrélées. Ce problème de débruitage est donc séparable en petits problèmes à deux dimensions qu'il est possible de résoudre analytiquement :

$$\mathbf{y}_{b,\ell}^{(t)} = \left(\mathbf{C}_{b,\ell}^{-1} + \rho_1 \mathbf{I} \right)^{-1} \cdot \left(\mathbf{C}_{b,\ell}^{-1} \cdot \mathbf{m}_{b,\ell} + \rho_1 \tilde{\mathbf{y}}_{b,\ell}^{(t)} \right), \quad (20)$$

où $\mathbf{C}_{b,\ell}$ est la matrice de covariance 2×2 associée à la mesure $\mathbf{m}_{b,\ell}$. La solution analytique de ce sous problème peut être calculée très rapidement en résolvant en parallèle pour chaque mesure $\mathbf{m}_{b,\ell}$ un système linéaire à deux inconnues.

4.2 Sous problème 2

Le sous problème 2, défini par l'équation (13) est un problème de débruitage avec *a priori* de parcimonie groupée. Ce problème est séparable par pixel et pour chacun des spectres, tous les calculs peuvent être réalisés indépendamment et en parallèle. Pour chaque spectre $\mathbf{z}_{n,\ell}$, résoudre le problème (13) revient à appliquer l'opérateur proximal [12] :

$$\text{prox}_{\alpha f}(\tilde{\mathbf{z}}) \stackrel{\text{def}}{=} \arg \min_{\mathbf{z} \in \mathbb{R}^N} \left\{ \alpha f(\mathbf{z}) + \frac{1}{2} \|\mathbf{z} - \tilde{\mathbf{z}}\|_2^2 \right\}. \quad (21)$$

qui dans le cas de la parcimonie groupée s'écrit [7] :

$$\text{prox}_{\alpha f_{\text{joint}}}(\tilde{\mathbf{z}})_{n,\ell} = \begin{cases} \left(1 - \frac{\alpha}{\tilde{\beta}_n} \right) \tilde{\mathbf{z}}_{n,\ell} & \text{si } \tilde{\beta}_n > \alpha; \\ 0 & \text{sinon} \end{cases} \quad (22)$$

avec $\beta_n \stackrel{\text{def}}{=} \left(\sum_{\ell} \tilde{\mathbf{z}}_{n,\ell}^2 \right)^{1/2}$ et $\alpha = \mu / \rho_2$.

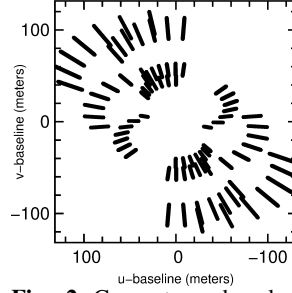


Fig. 2 Couverture du plan (u, v)

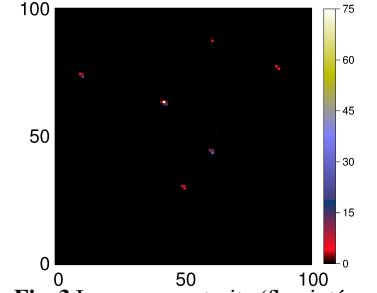


Fig. 3 Image reconstruite (flux intégré)

4.3 Sous problème 3

À la différence des sous problèmes 1 et 2, le sous problème défini par l'équation (15) n'est pas séparable et n'a pas de solution analytique. Toutefois, on reconnaît un problème classique de programmation quadratique sous contrainte de positivité. Nous le résolvons de manière itérative au moyen de l'algorithme d'optimisation continue VMLMB [14] qui permet de prendre en compte des contraintes de positivité. D'après le théorème d'Eckstein-Bertsekas [13], il n'est pas nécessaire de résoudre exactement ce problème pour assurer la convergence globale à condition que les approximations sur \mathbf{z} soient sommables. En pratique, une dizaine d'itérations suffisent.

5 Résultats

Nous avons testé notre méthode sur une simulations réalisées pour l'instrument GRAVITY du VLTI (avec les 4 télescopes principaux (UT)) comprenant 6 corps noirs de température entre $2000K$ et $15000K$. Les mesures sont composées de 240 canaux spectraux de $1.95 \mu\text{m}$ à $2.45 \mu\text{m}$ pour 42 ligne de base, soit 10080 visibilité complexes. La couverture du plan de Fourier (u, v) est représentée figure 2.

Nous avons choisi de reconstruire une image avec autant de canaux spectraux que les mesures (240) avec des pixels de taille $1 \times 1 \text{ mas}^2$ (mas = milliseconde d'arc), soit un champ reconstruit de 100×100 pixels. L'image reconstruite flux intégré est présenté Fig. 3. Les six étoiles sont reconstruites et il n'y a pas de fausse détection. La forme allongée des tâches reconstruites est due à la bien meilleure couverture du plan (u, v) dans une direction. La position de chacune des étoiles est estimée d'après le barycentre de ces tâches. L'erreur de positionnement indiquée en haut des spectres Fig. 4 est inférieure à 0.15 mas ($1/6^{\text{ème}}$ de pixel) et même dix fois moins pour les deux étoiles les plus brillantes. Les spectres reconstruits présentés Fig. 4

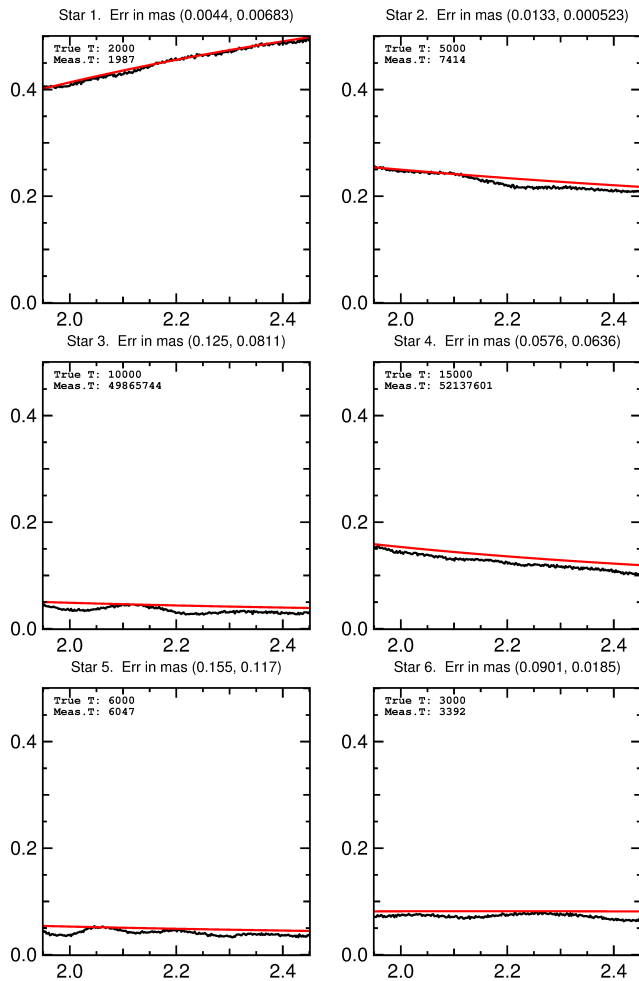


Fig. 4 Spectres reconstruits (en noir) et théoriques (en rouge) pour les six étoiles. Pour chaque spectre sont indiqués : la température ajustée sur le spectre et l’erreur de positionnement en milliseconde d’arc.

sont très proches des spectres théoriques.

Ces résultats sur données simulées réalistes¹ valident l’approche parcimonieuse par moindre norme L1 pour ce type d’observations. Afin de passer au traitement de mesures interférométriques réelles, nous sommes en train d’adapter le modèle (ici linéaire) aux mesures de visibilité différentielles. Lorsque le nombre de sources attendues est faible, nous envisageons de comparer les performances de l’approche convexe L1 à une méthode “gloutonne”.

Références

[1] É. Thiébaud, “MiRA : an effective imaging algorithm for optical interferometry,” in *SPIE : Astronomical Telescopes and Instrumentation*, 2008, vol. 7013, pp. 70131I–1. 1

[2] D. F. Buscher, “Direct maximum-entropy image reconstruction from the bispectrum,” in *IAU Symp. 158 : Very High Angular Resolution Imaging*, J. G. Robertson and W. J. Tango, eds. (1994), pp. 91–+. 1

[3] F. Baron and J. S. Young, “Image reconstruction at Cambridge University,” in *SPIE Conference Series*, (2008), vol. 7013, p. 144. 1

[4] S. Meimon, L. M. Mugnier, and G. le Besnerais, “Reconstruction method for weak-phase optical interferometry,” *Optics Letters*, **30**, 1809–1811 (2005). 1

[5] S. Gillessen, F. Eisenhauer, G. Perrin, W. Brandner, C. Straubmeier, K. Perraut, A. Amorim, M. Schöller, C. Araujo-Hauck, and H. Bartko, “Gravity : a four-telescope beam combiner instrument for the vlti,” in *SPIE*, 2010. 1

[6] É. Thiébaud and F. Soulez, “Multi-wavelength imaging algorithm for optical interferometry,” in *SPIE Astronomical Telescopes+ Instrumentation*, 2012, pp. 84451C–84451C. 1

[7] É. Thiébaud, F. Soulez, and L. Denis, “Exploiting spatial sparsity for multi-wavelength imaging in optical interferometry,” *Journal of the Optical Society of America A*, vol. to appear, 2013. 1, 3

[8] J. Keiner, S. Kunis and D. Potts “Using NFFT 3 - a software library for various nonequispaced fast Fourier transforms” *ACM Trans. Math. Software*, 36, Article 19, 1-30, 2009. 1

[9] D. Donoho, “For most large underdetermined systems of linear equations, the minimal ell-1 norm near-solution approximates the sparsest near-solution”, *Communications on Pure and Applied Mathematics* **59**, 907–934 (2006).

[10] F. Soulez, É. Thiébaud, S. Bongard, and R. Bacon, “Restoration of hyperspectral astronomical data from integral field spectrograph,” in *3rd IEEE WHISPERS*, (Lisbon, Portugal, 2011). 2

[11] Massimo Fornasier and Holger Rauhut, “Recovery algorithms for vector valued data with joint sparsity constraints,” *SIAM Journal on Numerical Analysis*, vol. 46, no. 2, pp. 577–613, 2008. 2

[12] P. L. Combettes and J.-C. Pesquet, *Proximal splitting methods in signal processing* (Springer, New York, 2011), chap. Fixed-Point Algorithms for Inverse Problems in Science and Engineering, pp. 185–212. 3

[13] J. Eckstein and D. Bertsekas, “On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators,” *Mathematical Programming* **55**, 293–318 (1992). 3

[14] É. Thiébaud, “Optimization issues in blind deconvolution algorithms,” in *Astronomical Data Analysis II.*, Jean-Luc Starck, Ed., dec 2002, vol. 4847, pp. 174–183. 3

[15] S. Boyd *et al.*, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Foundations and Trends in Machine Learning*, vol. 3, pp. 1–122, 2010. 1, 2

1. Nous remercions Thibaut Paumard pour cette simulation.