

Séparation myope de sources stellaires à partir d'images hyperspectrales par des méthodes bayésiennes : filtre particulière et MCMC

Ahmed SELLOUM^{1,2}, Emma VILLENEUVE^{1,2}, Hervé CARFANTAN^{1,2}, Yannick DEVILLE^{1,2}

¹Université de Toulouse; UPS-OMP; IRAP; Toulouse, France

²CNRS; IRAP; 14, avenue Edouard Belin, F-31400 Toulouse, France

{ahmed.selloum, emma.villeneuve, herve.carfantan, yannick.deville}@irap.omp.eu

Résumé – Nous nous intéressons au problème de séparation de spectres d'étoiles à partir de données hyperspectrales d'un champ stellaire dense obtenues par le futur spectro-imageur MUSE. La structure du mélange est connue (liée à la réponse impulsionnelle de l'imageur) et l'objectif est d'estimer les paramètres de cette structure du mélange ainsi que les positions précises des étoiles et leurs spectres. Nous proposons d'exploiter deux méthodes bayésiennes pour résoudre ce problème d'estimation non-linéaire : un filtre particulière et une méthode de Monte-Carlo par Chaîne de Markov. Les deux méthodes fournissent de très bons résultats d'estimation sur les paramètres non linéaires du modèle (position et mélange) malgré la multi-modalité de la loi a posteriori.

Abstract – We focus on the problem of stellar spectra separation from hyperspectral data for a dense stellar field observed with the future MUSE spectro-imager. It corresponds to a specific blind source separation problem, where the structure of the mixture is known (linked to the impulse response of the imager) and non-linearly depends on a reduced number of parameters. Our objective is to estimate simultaneously these structure parameters as well as the positions and spectra of the stars. We propose to exploit two Bayesian methods to solve this non-linear estimation problem: a particular filter and a Monte Carlo Markov Chain (MCMC). Both methods provide very good results for estimating the non-linear model parameters (position and mixture) despite the multi-modality of the posterior law.

1 Introduction

La séparation aveugle de sources (SAS) consiste à estimer un ensemble de sources à partir d'un ensemble de signaux mélangés observés, l'opérateur de mélange étant inconnu. On introduit généralement des informations a priori sur la nature de mélange et/ou sur les sources; on parle alors de séparation myope de sources. Ces méthodes ont été exploitées dans de nombreuses disciplines telles le traitement de signaux audio, la télédétection spatiale et l'astrophysique. En astrophysique, on peut classer les méthodes de SAS les plus utilisées en trois grandes familles : 1) les méthodes basées sur l'Analyse en Composantes Indépendantes (ACI) où l'on utilise l'hypothèse de l'indépendance statistique des sources [1]; 2) les méthodes basées sur la Factorisation en Matrices Non-négatives (FMN) où l'on exploite la positivité des sources et des coefficients de mélange [2]; 3) les méthodes bayésiennes où l'on cherche à estimer la loi a posteriori des sources connaissant le mélange observé et des lois a priori sur la nature de ce mélange et sur les sources. On peut distinguer deux types de méthodes bayésiennes : les méthodes de Monte-Carlo par Chaîne de Markov (MCMC) [3] et les méthodes de filtrage particulière (FP) [4].

Notre travail¹ s'inscrit dans le cadre du projet européen MUSE (Multi Unit Spectroscopic Explorer) [5]. L'imageur hyperspectral MUSE fournira des cubes de données constitués de 300×300 pixels et jusqu'à 4000 longueurs d'onde du spectre. L'objectif de notre travail est de développer une méthode de séparation myope de spectres d'étoiles à partir de l'observation d'un champ stellaire dense. Les spectres des étoiles peuvent être très corrélés entre eux ce qui est incompatible avec les méthodes de type ACI. D'autre part, les performances des méthodes basées sur la FMN dépendent de l'initialisation et ne permettent pas de contraindre la matrice de mélange à prendre une forme particulière. Aussi, nous nous sommes dirigés vers les méthodes bayésiennes de type FP et MCMC qui permettent aisément de prendre en compte la forme a priori du mélange. Après une présentation du problème et des informations a priori, nous détaillons les deux méthodes bayésiennes proposées. Enfin, nous comparons les résultats des deux algorithmes sur des données simulées d'un mélange réaliste de spectres d'étoiles.

1. Ce travail a été partiellement financé par le projet ANR 08-BLAN-0253-01 DAHLIA - Dedicated Algorithms for HyperspectraL Imaging in Astronomy.

2 Problématique

2.1 Modèle de mélange

Nous considérons un champ stellaire dense, constitué de N_s étoiles, supposées ponctuelles, situées aux positions spatiales $\{z_i\}_{i \in [1; N_s]}$. La convolution du champ observé par la PSF (*Point Spread Function*), qui modélise la réponse impulsionnelle de l'ensemble de la chaîne d'acquisition (atmosphère, télescope, système d'optique adaptative et instrument), produit un mélange des spectres des étoiles dans les données. En première approximation, les effets spatiaux et spectraux de la PSF peuvent être séparés [6]. Nous ne considérons ici que l'étalement spatial de la PSF et essayons d'estimer $\{s_{i,\lambda}\}_{i \in [1; N_s], \lambda \in [1; N_\lambda]}$, correspondant aux spectres stellaires convolués par la PSF spectrale. Par conséquent le modèle de mélange peut s'écrire :

$$y_\lambda[p] = \sum_{i=1}^{N_s} \{\mathbf{M}_\lambda\}_{p,i} s_{i,\lambda} + \epsilon_\lambda[p] \Leftrightarrow \mathbf{Y}_\lambda = \mathbf{M}_\lambda \mathbf{S}_\lambda + \epsilon_\lambda \quad (1)$$

où $\mathbf{Y}_\lambda = [y_\lambda[1], \dots, y_\lambda[N_p]]^T$ est le vecteur de mélange, N_p le nombre de pixels, $\mathbf{S}_\lambda = [s_{1,\lambda}, \dots, s_{N_s,\lambda}]^T$ le vecteur source et $\epsilon_\lambda = [\epsilon_\lambda[1], \dots, \epsilon_\lambda[N_p]]^T$ le vecteur bruit. L'élément (p, i) de la matrice \mathbf{M}_λ correspond à la valeur, au pixel p , de la PSF centrée en la position z_i . Cette PSF peut être modélisée par une fonction de Moffat caractérisée par les paramètres α_λ et β_λ qui varient spectralement [7], ce qui donne :

$$\{\mathbf{M}_\lambda\}_{p,i} = \frac{\beta_\lambda - 1}{\pi \alpha_\lambda^2} \left[1 + \frac{\|p - z_i\|^2}{\alpha_\lambda^2} \right]^{-\beta_\lambda} \quad (2)$$

Le bruit d'observation peut être modélisé comme un bruit additif gaussien, centré et indépendant spatialement et spectralement. Sa variance est variable spatialement et spectralement [7], mais on dispose d'une estimation de cette variance, notée $\hat{\sigma}_\lambda^2[p]$ au pixel p et à la longueur d'onde λ .

2.2 Estimation non bayésienne

Notre objectif est de séparer les spectres \mathbf{S}_λ à partir des observations \mathbf{Y}_λ . Notons que l'Eq. (1) sous forme matricielle est typique des problèmes de SAS, mais on dispose ici d'une contrainte forte sur la structure de la matrice de mélange \mathbf{M}_λ , dépendant non linéairement d'un faible nombre de paramètres. Dans le cas où les positions des étoiles $\{z_i\}_{i \in [1; N_s]}$ et les paramètres de Moffat α_λ et β_λ (et donc la matrice \mathbf{M}_λ) sont connus, nous pouvons séparer les spectres par une simple estimation par maximum de vraisemblance (moindres carrés pondérés) :

$$\hat{\mathbf{S}}_\lambda = \Sigma_\lambda \mathbf{M}_\lambda^T \Sigma_{\epsilon_\lambda}^{-1} \mathbf{Y}_\lambda, \quad \text{avec} \quad \Sigma_\lambda = (\mathbf{M}_\lambda^T \Sigma_{\epsilon_\lambda}^{-1} \mathbf{M}_\lambda)^{-1} \quad (3)$$

où $\Sigma_{\epsilon_\lambda} = \text{diag}(\{\hat{\sigma}_\lambda^2[p]\}_p)$.

Mais lorsque ces paramètres sont inconnus, leur estimation conjointe, au sens du maximum de vraisemblance (MV), est délicate à cause de la multi-modalité de la vraisemblance liée à la non linéarité du modèle.

En abordant ce problème de SAS sans prendre en compte la structure de la matrice, par exemple avec une approche de type FMN, les résultats ne sont pas très satisfaisants [6, p.139]. L'estimation par MV, en alternant l'optimisation sur les spectres et sur les autres paramètres, a été proposée dans [6] aux positions d'étoiles fixées et dans [8] en estimant également les positions des étoiles. Cependant, ces méthodes sont sensibles à l'initialisation de l'algorithme.

2.3 Informations a priori

Les paramètres de la PSF et les positions des étoiles ne peuvent être connus parfaitement, mais on peut disposer d'information sur ces paramètres. Ainsi, l'observation de ce même champ stellaire par un imageur à haute résolution spatiale tel que le télescope spatial Hubble, permettra d'obtenir une plage de valeurs pour les positions des étoiles $z_i \in [\bar{z}_i - \Delta_z, \bar{z}_i + \Delta_z]$. De même, l'analyse des conditions d'observation (*seeing*) permettra d'obtenir des plages de valeurs pour les paramètres de la PSF $\alpha_\lambda \in [\alpha_{\min}, \alpha_{\max}]$ et $\beta_\lambda \in [\beta_{\min}, \beta_{\max}]$. Il est donc possible de prendre en compte, pour ces paramètres, des lois a priori indépendantes uniformes sur ces intervalles. En revanche, nous ne disposons pas d'information a priori particulière² sur les spectres des étoiles ce qui revient à considérer des lois a priori (dégénérées) uniformes sur \mathbb{R} .

3 Filtre particulière

Un filtre particulière vise à approximer, par un ensemble de N particules $(\xi_\lambda^{(q)}[p])_{q \in [1; N]}$ et leurs poids associés $(w^{(q)}[p])_{q \in [1; N]}$, la loi a posteriori $f_{\mathbf{x}_\lambda[p]|\mathbf{y}_\lambda[1:p]}$ d'un vecteur de paramètres cachés $\mathbf{x}_\lambda[p]$ conditionnellement à des observations $\mathbf{y}_\lambda[1:p] = \{y_\lambda[k]\}_{k \in [1;p]}$. Il procède par itération sur les données, à partir d'une initialisation aléatoire des particules suivant la loi a priori sur $\mathbf{x}_\lambda[0]$, en propageant les particules, prenant en compte à chaque itération p l'observation $y_\lambda[p]$. Pour cela, le FP s'appuie sur la loi d'observation $f_{y_\lambda[p]|\mathbf{x}_\lambda[p]}$ et sur la loi d'évolution $f_{\mathbf{x}_\lambda[p+1]|\mathbf{x}_\lambda[p]}$. A partir de cette approximation, il est aisé d'estimer des caractéristiques sur les paramètres du vecteur caché $\mathbf{x}_\lambda[p]$ tels que leur moyenne et variance, par simples moyennes pondérées sur les particules.

Pour notre problème, le vecteur des paramètres à estimer est $\mathbf{x}_\lambda[p] = [\alpha_\lambda[p], \beta_\lambda[p], \{z_i[p], s_{i,\lambda}[p]\}_{i \in [1; N_s]}]$. D'après les hypothèses de la § 2, la loi d'observation $f_{y_\lambda[p]|\mathbf{x}_\lambda[p]}$ est gaussienne $\mathcal{N}(g(\mathbf{x}_\lambda[p]), \hat{\sigma}_\lambda^2[p])$ où g est la fonction qui lie $y_\lambda[p]$ à $\mathbf{x}_\lambda[p]$ (Eq. 1 et Eq. 2). $\mathbf{x}_\lambda[p]$ étant constant par rapport aux pixels p , la loi d'évolution est une loi déterministe. Aussi, un simple filtre particulière de type bootstrap ne fonctionnera pas correctement à cause de phénomènes de dégénérescence des particules. En effet, comme le bruit d'évolution est inexistant, les redistributions répétées appauvrissent de plus en plus le nuage de

² En dehors de leur positivité qu'il est inutile de prendre en compte ici, les spectres estimés étant toujours positifs en pratique.

particules, jusqu'à le réduire à un seul point. Pour éviter ce phénomène, une solution consiste à ajouter un processus stochastique artificiel à la loi d'évolution, sous contrainte de ne pas perdre d'information en terme de moyenne et variance dans l'étape d'évolution [10, p. 202]. On utilise en pratique une loi d'évolution $f_{\mathbf{x}_\lambda[p+1]|\mathbf{x}_\lambda[p]}$ gaussienne :

$$\mathcal{N}(a\mathbf{x}_\lambda[p] + (1-a)\overline{\mathbf{x}_\lambda[p]}, (1-a^2)V[p]) \quad (4)$$

où a est un paramètre à régler ($a < 1$) et $\overline{\mathbf{x}_\lambda[p]}$ et $\mathbf{V}_\lambda[p]$ sont respectivement la moyenne et la variance a posteriori du vecteur caché \mathbf{x}_λ au pixel p .

Les différentes étapes de l'algorithme de FP sont détaillées dans [11]. Pour chaque image à la longueur d'onde λ , le FP donne une approximation de $f_{\mathbf{x}_\lambda[N_p]|\mathbf{y}_\lambda[1:N_p]}$ (la loi a posteriori) définie par un ensemble de particule $\boldsymbol{\xi}_\lambda^{(q)}[N_p]$ pondérées par les poids $w^{(q)}[N_p]$. A partir de cette approximation particulière, on estime les paramètres recherchés au sens de la moyenne a posteriori :

$$\widehat{\mathbf{x}}_\lambda = \sum_{k=1}^N w^{(k)}[N_p] \boldsymbol{\xi}_\lambda^{(k)}[N_p]$$

avec pour estimateur de matrice de covariance associé :

$$\text{Cov}(\widehat{\mathbf{x}}_\lambda) = \sum_{k=1}^N w^{(k)}[N_p] (\boldsymbol{\xi}_\lambda^{(k)}[N_p] - \widehat{\mathbf{x}}_\lambda)(\boldsymbol{\xi}_\lambda^{(k)}[N_p] - \widehat{\mathbf{x}}_\lambda)^T$$

4 Méthode MCMC

L'algorithme de Metropolis-Hastings (MH) est une méthode de Monte-Carlo par Chaîne de Markov qui produit N_e échantillons distribués suivant une chaîne de Markov dont la distribution tend asymptotiquement vers la loi cible. L'objectif est donc ici d'exploiter un tel algorithme afin de générer des échantillons du vecteur de paramètres \mathbf{x}_λ suivant la loi a posteriori $f_{\mathbf{x}_\lambda|\mathbf{y}_\lambda}$ pour toutes les données observées. L'estimateur de la moyenne a posteriori et la variance associée se calculent alors simplement par moyenne empirique sur les échantillons générés (en supprimant une première moitié des échantillons sensibles à l'initialisation de la chaîne).

L'algorithme MH est relativement générique et simple à appliquer à notre problème [9]. Cependant, afin de réduire la dimension de l'espace des paramètres, nous proposons ici de générer les échantillons suivant la loi a posteriori marginale $f_{\boldsymbol{\zeta}_\lambda|\mathbf{y}_\lambda}$ des paramètres $\boldsymbol{\zeta}_\lambda = [\alpha_\lambda, \beta_\lambda, \{z_i\}_{i \in [1;N_s]}]$. En effet, pour une vraisemblance gaussienne et un modèle linéaire en les spectres des étoiles $\{s_{i,\lambda}\}_{i \in [1;N_s]}$, l'intégration de la loi a posteriori $f_{\mathbf{x}_\lambda|\mathbf{y}_\lambda}$ par rapport à ces paramètres se fait aisément. Afin d'explorer au maximum l'espace des paramètres tout en ayant un taux d'acceptation assez élevé, nous exploitons pour loi de proposition un mélange entre une loi uniforme (à 10%) sur les intervalles a priori pour les paramètres et une marche aléatoire gaussienne (à 90%) de faible écart-type (1/100ème de la largeur de l'intervalle a priori). Une fois la position des étoiles et les paramètres $(\alpha_\lambda, \beta_\lambda)$ estimés grâce aux

échantillons de la chaîne, et donc la matrice \mathbf{M}_λ connue, les spectres $\{s_{i,\lambda}\}_{i \in [1;N_s]}$, sont estimés par maximum de vraisemblance (Eq. 3).

5 Résultats de simulations

Les données simulées correspondent à des images de 16×16 pixels à 3400 longueurs d'onde d'un champ contenant $N_s = 11$ étoiles de spectres réalistes (cf. Fig. 1). Les paramètres de la PSF varient linéairement en λ . Suivant le modèle décrit par l'Eq. (1), nous générons des données non bruitées $y_\lambda^{\text{true}}[p]$, auxquelles nous ajoutons du bruit additif gaussien centré indépendant. Selon un modèle d'observation réaliste [7], le bruit a pour variance $\sigma_\lambda^2[p] = y_\lambda^{\text{true}}[p] + s_{\text{sky},\lambda}$ où $s_{\text{sky},\lambda}$ est un fond constant correspondant au spectre de ciel, supposé connu. Notons qu'en raison de raies d'émission du spectre de ciel, le rapport signal sur bruit (RSB) des données est très faible à certaines longueurs d'onde. De telles simulations sont illustrées Fig. 1. Pour l'estimation, la variance du bruit est estimée par $\widehat{\sigma}_\lambda^2[p] = y_\lambda[p] + s_{\text{sky},\lambda}$, les positions des étoiles (constantes en λ) sont supposées connues à plus ou moins $\Delta_z = \frac{1}{5}$ pixel et $\alpha \in [1; 2]$ et $\beta \in [1, 5; 2, 5]$.

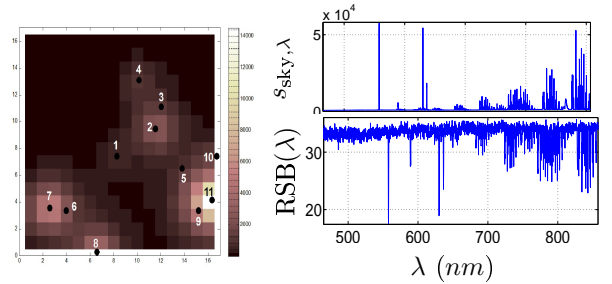


FIGURE 1 – Image simulée et positions des étoiles; spectre du ciel (en photons) et RSB en dB fonction de λ .

Afin de réduire le coût de calcul, les deux méthodes sont utilisées sur les 20 premières longueurs d'onde afin d'obtenir une estimation suffisamment précise de la position des étoiles que l'on fixera par la suite pour l'estimation des paramètres de la PSF et des spectres des étoiles. Notons ici que pour utiliser le FP il faut fixer les paramètres (a, N) et pour MH il faut fixer le nombre d'échantillons N_e (en plus de choisir la loi de proposition). Pour un bon compromis entre précision et temps de calcul, les paramètres du FP sont fixés à $N = 2500$ particules et $a = 0.98$ et pour MH à $N_e = 5.10^5$ échantillons. A titre d'illustration des résultats, les valeurs absolues d'erreur d'estimation de la position en ordonnée des 6^{ème} et 7^{ème} étoiles (deux étoiles proches une brillante et l'autre faible) sont illustrées sur la Fig. 2.

Ensuite, nous avons estimé les paramètres de la PSF ainsi que les spectres stellaires pour toutes les longueurs d'onde λ (conditionnellement aux positions d'étoiles fixées à $\{\widehat{z}_i\}_{i \in [1;11]}$). On a fixé ici le nombre de particules du FP à $N = 1500$ et pour MH à $N_e = 2000$ échantillons.

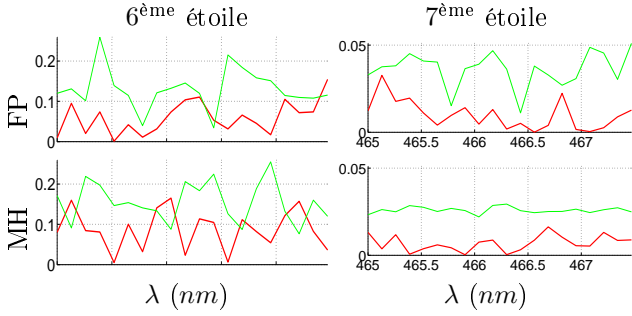


FIGURE 2 – Valeur absolue d’erreur d’estimation de la position en ordonnée des 6^{ème} et 7^{ème} étoiles ; moyenne en rouge et enveloppe de confiance à 3σ en vert.

Les valeurs absolues des erreurs relatives d’estimation pour α_λ et β_λ ainsi que sur les spectres de la 6^{ème} et de la 7^{ème} étoiles sont représentées en Fig. 3.

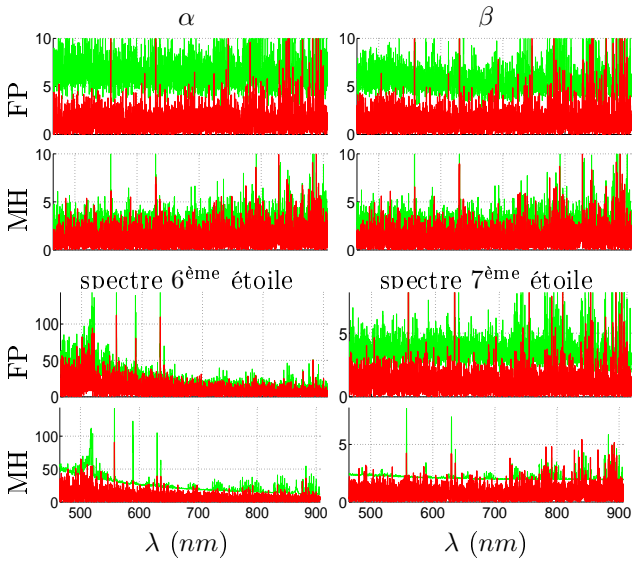


FIGURE 3 – Valeur absolue de l’erreur relative d’estimation (en %) des paramètres α_λ et β_λ et des spectres de la 6^{ème} et la 7^{ème} étoiles (mêmes couleurs que Fig. 2).

6 Discussion et Conclusion

En comparant les résultats de deux méthodes, on constate qu’en localisation des étoiles, le FP et MH donnent des résultats similaires, en moyenne sur les 11 étoiles on obtient une erreur (en pixel) de 0.015 pour le FP et de 0.02 pour MH.

Concernant l’estimation des paramètres de structure du mélange ($\alpha_\lambda, \beta_\lambda$), on constate qu’en moyenne le FP et MH donnent des résultats similaires pour $(\hat{\alpha}_\lambda, \hat{\beta}_\lambda)$, mais que l’enveloppe de confiance à 3σ de MH est divisé par 2 par rapport à celle du FP. On pourrait penser que cette augmentation de variance est due à la prise en compte d’une loi d’évolution artificielle, mais en augmentant le nombre de particules N et en faisant tendre a vers 1, en moyenne le FP donne toujours une enveloppe de confiance 2 fois plus grand que celle de MH. Or, d’un point de vue théorique on s’attend à obtenir des résultats équivalents.

Pour l’estimation des spectres, les estimations fournies par le FP de $\{\hat{s}_{i,\lambda}\}_{i \in [1;11]}$ et des enveloppes de confiance à 3σ semblent plus variables que celles de MH. Cela est dû au mode d’estimation des spectres par la méthode proposée pour MH, puisque l’estimation se fait par maximum de vraisemblance, une fois les autres paramètres estimés par moyenne a posteriori sur la loi marginalisée pour les spectres, tandis que pour le FP l’estimation est effectuée à partir de la loi a posteriori des spectres $f_{\{s_{i,\lambda}\}_{i \in [1;11]}} \mathbf{y}_\lambda$. Néanmoins, si l’on fait la moyenne sur l’ensemble des longueurs d’onde, les moyennes et enveloppes estimées sont du même ordre de grandeur. Enfin, rien n’empêcherait d’estimer les moyennes et enveloppes à 3σ des spectres de la même façon pour le FP que pour MH, ce qui donnerait des résultats similaires.

Pour améliorer les résultats, on souhaite à l’avenir prendre en compte la corrélation des paramètres entre longueurs d’onde, en particulier en ce qui concerne les paramètres de la structure de mélange, ces paramètres évoluant lentement en longueur d’onde.

Références

- [1] C. Baccigalupi, et al, "Extracting cosmic microwave background polarisation from satellite astrophysical maps", in Monthly Notices of the Royal Astronomical Society, 354 :55-70, 2004.
- [2] O. Berné, et al, "Analysis of the emission of very small dust particles from Spitzer spectro-imagery data using Blind Signal Separation methods", in Astronomy and Astrophysics, 469, pp. 575-586, 2007.
- [3] E. Kuruoglu, "Bayesian source separation for cosmology", in IEEE Signal Processing Mag., vol. 27, pp. 43-54, 2010.
- [4] M. Costagli, et al, "Astrophysical Source Separation Using Particle Filters", in Independent Component Analysis and Blind Signal Separation, Lecture Notes in Computer Science Vol. 3195, pp 930-937, 2004.
- [5] R. Bacon, et al., *The MUSE second-generation VLT instrument*, SPIE Conf. Series, vol. 7735, July 2010.
- [6] I. Meganem, "Méthodes de Séparation Aveugle de Sources pour l’imagerie hyperspectrale. Application à la télédétection urbaine et à l’astrophysique", Thèse de doctorat, Université Toulouse III - Paul Sabatier, décembre 2012.
- [7] E. Villeneuve, et al, "PSF Estimation of hyperspectral data acquisition system for ground-based astrophysical observations", IEEE WHISPERS, Lisbon, Portugal, 2011.
- [8] S. Kamann, et al, "Resolving stellar populations with crowded field 3D spectroscopy", in Astronomy and Astrophysics, Vol. 549, 2013.
- [9] W. R. Gilks, et al, "Markov Chain Monte Carlo in Practice," Chapman & Hall, London, UK, 1996.
- [10] A. Doucet, et al, "Sequential Monte Carlo Methods in Practice," Springer, New York, 2001.
- [11] A. Selloum, Y. Deville, H. Carfantan, "Separation of stellar spectra from hyperspectral images using particle filtering constrained by a parametric spatial mixing model," in IEEE ECMSM, Toulouse, France, 2013.